University of Warsaw

Noise and signal correlations in quantum metrology and imaging

Stanisław Kurdziałek

Doctoral thesis written under the supervision of
prof. dr hab. Rafał Demkowicz-Dobrzański

February 2026

Abstract

Quantum metrology exploits non-classical phenomena, such as coherence and entanglement, to enhance measurement precision. This thesis advances theoretical metrology in two different directions. First, I develop general theoretical tools to study fundamental metrological limits in the relatively unexplored regime of correlated signal and noise. Second, I focus on quantum-inspired superresolution imaging and demonstrate how temporal optical signal correlations serve as a valuable metrological resource.

The first part (Chapters 2–4) develops techniques for quantum Fisher information (QFI) optimization over general adaptive metrological protocols in the presence of correlated noise. Exact solutions, demonstrated in previous works, are feasible only for small channel numbers (typically $N \leq 5$). In this thesis, the large-N limit is explored, for which lower and upper bounds for the optimal QFI are derived.

Chapter 2 presents tensor-network-based methods that handle adaptive protocols under ancilla dimension constraints. Applicable to both correlated and uncorrelated scenarios, these methods provide constructive lower bounds for significantly larger systems than previously possible.

Chapter 3 derives universal upper bounds for adaptive QFI in uncorrelated models. The derived bounds are proven to be asymptotically tight. This enables the proof of the general asymptotic equivalence between the metrological performance of parallel and adaptive schemes, which was a long-standing open question.

Chapter 4 extends the bounds derived in Chapter 3 to the fully general case involving signal and noise correlations. This allows us to study two canonical physical cases involving correlations—phase estimation in the presence of correlated dephasing noise, and collisional quantum thermometry.

Chapter 5 introduces a universal measure of the vulnerability of quantum estimation precision to measurement imperfections, which we call Fisher information measurement noise susceptibility (FI MeNoS). We demonstrate the usefulness of this quantity in analyzing interferometry and superresolution imaging scenarios. We show that sub-Poissonian photon statistics, arising from antibunching, enable to construct a less noise-susceptible measurements for a binary source separation estimation.

Chapter 6 shows how temporal correlations of optical signal resulting from emitter blinking enhance quantum-inspired superresolution. We consider estimation of spatial moments for arbitrary objects. Correlations are shown to enhance precision of higher

spatial moments estimation and, more importantly, to enable extraction of the information about these higher moments using significantly simpler measurements.

Streszczenie

Metrologia kwantowa wykorzystuje nieklasyczne zjawiska, takie jak koherencja i splątanie kwantowe, do zwiększania precyzji pomiarów. Wyniki zawarte w niniejszej rozprawie odnoszą się do dwóch różnych kierunków badań w tej dziedzinie. W pierwszej części rozprawy opracowuję ogólne narzędzia teoretyczne do badania fundamentalnych ograniczeń metrologicznych w dotychczas słabo zbadanym reżimie skorelowanego sygnału i szumu. Następnie, skupiam się na kwantowo inspirowanym obrazowaniu nadrozdzielczym i pokazuję, że czasowe korelacje sygnału stanowią metrologicznie wartościowy zasób.

W pierwszej części rozprawy (rozdziały 2–4) rozwijane są uniwersalne techniki optymalizacji kwantowej informacji Fishera (QFI) po protokołach adaptatywnych w obecności skorelowanego szumu. Dokładne wykonanie tej optymalizacji jest możliwe jedynie w przypadku małej liczby estymowanych kanałów (typowo $N \leq 5$). Celem tej rozprawy jest zbadanie obszaru dużych N, w którym możliwe jest jedynie wyprowadzenie dolnych i górnych ograniczeń na wynik optymalizacji.

Rozdział 2 przedstawia metody oparte na sieciach tensorowych, pozwalających efektywnie optymalizować protokoły adaptatywne przy dodatkowym ograniczeniu rozmiaru układu pomocniczego (*ancilli*). Metody te, mające zastosowanie zarówno do scenariuszy skorelowanych, jak i nieskorelowanych, pozwalają uzyskać konstruktywne dolne ograniczenia kwantowej informacji Fishera dla układów, których rozmiar znacznie przekracza możliwości wcześniejszych algorytmów.

W rozdziale 3 wyprowadzone są uniwersalne górne ograniczenia adaptatywnej kwantowej informacji Fishera dla przypadku modeli nieskorelowanych. Wyprowadzone ograniczenia są asymptotycznie wysycalne. Pozwala to wykazać ogólną asymptotyczną równoważność metrologicznych schematów równoległych i adaptatywnych, co przez dłuższy czas stanowiło otwarty problem.

Rozdział 4 rozszerza ograniczenia wyprowadzone w Rozdziale 3 do w pełni ogólnego przypadku obejmującego korelacje sygnału i szumu. Pozwala to zbadać dwa kanoniczne fizyczne przypadki, w których kluczowa jest rola korelacji – estymację fazy w obecności skorelowanego szumu defazującego oraz kolizyjną termometrię kwantową.

Rozdział 5 wprowadza uniwersalną miarę wrażliwości precyzji estymacji kwantowej na niedoskonałości pomiarowe, którą nazywamy podatnością informacji Fishera na szum pomiarowy (*Fisher information measurement noise susceptibility*, FI MeNoS). Wielkość ta okazuje się być użyteczna do analizy kwantowej interferometrii i kwantowo-

inspirowanego obrazowania nadrozdzielczego. Pokazuję, że sub-Poissonowska statystyka fotonów, wynikająca ze zjawiska antygrupowania, umożliwia skonstruowanie pomiarów mniej podatnych na szum w przypadku estymacji rozsunięcia dwóch obrazowanych źródeł.

Rozdział 6 pokazuje, jak czasowe korelacje sygnału optycznego wynikające z migania emiterów poprawiają kwantowo inspirowaną nadrozdzielczość. Rozważam estymację momentów przestrzennych obiektów o dowolnym kształcie. Pokazuję, że korelacje zwiększają precyzję estymacji wyższych momentów przestrzennych oraz, co ważniejsze, umożliwiają uzyskanie informacji o tych wyższych momentach przy użyciu znacznie prostszych pomiarów.

Podziękowania

W pierwszej kolejności chciałbym podziękować mojemu promotorowi, Rafałowi Demkowiczowi-Dobrzańskiemu, który kształtował mnie jako naukowca przez ostatnich 8 lat. Dziękuję za niezliczone inspirujące rozmowy naukowe i pozanaukowe, za danie mi wolności podążania za moimi pomysłami już na bardzo wczesnym etapie kariery badawczej, ale też za motywację, dzięki której pomysły zmieniały się w publikacje w skończonym czasie.

W czasie mojej akademickiej ścieżki napotkałem również innych życzliwych mentorów. Dziękuję Radkowi Łapkiewiczowi za to, że w jego laboratorium miałem okazję *dotknąć* interferometru i mikroskopu, a nie tylko rozważać te układy teoretycznie. Nabyte tam doświadczenie stanowiło wartościową inspirację moich dalszych badań. Dziękuję również Konradowi Banaszkowi za jego wsparcie moich badań dotyczących teorii mikroskopii nadrozdzielczej. Dziękuję Jankowi Kołodyńskiemu za rozmowy i porady dotyczące ścieżki naukowej.

Chciałbym również podziękować wszystkim tym z którymi współpracowałem i od których wiele się nauczyłem. Wojtek Górecki i Francesco Albarelli towarzyszyli mi i służyli pomocą, gdy stawiałem pierwsze kroki w dziedzinie teorii kwantowej metrologii. Moja współpraca z Piotrem Dulianem rozpoczęła się nieco później, jednak dzięki niemu postawiłem pierwsze pół kroku w dziedzinie programowania nie-jak-fizyk. Wspólnie ze mną meandry kwantowej metrologii i optyki na Wydziale Fizyki UW zgłębiali również Arek Kobus, Joanna Majsak, Staszek Sieniawski, Srijon Gosh, Arpan Das, Sagnik Chakraborty, Giovanni Scala, Sanjukta Kundu oraz Jerzy Szuniewicz, którym wyrażam wdzięczność za ciekawe dyskusje i owocne interakcje.

Dziękuję też kolegom ze studiów Staśkowi Żukowskiemu, Krzysztofowi Mękale, Mateuszowi Homendzie i Ignacemu Nałęczowi – nasze wspólne zmagania z sesją przerodziły się obecnie we wspólne zmagania z doktoratem. Dziękuję Jankowi Kołodziejczykowi, Grześkowi Jaczewskiemu, Teresie Kulce i Wojtkowi Góreckiemu za wspólne tworzenie atmosfery pokoju 5.28 na Wydziale Fizyki.

Rozprawa ta nie powstałaby bez bliskich mi osób spoza środowiska naukowego, poczynając od rodziców, którzy zawsze wspierali mnie w mojej pasji i we wszystkich moich wyborach. Chciałbym też serdecznie podziękować mojej dziewczynie Gabrysi za wsparcie mentalne i duchowe oraz przeżyte razem piękne chwile. Dziękuję też mojej siostrze Ewie za wsparcie w czasie pisania tej rozprawy.

Nie związałbym swojej kariery zawodowej z fizyką, gdy nie pan Piotr Kononowicz,

który uczył mnie w czasach gimnazjum i liceum. Powiedzieć, że uczył mnie fizyki, to jak nic nie powiedzieć – był dla mnie przez te lata mentorem, który pokazał mi prawdziwe piękno i głębię nauki.

Manuscripts comprising the thesis

The thesis contains the following publications:

[I] Quantum metrology using quantum combs and tensor network formalism
Stanisław Kurdziałek, Piotr Dulian, Joanna Majsak, Sagnik Chakraborty, and Rafał Demkowicz-Dobrzański
New Journal of Physics **27**, 013019 (2025)
DOI: 10.1088/1367-2630/ada8d1

[II] Using Adaptiveness and Causal Superpositions Against Noise in Quantum Metrology
Stanisław Kurdziałek, Wojciech Górecki, Francesco Albarelli, and Rafał Demkowicz-Dobrzański
Physical Review Letters **131**, 090801 (2023)
DOI: 10.1103/PhysRevLett.131.090801

[III] Universal Bounds for Quantum Metrology in the Presence of Correlated Noise
Stanisław Kurdziałek, Francesco Albarelli, and Rafał Demkowicz-Dobrzański
Physical Review Letters **135**, 130801 (2025)
DOI: 10.1103/jy3v-wkcb

[IV] Measurement Noise Susceptibility in Quantum Estimation
Stanisław Kurdziałek and Rafał Demkowicz-Dobrzański
Physical Review Letters **130**, 160802 (2023)
DOI: 10.1103/PhysRevLett.130.160802

[V] Super-resolution microscopy via fluctuation-enhanced spatial mode demultiplexing
Stanisław Kurdziałek
arXiv preprint: 2511.20790 (2025)

The publications are presented in Chapters 2–6 of this thesis, respectively.

Contents

Acronyms

CRB Cramér-Rao Bound.

DI Direct Imaging.

FI Fisher Information.

FI MeNoS Fisher Information Measurement Noise Susceptibility.

GKSL Gorini-Kossakowski-Sudarshan-Lindblad.

H-G Hermite-Gaussian.

HS Heisenberg Scaling.

III Image Inversion Interferometry.

ISS Iterative See-Saw.

MOP Minimization Over Purifications.

MPS Matrix Product State.

MSE Mean Squared Error.

QCRB Quantum Cramér-Rao Bound.

QFI Quantum Fisher Information.

SDP Semidefinite Program.

SS Standard Scaling.

Chapter 1

Introduction

1.1 Motivation and structure

The ultimate goal of *quantum metrology* is to exploit quantum effects in order to enhance the precision of measurements of physical quantities [Dem15; Deg17; Pir18; Pez18; Pol20; Bar22]. Microscopic probes—such as photons, atoms, or molecules—serve as sensitive detectors of parameters including, for example, time, length, magnetic and gravitational fields, etc.

Quantum metrology is currently among the most promising areas of quantum technologies [Dow03; Ací18]. A prominent illustration is the LIGO gravitational-wave detector [Abb16], in which squeezed states of light are employed to enhance interferometric measurement precision beyond the classical limit [Aas13; Tse19]. Remarkably, it has been shown that, in the presence of photon losses, squeezed states are the optimal choice for phase sensing among *all* states of light allowed by quantum mechanics [Cav81; Dem13], even the most exotic ones! In this sense, the success of squeezed-light metrology can be viewed as a very rare, fortunate coincidence: squeezed states belong to the very narrow subset of non-classical states of light that can be generated reliably in experimental settings, even for relatively large number of photons involved.

Other applications of quantum metrology include ultra-precise time measurements via atomic and optical clocks [Kat11; Hin13; Lud15; Bel21], atomic [Bud07; Sew12; Hor12; Bra15] and nitrogen-vacancy (NV) centers [Maz08; Aco09; Ron14; Bar20; Du24] magnetometry and spectroscopy, quantum-enhanced gravimetry [Szi20], and super-resolution optical imaging [Tsa16; Mor19]. In numerous of these applications, the foundational principles are well-known, but connecting fundamental predictions with practical applications still remains a challenge. One can distinguish different research directions aiming to take up this challenge and push technology to theoretical limits.

First, experimental works aim to implement increasingly broad class of metrological protocols. This is typically achieved by reducing noise, enhancing the fidelity of preparation of metrologically useful states, improving coherence times of quantum

probes, etc.

This direction is extensively supported by close-to-experiment theoretical works, whose goal is to thoroughly analyze specific laboratory setups, and identify practical routes to achieving quantum advantage. Such analyses usually rely on the theory of open quantum systems [Gor76; Lin76; Bre02] in order to properly account for decoherence effects [Zur03].

Finally, a complementary line of theoretical research focuses on establishing fundamental precision bounds and on developing general numerical or analytical tools for identifying optimal metrological protocols. Rather than targeting specific quantum systems, these works aim to advance methodologies applicable across a broad range of experimental platforms. The current challenge is to extend the range of applicability of these fundamental techniques, so they take into account complex structure of noise, wide range of possible metrological protocols, and potentially limited access to quantum resources.

The core of the research presented in this thesis (Chapters 2–4) belongs to this latter category. The primary—though not exclusive—aim of the thesis is to develop general theoretical tools suitable for analyzing scenarios involving *correlated signal or noise*. Such scenarios are commonly present in practice, and have been analyzed in various case-studies [Tsa11; Jes14; Bea18; Sea19; Cza19; Rib22; Pla22; Yan24]. Nevertheless, universal bounds and algorithms to optimize metrological protocols in the presence of correlations have been missing for a long time. The presented research fills this gap, which significantly broadens the applicability of universal metrological tools.

The work presented in Chapter 5 can also be classified as extending the general toolbox of quantum metrology. However, this study was inspired by a particular application—super-resolution imaging. Using introduced methodology, we observe that signal correlations arising from non-Poissonian light statistics can enhance quantum-inspired superresolution imaging.

Chapter 6 develops this concept further through close-to-experiment theoretical work that proposes practical improvements to super-resolution imaging techniques based on signal correlations resulting from emitters blinking.

In what follows, we provide a more detailed description of the thesis structure.

Chapter 1 introduces the general classification of quantum metrological schemes (Sec. 1.2), discusses the mathematical formulation of metrological protocol optimization (Sec. 1.3), and provides a brief overview of the most important results in the field (Sec. 1.4). This allows the results presented in Chapters 2–4 to be placed in the context of state-of-the-art research. The final section of this chapter (Sec. 1.5) introduces the concept of quantum-inspired super-resolution imaging and discusses the impact of signal correlations caused by non-Poissonian light statistics.

Chapter 2 contains the paper *Quantum metrology using quantum combs and tensor network formalism* [Kur25c] published in *New Journal of Physics*. This work provides efficient tensor-network based tools for optimizing adaptive metrological protocols in the limit of large numbers of channel uses. Importantly, the method enables studying

scenarios with correlated noise and signal. The paper also includes an extended theoretical introduction to concepts used throughout Chapters 2–4.

Chapter 3 contains the paper *Using adaptiveness and causal superpositions against noise in quantum metrology* [Kur23b] published in *Physical Review Letters*. This work derives universal bounds for the estimation precision of adaptive metrological protocols. The asymptotic tightness of these bounds enables us to prove the asymptotic equivalence between parallel and adaptive quantum metrological protocols, which was a long-standing open problem. Although the results are limited to uncorrelated noise, the developed tools are crucial for constructing bounds involving correlation effects.

Chapter 4 contains the paper *Universal bounds for quantum metrology in the presence of correlated noise* [Kur25b] published in *Physical Review Letters*. In this work, universal metrological precision bounds valid for noise and signal correlations of any type are derived. Two canonical models involving correlations are then studied—phase estimation in the presence of correlated dephasing and quantum collisional thermometry.

Chapter 5 contains the paper *Measurement noise susceptibility in quantum estimation* [Kur23a] published in *Physical Review Letters*. In this work, we introduce Fisher information measurement noise susceptibility (FI MeNoS), which quantifies the potential loss of the estimation precision due to small measurement disturbance of any type. We then study the quantum-inspired super-resolution imaging case, and conclude, that signal correlations caused by quantum light statistics may lead to increase of measurement robustness against noise.

Chapter 6 contains the paper *Super-resolution microscopy via fluctuation-enhanced spatial mode demultiplexing* [Kur25a], currently available as an arXiv preprint. This is the most experimentally oriented work in this collection. It shows how to extract additional information from signal correlations caused by emitter blinking in techniques utilizing quantum-inspired optimal superresolving measurements.

Chapter 7 summarizes all the results and provides an outlook on future work.

Each of Chapters 2–6 begins with a short introduction to the presented paper, specifying my contribution to the work.

Appendix A describes my other research papers that are not included in this thesis.

1.2 Hierarchy of metrological protocols

Irrespectively of the specific quantity being measured or the nature of the quantum probes used, metrological scenarios can be classified according to two criteria: their *real-world faithfulness* and their *degree of quantumness* (see Fig. 1.1).

We begin by discussing the classification based on real-world faithfulness. The simplest metrological scenarios to analyze are those in which no noise is present. Although such situations are rarely present in practice, their study provided an important theoretical baseline for the whole field of quantum metrology [Gio06]. In

the noiseless case, the parameter to be estimated denoted by φ, is encoded in the probe state through a φ-dependent unitary channel U_φ, see Fig. 1.1(A1, B1, C1). We assume that φ is a single real number, and represents the only unknown degree of freedom of the examined system—all the results in this thesis are derived using this single-parameter setup.

A canonical noiseless scenario involves the estimation of the strength of a static magnetic field oriented along the z direction using a spin-1/2 probe. Then, the parameter is imprinted on the probe via the unitary evolution $U_\varphi = e^{-i\sigma_z\varphi/2}$, where σ_z denotes the corresponding Pauli matrix. Another widely studied physical case is phase estimation in an optical interferometer under the assumptions of phase stability and no photon losses.

In real experiments, however, the probes are affected by unwanted interactions with the environment, leading to decoherence [Zur03; Sut16]. As a consequence, the probe dynamics is usually non-unitary. To capture this more realistic level of description, one should model the probe evolution using a general, φ-dependent quantum channel Λ_φ, which is mathematically represented as a completely positive, trace preserving (CPTP) map [Nie10], see Fig. 1.1(A2, B2, C2).

This approach allows one to account for a wide range of experimentally relevant effects, such as photon losses in optical interferometers, dephasing or depolarization affecting spin probes, or amplitude damping of atomic probes caused by spontaneous photon emission [Dem12]. Although the environment plays a crucial role in these processes, it is not explicitly present in our description—channel Λ_φ represents the effective evolution of the probe when the environment is inaccessible. This greatly simplifies the whole analysis, especially when environmental system is large compared to the probe system. For example, the problem of magnetic field sensing with a spin-1/2 probe subject to dephasing noise can be formulated as the estimation of the parameter φ encoded in a channel

$$\Lambda_\varphi(\bullet) = \sum_{k=1}^{2} K_{k,\varphi} \bullet K_{k,\varphi}^\dagger, \quad K_{k,\varphi} = K_k e^{-i\sigma_z\varphi/2}, \quad K_1 = \sqrt{p}\mathbb{1}, \quad K_2 = \sqrt{1-p}\sigma_z, \tag{1.1}$$

where K_k are Kraus operators of dephasing, p parametrizes the dephasing strength; $p = 0.5$ corresponds to full dephasing, $p = 1$ corresponds to a noiseless case. Even though the dephasing process is caused by interactions with environment, we are only interested in its effect on a probe—such effect is described by the above qubit channel, and the environment does not have to be explicitly present in the description.

However, this simplification relies on an important assumption: each application of a channel Λ_φ needs to be an independent process, with its own, fresh environment.

Nevertheless, in multiple physically relevant situations, this assumption breaks down [Bre16; Veg17]. Therefore, to achieve the highest level of real-world faithfulness, we should consider that subsequent probes interact with the common environment, so the noise acting on different probes is correlated. Then, one needs to explicitly take into account the environment memory in the description of the probe evolution.

Therefore, one should consider channels Λ_φ acting both on a probe and environmental subsystems $\mathcal{E}$, as shown in Fig. 1.1(A3, B3, C3). The whole parameter decoding process is then a chain of channels, such that the environmental output of one channel is an environmental input of the next channel. Such a chain can be mathematically described as a *quantum comb* [Chi08; Chi09]—more details will be given in Chapter 2.

This framework allows to describe the broadest class of real-world effects involving noise and signal correlations of both classical and quantum origin. For example, one can consider the estimation of a constant component of a magnetic field with spin probes, that are additionally affected by fluctuating magnetic field. These fluctuations may be modeled as a stochastic process with memory, such that subsequent spins are subject to correlated rotations. For a single spin, such random rotations amount to simple dephasing noise. However, the description of a joint state of multiple rotated spins is more complex, since it must take into account their mutual correlations. More details about correlated dephasing model can be found in Chapters 2 and 4.

Let us now analyze another classification according to the *degree of quantumness* used to implement a given protocol.

The most rudimentary quantum metrological resource is *quantum coherence* [Str17], which is the ability of preparing probes in superposition states. This resource is inevitable in all phase estimation tasks—for example, when evolution is given by a unitary operation $U_\varphi = e^{-i\sigma_z\varphi/2}$, then eigenstates $|0\rangle, |1\rangle$ of a generator σ_z and their incoherent mixtures are invariant under the action of U_φ. Consequently, to estimate φ, one should prepare a probe in a coherent superposition of generator's eigenstates. Formally, the density matrix ρ describing the probe quantum state must have non-vanishing off-diagonal terms when expressed in the eigenbasis of the generator σ_z. In the simplest metrological protocols, coherence is the only quantum resource used, and independently prepared probes ρ interact with a measured system in parallel, see Fig. 1.1(A1, A2, A3). Thereafter, the measurement $\boldsymbol{M}$ is performed on probes to get information about the parameter φ. When we additionally assume that there are no signal or noise correlations, the precision of φ estimation is limited by the *shot noise*—the standard deviation of estimator $\tilde{\varphi}$ scales as $\Delta\tilde{\varphi} \sim 1/\sqrt{N}$, where N is the number of probes used. This is called a *standard scaling* (SS).

The estimation precision can be significantly increased if different probes are prepared in an *entangled* state [Hor09], see Fig. 1.1(B1, B2, B3). This extra degree of quantumness allows to beat the standard scaling. For example, when N probes are prepared in a so-called N00N state [Dow08], $\rho = |\psi\rangle\langle\psi|$, $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$, and each probe evolves unitarily via $U_\varphi = e^{-i\sigma_z\varphi/2}$, then it is possible to achieve a so-called *Heisenberg scaling* (HS) of estimation error, $\Delta\tilde{\varphi} \sim 1/N$. Unfortunately, N00N states are very fragile and thus not useful in the presence of noise. However, entanglement-based schemes can be beneficial in noisy scenarios—for example, squeezed states reduce the prefactor in the presence of dephasing or photon losses [Cav81; Dem12].

Another way to use entanglement in quantum metrology is to initially entangle probes with an ancillary system $\mathcal{A}$, which remains intact through the whole protocol, up to a final measurement $\boldsymbol{M}$. This means that ancillas do not interact with the measured

system and do not harvest information about φ—nevertheless, the correlations between probes and ancillas resulting from initial entanglement might be useful to increase the precision of estimation of φ in noisy scenarios [Fuj01].

In the schemes considered so far, probes evolve through different channels in *parallel*, implying that the information gathered by one probe could not influence the initial states of the others. To be more general, and exploit the potential of quantum mechanics further, one can consider *adaptive* protocols, in which different quantum channels are probed sequentially, and arbitrary control channels C_i can be applied between subsequent uses, see Fig. 1.1(C1, C2, C3). Importantly, the input probe state can be entangled with an arbitrarily large ancillary system $\mathcal{A}$. Therefore, parallel schemes constitute a subset of adaptive schemes—one can simulate a parallel scheme using adaptive setup with all probes included in an initial state, and choosing quantum controls C_i as appropriate SWAP operations, such that each probe is sent to a corresponding channel . The opposite is not true—adaptive schemes are more powerful than parallel, as it sometimes might be useful to send a single probe through the channels sequentially, to acquire the signal coherently—the role of control operations C_i and entangled ancillary system is then crucial to fight against noise using quantum error correction codes.

The distinction between three levels of quantumness—coherence, entanglement, quantum adaptiveness—can be also made for correlated noise cases. Then, however, one may sometimes expect scaling different than $\Delta\tilde{\varphi} \sim 1/\sqrt{N}$ even for independent probes, as output probe states may become correlated. One needs to also remember, that in such a case, the term *parallel* only refers to the probe preparation—the noise or signal correlations may be caused by a sequential process.

Interestingly, recent studies showed that quantum theory in principle allows for an even broader class of protocols involving superpositions of different causal orders of channel probing [Ara14; Gos18; Fre19; Wec21; Bav21; Roz24]. This establishes a *fourth* level of quantumness—*quantum non-causality*. While their experimental realizability remains unclear, they have been shown to be capable of providing metrological advantages in certain scenarios [Zha20; Cha21]. For now, we will neglect this theoretically possible quantum resource, but we discuss it further in Chapter 3.

To summarize, metrological schemes can be classified according to their real-world faithfulness (no noise, uncorrelated noise, or correlated noise) and their level of quantumness (coherence, entanglement, or adaptiveness), leading to the nine cases shown in Fig. 1.1. According to this classification, the most general class of metrological schemes—encompassing the widest range of real-world effects and quantum metrological tools—is the class of adaptive schemes with correlated noise, C3. All the remaining classes are subsets of this most general class. Consequently, any metrological upper bounds derived for class C3 are also valid for all nine cases shown in Fig. 1.1.

On the other hand, it is often useful to analyze these less general classes independently, either to account for limited access to quantum resources or to obtain tighter bounds by exploiting simplifying assumptions about the structure of the noise. As discussed in the next section, the historical development typically proceeded from less general

to more general schemes.

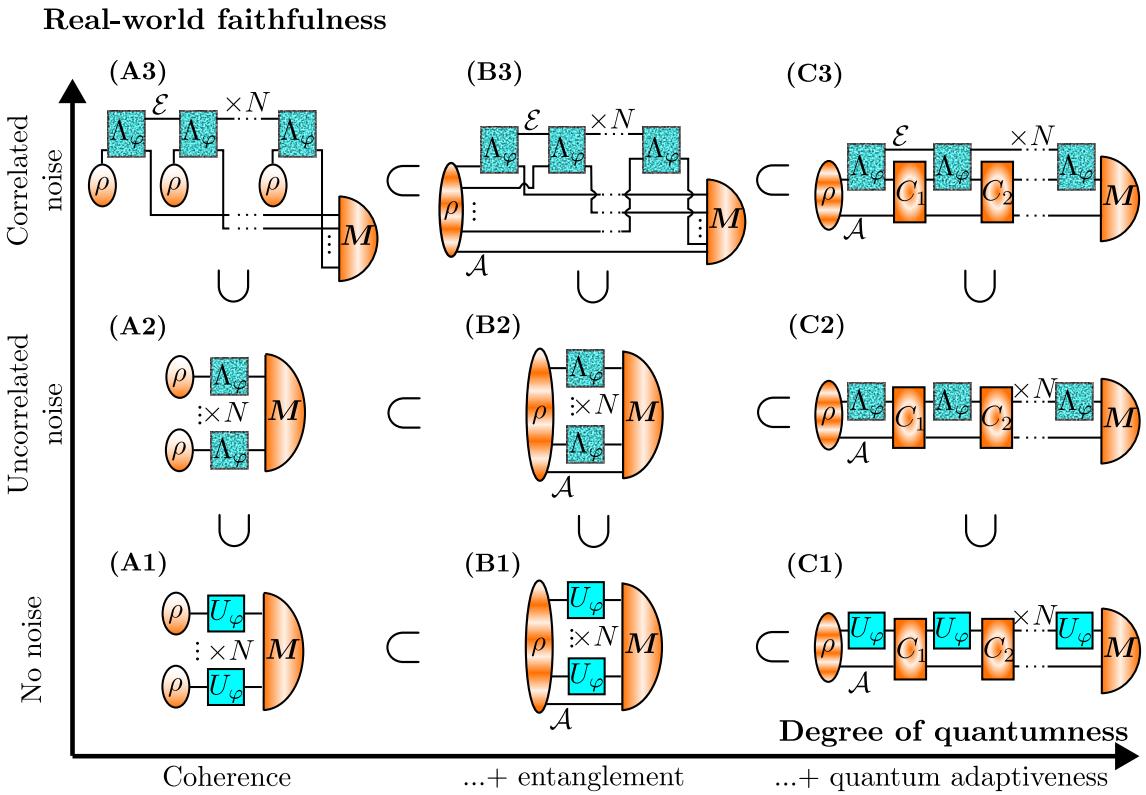


Figure 1.1: Metrological schemes can be classified according to their real-world faithfulness and degree of quantumness, resulting in a hierarchical classification with different levels of generality. Regarding degree of quantumness, the least general schemes (A1, A2, A3) rely solely on coherence of single probes; more general schemes (B1, B2, B3) additionally exploit entanglement between input probes; the most general schemes (C1, C2, C3) further incorporate adaptive quantum protocols. Regarding real-world faithfulness, the simplest schemes (A1, B1, C1) assume noiseless unitary parameter encoding; more realistic schemes (A2, B2, C2) incorporate uncorrelated noise; the most general schemes (A3, B3, C3) account for arbitrary noise and signal correlations. Class C3, the most general, encompasses all other presented schemes as subclasses.

1.3 Quantum Metrology as an optimization problem

In all the schemes considered, one can distinguish the following groups of components:

1. Fixed components—the channels U_φ or Λ_φ, which describe signal encoding and noise. We denote the collection of such components for a given protocol by $\mathcal{S}_\varphi$.

2. Adjustable, fully quantum components—input probes state ρ, and, in the case of adaptive protocols, the control operations C_i. We denote the set of these components for a given protocol by $\mathcal{C}$.

3. Adjustable measurement $\boldsymbol{M}$—a quantum-to-classical map that transforms a final quantum state into a classical outcome

The fundamental question of quantum metrology can be formulated as follows:

> *What is the optimal choice of adjustable components, and what is the resulting precision of estimation of φ?*

Let us now formulate this question as a well-defined mathematical problem. In each protocol, the probes evolve through φ-encoding channels, and, when applicable, through control operations. This results in a φ-dependent quantum state of all probes (and possibly of ancillary systems initially entangled with them), represented by a density matrix ρ_φ. To emphasize that ρ_φ depends on elements of the protocol belonging to both $\mathcal{C}$ and $\mathcal{S}_\varphi$, we write

$$\rho_\varphi = \mathcal{C} * \mathcal{S}_\varphi. \tag{1.2}$$

Intuitively, this notation indicates that ρ_φ is obtained by concatenating the inputs and outputs of elements from $\mathcal{C}$ with the corresponding outputs and inputs of elements from $\mathcal{S}_\varphi$. The strict mathematical definition of the operation $*$ (called a *link product* [Chi08]) will be given in Chapter 2, where we will also show how all elements of $\mathcal{C}$ and $\mathcal{S}_\varphi$ can be represented as single mathematical objects. For the moment, $\mathcal{C} * \mathcal{S}_\varphi$ denotes as-yet-undefined function that outputs density matrix ρ_φ for a specified protocol.

Information about φ is extracted from ρ_φ via a measurement $\boldsymbol{M}$, which returns a classical outcome x with a φ-dependent probability $p(x|\varphi)$. More precisely, $\boldsymbol{M} = \{M_x\}_{x\in X}$, where X is a set of all outcomes, operators M_x act on the same Hilbert space as ρ_φ, and form a positive operator-valued measure: the operators M_x are positive semidefinite and $\int_x \mathrm{d}x M_x = \mathbb{1}$. The conditional probability of an outcome x is then given by

$$p(x|\varphi) = \mathrm{Tr}(\rho_\varphi M_x). \tag{1.3}$$

After performing the measurement $\boldsymbol{M}$, the problem reduces to a classical task of estimation of the value of φ using an outcome x. In other words, we need to construct an estimator $\tilde{\varphi}(x)$, which is a function of classical outcome that is supposed to predict the value of φ as closely as possible.

There are two approaches to quantify the precision of estimation of φ.

In the *Bayesian* approach [Kay93], one needs to specify two elements: (i) a prior probability distribution $p_0(\varphi)$ representing initial knowledge about φ, and (ii) a cost function $C(\tilde{\varphi}(x), \varphi)$ quantifying the deviation of the estimator from the true value. The most common choice is a quadratic cost, $C(\tilde{\varphi}(x), \varphi) = (\tilde{\varphi}(x) - \varphi)^2$. The estimator precision is quantified by the average cost

$$\bar{C} = \int \mathrm{d}\varphi \int \mathrm{d}x p_0(\varphi) p(x|\varphi) C(\tilde{\varphi}(x), \varphi), \tag{1.4}$$

which is the cost function averaged over all values of φ, weighted by the prior distribution, and over all outcomes x, weighted by their conditional probabilities.

Let us define the *quantum cost* $C_Q(\mathcal{S}_\varphi)$, as the averaged cost minimized over all adjustable elements ($\mathcal{C}$, $\boldsymbol{M}$, $\tilde{\varphi}$) for a fixed signal encoding and noise. Taking into account (1.3), (1.4), we get

$$C_Q(\mathcal{S}_\varphi) = \min_{\{\mathcal{C},\boldsymbol{M},\tilde{\varphi}\}} \int \mathrm{d}\varphi \int \mathrm{d}x p_0(\varphi) \mathrm{Tr}(\rho_\varphi M_x) C(\tilde{\varphi}(x), \varphi), \quad \rho_\varphi = \mathcal{C} * \mathcal{S}_\varphi, \tag{1.5}$$

which is the formulation of a general quantum metrological problem in the Bayesian framework. The arguments ($\mathcal{C}$, $\boldsymbol{M}$, $\tilde{\varphi}$), for which the minimum is achieved, can be interpreted as optimal protocol, measurement and estimator respectively.

Let us now discuss an alternative, *frequentist* approach [Kay93]. We define the expected value of an estimator as

$$\mathbb{E}[\tilde{\varphi}|\varphi] = \int \mathrm{d}x p(x|\varphi)\tilde{\varphi}(x), \tag{1.6}$$

where the notation emphasizes the dependence on the true parameter's value φ. We call an estimator *globally unbiased* if its expected value always follows the estimated parameter,

$$\forall_\varphi \quad \mathbb{E}[\tilde{\varphi}|\varphi] = \varphi. \tag{1.7}$$

The estimator is *locally unbiased* around $\varphi = \varphi_0$ if

$$\mathbb{E}[\tilde{\varphi}|\varphi_0] = \varphi_0, \quad \left.\frac{\partial}{\partial\varphi}\right|_{\varphi=\varphi_0} \mathbb{E}[\tilde{\varphi}|\varphi] = 1, \tag{1.8}$$

so the estimator's expected value coincided with φ for $\varphi = \varphi_0 + \delta\varphi$ up to 1st order terms in $\delta\varphi$.

In the frequentist approach, we do not specify a prior distribution, but instead we focus on the estimation at the certain value of the parameter $\varphi = \varphi_0$, and optimize the mean-squared error (MSE), defined as

$$\Delta^2\tilde{\varphi} = \int \mathrm{d}x p(x|\varphi_0)(\tilde{\varphi}(x) - \varphi_0)^2, \tag{1.9}$$

over all estimators which are *locally unbiased* around $\varphi = \varphi_0$. Hence, we can define the frequentist cost function as a solution of the following optimization problem

$$C_{\mathrm{freq}} = \min_{\tilde{\varphi}} \Delta^2\tilde{\varphi} \quad \text{s.t.} \quad \mathbb{E}[\tilde{\varphi}|\varphi_0] = \varphi_0, \quad \left.\frac{\partial}{\partial\varphi}\right|_{\varphi=\varphi_0} \mathbb{E}[\tilde{\varphi}|\varphi] = 1. \tag{1.10}$$

This problem can be solved using a Cramér-Rao bound (CRB) [Kay93], which says

that for any locally unbiased estimator $\tilde{\varphi}$,

$$\Delta^2\tilde{\varphi} \geq \frac{1}{F[p(x|\varphi)]}, \quad F[p(x|\varphi)] = \int \mathrm{d}x \frac{\left(\frac{\partial}{\partial\varphi}\Big|_{\varphi=\varphi_0} p(x|\varphi)\right)^2}{p(x|\varphi_0)}, \tag{1.11}$$

where F is called a *Fisher information* (FI). Henceforth, we drop the φ_0 notation and simply write φ.

We now discuss the saturability of the CRB. First, it can be always saturated by a locally unbiased estimator, which leads to the conclusion that

$$C_{\text{freq}} = \frac{1}{F[p(x|\varphi)]}. \tag{1.12}$$

However, the operational interpretation of this statement requires care. An estimator that attains the CRB is, in general, guaranteed to be unbiased only *locally*. As a result, the CRB is not necessarily saturable in practical scenarios, when the initial knowledge about the parameter is limited.

In practice, one often repeats an experiment multiple times, and construct an estimator $\tilde{\varphi}(x_1, ...x_k)$ based on outcomes from k independent repetitions. The resulting CRB reads

$$\Delta^2\tilde{\varphi} \geq \frac{1}{kF[p(x|\varphi)]}, \tag{1.13}$$

where F is a FI for a single repetition. This version of the bound is more operational—in the asymptotic limit $k \to \infty$, one can construct *globally* unbiased estimator saturating the bound [Kay93]. In the limit of finite k, the bound (1.13) is still valid, and often informative, but not necessarily saturable.

The FI is central quantity in the frequentist framework. In this setting, the general quantum metrological problem is formulated as maximizing the FI over $\mathcal{C}$ and $\boldsymbol{M}$:

$$F_Q(\mathcal{S}_\varphi) = \max_{\boldsymbol{M},\mathcal{C}} F[p(x|\varphi)], \quad p(x|\varphi) = \mathrm{Tr}(\rho_\varphi M_x), \quad \rho_\varphi = \mathcal{C} * \mathcal{S}_\varphi, \tag{1.14}$$

where $F_Q(\mathcal{S}_\varphi)$ is called a *protocol quantum Fisher information* (protocol QFI).

When $\mathcal{C}$ is fixed, the problem reduces to maximizing the FI over all measurements $\boldsymbol{M}$ acting on the given φ-dependent output state ρ_φ. The resulting quantity is the *state quantum Fisher information* (state QFI or simply QFI), defined as

$$F_Q(\rho_\varphi) = \max_{\boldsymbol{M}} F[p(x|\varphi)], \quad p(x|\varphi) = \mathrm{Tr}(\rho_\varphi M_x). \tag{1.15}$$

Interestingly, there is a closed formula for the state QFI [Hel69],

$$F_Q(\rho_\varphi) = \mathrm{Tr}(\rho_\varphi L_\varphi^2), \quad \dot{\rho}_\varphi = \frac{1}{2}(\rho_\varphi L_\varphi + L_\varphi \rho_\varphi), \tag{1.16}$$

where $\dot{\rho}_\varphi = \frac{\partial}{\partial\varphi}\rho_\varphi$, and the right-hand-side part of the above equation is the definition

of a matrix L_φ, which is called a *symmetric logarithmic derivative* (SLD).

For pure states, the formula for the QFI is more straightforward,

$$F_Q(|\psi_\varphi\rangle) = 4\left(\langle\dot{\psi}_\varphi|\dot{\psi}_\varphi\rangle - \left|\langle\dot{\psi}_\varphi|\psi_\varphi\rangle\right|^2\right). \tag{1.17}$$

Importantly, the state QFI is additive for independent quantum states,

$$F_Q(\rho_\varphi \otimes \sigma_\varphi) = F_Q(\rho_\varphi) + F_Q(\sigma_\varphi). \tag{1.18}$$

When k independent copies of ρ_φ are measured, then the MSE of any locally unbiased estimator of φ satisfies

$$\Delta^2\tilde{\varphi} \geq \frac{1}{kF_Q(\rho_\varphi)} \tag{1.19}$$

for any measurement $\boldsymbol{M}$. This inequality is called *Quantum Cramér-Rao bound* (QCRB).

The protocol QFI (1.14) can be thus expressed as the following maximization of the state QFI

$$F_Q(\mathcal{S}_\varphi) = \max_{\mathcal{C}} F_Q(\mathcal{C} * \mathcal{S}_\varphi). \tag{1.20}$$

Let us compare the Bayesian and frequentist approaches. The operational meaning of the Bayesian cost is better justified—the initial knowledge about the parameter is well defined by the prior distribution. In constract, for the frequentist approach, one gets an upper bound for the MSE (1.13), which is only guaranteed to be saturable for infinitely large number of experiment's repetitions k, or when we sense infinitesimal changes of a parameter around a well-known value.

On the other hand, the frequentist approach leads to a much simpler optimization problem because the optimization over estimator and measurement can be done analytically, by calculating classical FI or state QFI respectively—the only optimization left is the optimization over $\mathcal{C}$, see (1.20). Moreover, the results obtained using the frequentist approach may be regarded as universal bounds applicable even when it is not clear how to define a prior distribution.

Importantly, the results obtained using the two approaches can be connected. For example, the van Trees inequality bounds the Bayesian average quadratic cost using the FI averaged over the prior [Jar15; Dem20]. The connection between frequentist and Bayesian results can also be established through the relation between the FI and the mutual information [Gór25].

In the study of complex metrological scenarios, a frequentist approach is often the first choice, as it leads to tractable bounds and offers good intuition regarding the protocols leading to quantum advantage. The Bayesian approach then serves as a validation of operational relevance of this advantage. The comparison between the two approaches sometimes leads to surprising conclusions, such as correction of truly operational Heisenberg limit by a factor π [Gór20].

Even in the frequentist setting, the optimization of the state QFI over protocols (1.20)

is often a highly non-trivial task, especially for large number of probed channels N, and for more complex metrological scenarios. In fact, this is one of the main research goals of this thesis—to use the frequentist approach to extend the applicability of general metrological tools to broader class of scenarios.

1.4 Context and contributions

Following this brief theoretical introduction, we now provide an overview of the historical development of quantum metrology theory and place the results of this thesis in the context of existing literature.

The foundations of quantum estimation theory were laid in the seminal work of Helstrom [Hel69], and later in Refs. [Bra94; Hol11]. The introduced framework provides the conceptual basis of quantum metrology: although it does not address the optimization of probe states, it introduces the essential tools for metrological analysis, such as state QFI, or Bayesian cost defined for parameter encoded in a quantum state.

In Ref. [Cav81], Caves proposed the first explicit quantum-enhanced metrological protocol, showing that squeezed states of light [Yue76] can be employed to reduce quantum noise in optical interferometers and thereby enhance phase measurement precision. This work was motivated by the early proposals for gravitational-wave detection experiments [Tho80] and the studies of fundamental limitations of such systems [Hol79; Cav80]. The idea of Caves was then further developed [Bon84; Yur86b], demonstrated experimentally [Xia87], and extended to spin systems [Win92; Kit93]. The proposed schemes offer the HS of the precision in the absence of noise, and belong to class B1 (we use classification from Fig. 1.1), as squeezed states belong to the family of entangled states [Dem15]. Interestingly, as early as in Ref. [Cav81], the role of photon losses was studied, so some studied scenarios belong to class B2. More than 30 years after the original proposal, squeezed states were injected into the LIGO gravitational-wave detector [Aas13], leading to the first squeezing-enhanced observation of a gravitational-wave signal in 2019[1] [Tse19]. One can, therefore, admire the fact that historically first proposed metrological scheme is one of the most striking example of practical quantum advantage nowadays!

Going back to 90s, other early-studied metrological scenario was a frequency measurement with two-level probes such as spins 1/2 [Win92] or ions with two available energy levels [Ita93]—these scenarios belong to classes A1 (when probes are independent, as in Ref. [Ita93]) or B1 (when probes are entangled as in Ref. [Win92]).

Soon after the advent of squeezing-based metrology, it was shown that other types of entangled states can also enhance the precision of interferometric [Yur86a; Hol93; Dow98] and spectroscopic measurements [Bol96; Hue97]. The notable examples are twin-Fock states resulting from the interference of two Fock states on a beam-splitter [Hol93] and already mentioned N00N states [Bol96; Bot00; Lee02; Nag07; Dow08].

[1] It should not be confused with first-ever gravitational wave detection without squeezed states in 2015 [Abb16].

The growing number of examples of metrological protocols prompted more general considerations about the ultimate metrological limits. In Ref. [Bol96], it was argued, using uncertainty relations, that the HS is the optimal scaling in noiseless, entanglement-based phase estimation task (class B1), and that N00N states offer the optimal phase estimation precision. The phase estimation problem was also examined more rigorously using the Bayesian framework [Ber00], which enabled construction of optimal probe states for practical estimation scenarios with limited prior knowledge.

The formal analogy between optical interferometers, quantum spectroscopy, and quantum circuits [Lee02] enabled the introduction of a unified framework for *noiseless* quantum metrology [Gio04; Gio06]. In Ref. [Gio06], the distinction between schemes A1, B1, and C1 was explicitly introduced, and the corresponding fundamental precision limits were rigorously derived within a frequentist framework. Interestingly, this led to the conclusion that adaptive schemes (C1) offer no fundamental advantage over parallel schemes (B1) in the absence of noise—in both B1 and C1, the HS with the same constant is achievable, while only the SS can be obtained in A1. The Ref. [Gio06] solves the frequentist optimization problem (1.20) for noiseless cases (A1, B1, C1) in full generality. Nevertheless, extending these results to the Bayesian setting remained a significant challenge, addressed in later works [Dem11; Jar15; Gór20; Hay25].

Developing a comprehensive theory of *noisy* quantum metrology proved substantially more challenging even within frequentis framework. In Ref. [Hue97], entangled input states were numerically optimized for phase estimation under dephasing noise, providing one of the earliest examples of optimization within class B2. This result, however, was restricted to a specific noise model and small probe numbers N. Ref. [Fuj01] subsequently showed that entangling the probe with an ancillary system $\mathcal{A}$ can enhance estimation precision, revealing features of noisy channel estimation absent in state or unitary estimation. Subsequently, Ref. [Sar06] derived a general upper bound for ancilla-assisted single-channel estimation (B2 with $N = 1$), though the tightness of the bound was not analyzed. This work laid the foundations of the so-called *minimization over purification* (MOP) technique for computing upper bounds for the QFI of noisy channels. This technique was significantly improved in Ref. [Fuj08], in which the single-channel bound from Ref. [Sar06] was shown to be saturable for sufficiently large ancilla $\mathcal{A}$. Even more importantly, the bound was extended to large-N (and even asymptotic) cases, for which, however, the tightness remained unclear.

The idea of MOP-based bounds for scheme B2 was further developed in Ref. [Esc11], and its full practical potential was realized in Refs. [Dem12; Koł13], where efficient bounds computation via semidefinite programs (SDPs) [Boy04] was demonstrated. This enabled obtaining informative bounds for physically interesting noise models, such as dephasing, depolarization, amplitude damping, and photon losses. Ref. [Dem12] demonstrated that HS is usually asymptotically suppressed in the presence of noise. However, even in such cases, probe entanglement remains metrologically useful, as it may improve the constant in the SS.

The derived upper bounds for QFI were not constructive, and their tightness was not guaranteed. A constructive method for input probe state optimization, called *iterative see-saw* (ISS) optimization, was therefore proposed for Bayesian [Dem11; Mac14] and frequentist [Mac13] frameworks. Later, it was also shown how to find optimal probes within frequentist framework using MOP [Zho20; Zho21].

A detailed comparison of the MOP and ISS approaches is provided in Chapter 2. Importantly, for both methods, optimal probes can be found only for relatively small ($N \lesssim 5$) numbers of channels due to the *curse of dimensionality*—the Hilbert space of entangled probes grows exponentially with N. Even though MOP allows to compute QFI upper-bounds for large N as well, those bounds do not give any clue about the optimal input state.

To remedy this, the QFI ISS optimization was supplemented with the tensor-network based matrix product state (MPS) [Aff88; Fan92; Öst95] representation of the input probe in Refs. [Cha20; Cha22]. This approach was motivated by the finding that low-rank MPSs are often optimal in noisy metrological scenarios [Jar13]. The use of MPS enabled efficient probe optimization for schemes A2 and B2 with moderately large channel numbers ($N \lesssim 100$), and additionally for schemes A3 and B3 with $N \lesssim 100$ and relatively simple noise correlation models. The method was recently extended to more general correlations [Dul26].

Finally, Ref. [Zho21] completed the theoretical understanding of scheme B2 in the frequentist setting by proving that MOP-based fundamental bounds are asymptotically saturable through direct construction of asymptotically optimal metrological protocols.

Understanding metrological limitations of adaptive schemes with uncorrelated noise (C2) proved more challenging. This class of schemes was first studied in Ref. [Esc11], which showed, using a frequentist approach, that the scaling type (HS or SS) is identical for B2 and C2. It remained unclear, however, whether adaptiveness improves the *constant* in the precision scaling.

A fundamental upper bound for the QFI of schemes C2 was derived using the MOP technique in Ref. [Dem14]. The bound was asymptotically equivalent to the parallel bound (covering schemes B2) for SS cases but yielded a larger coefficient for HS. Whether this discrepancy reflected a genuine advantage of adaptiveness or simply a loose bound remained a long-standing open question, though the authors conjectured that B2 and C2 achieve the same asymptotic precision.

The adaptive bound was then explored further, with special emphasis on the continuous-time limit [Dem17; Sek17; Zho18], where the parameter is encoded in a Hamiltonian, noise is modeled via the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation, and arbitrarily fast control is allowed. Note that this continuous-time limit is in fact a special case of the discrete scenario C2, where channels Λ_φ represent the evolution given by the GKSL equation integrated over an infinitesimal time interval. The converse is not true—there are channels Λ_φ that cannot be obtained from integrating a time-independent GKSL equation; moreover, the parameter may be encoded in the noise rather than in the Hamiltonian. For the

continuous-time scenario with unitary parameter encoding, universal error-correction codes were constructed [Zho18; Zho20], and asymptotically tight upper bounds for the QFI were subsequently derived [Wan22].

Alternative approaches to upper bound the QFI of schemes B2 for general cases with arbitrary noise and parameter encoding have also been proposed [Pir17; Per21]. However, these approaches lacked either generality or tightness, leaving the asymptotic equivalence between B2 and C2 unresolved.

This long-standing conjecture was resolved in our work [Kur23b], presented in **Chapter 3** of this thesis, where we prove that adaptive and parallel schemes achieve identical asymptotic precision for uncorrelated noise. More precisely, we derived a new upper bound for the QFI of adaptive schemes (C2) that is tighter than previous bounds and asymptotically equivalent to the existing bound for parallel schemes (B2). Since the B2 bound was known to be asymptotically tight [Zho21], and B2 forms a subclass of C2, our bound for C2 must also be asymptotically tight. This establishes that adaptive schemes offer no asymptotic advantage over parallel schemes for uncorrelated noise of any type.

The lack of asymptotic advantage does not mean that adaptive strategies are useless. Not only can they offer advantages over parallel protocols for finite N, but more importantly, they may be significantly easier to implement in certain circumstances, for example, when the optimal parallel strategy requires large-scale entanglement. Therefore, finding a general recipe for optimal quantum adaptive strategies remained an important problem.

Solutions based on the MOP technique were provided in Refs. [Alt21; Liu23], where it was shown how to optimize the QFI over all adaptive strategies, represented by quantum combs, using a single SDP. Moreover, the proposed procedures also accommodated correlated signal and noise, thus providing universal solutions to the QFI optimization problem schemes C2 and C3, as well as for the parallel schemes with correlated noise, B3.

Unfortunately, the proposed methods suffered from the curse of dimensionality, restricting their applicability to small numbers of probed channels, $N \lesssim 5$. As in the parallel case, the remedy was to employ a tensor-network representation of the optimized strategy. This time, however, the optimized object is a chain of control operations C_i rather than a many-particle entangled input state ρ, requiring a different type of tensor network [Whi22; But24] to represent the strategy.

Tensor-network based optimization of the QFI for adaptive strategies was developed independently in two concurrent works: ours [Kur25c], presented in **Chapter 2** of this thesis, and that of Liu et al. [Liu24]. The ISS method enabled iterative optimization over different network nodes representing subsequent control operations, thereby circumventing the curse of dimensionality and enabling optimization of adaptive metrological protocols for $N \lesssim 100$. This was only feasible for relatively low bond dimensions of the network representing an adaptive strategy—this bond dimension can be interpreted as the dimension of the ancillary system or, equivalently, as the amount of quantum memory available for the estimation protocol. While

this restriction to low ancilla dimensions is a limitation, low-memory protocols are typically much easier to implement experimentally, making the search through such protocols practically beneficial.

The tensor-network based optimization presented in Chapter 2 is complementary to the MOP-based bound presented in Chapter 3. When the QFI of the protocol represented by the tensor network converges to the fundamental bound, this validates both the saturability of the upper bound and the optimality of the found protocol.

The chain of correlated signal-encoding channels Λ_φ can be represented as a tensor network of the same type as that used to represent control operations. Therefore, as we show in Chapter 2, our optimization method also applies to correlated signal and noise (schemes C3). A general pattern emerges: understanding the structure of quantum adaptive strategies facilitates handling correlated Λ_φ cases, as both require similar mathematical structures. This also explains why, paradoxically, the less general scheme C2 is slightly harder to handle than the more general scheme C3—in C2, two inherently different structures must be connected. Nevertheless, it was shown in Ref. [Dul26] that such a connection is possible: one can optimize an entangled input represented by an MPS evolving through correlated channels represented by a network of a different type.

Having mentioned correlations in the context of the results presented in Chapter 2, let us now provide a more comprehensive review of quantum metrology with signal and noise correlations.

The effects of non-Markovian system-environment interactions have been studied for dephasing noise [Mat11; Chi12] and then for more general noise models [Mac15; Smi16; Haa18]. It was shown that the memory effects dominate probe dynamics at sufficiently short time scales, known as the Zeno regime [Ita90]. This can be exploited metrologically: when the total measurement time T is divided into short time intervals t below the Zeno scale, with measurements performed after each interval, the precision for estimating a unitary parameter scales as $\Delta\varphi \sim 1/N^{3/4}$ for N entangled probes and optimally chosen t. Therefore, the standard scaling $\Delta\varphi \sim 1/\sqrt{N}$, achievable in the markovian dephasing regimes, is surpassed.

However, in all the mentioned studies, non-Markovian effects were considered only within individual time intervals t (between consecutive measurements), while correlations *between intervals* were neglected. That is, the environment was assumed to reset completely after each interval t, eliminating correlations between successive experimental runs. This reduced the problem to $k = T/t$ independent repetitions of scheme B2, with additional optimization over t for fixed T and N [Smi16]. This simplification, unfortunately, is not well justified in the analyzed small-t regime, where the environment retains memory across intervals and cannot be treated as resetting between runs [Zho17].

The impact of *temporal correlations* caused by environmental non-Markovianity was studied more rigorously in Ref. [Yan24] for multiple types of noise spectra corresponding to different experimental platforms, such as semiconductor quantum dots and NV centers. In this work, environment-memory-induced correlations were

addressed properly, and numerical optimization over a certain class of control-enhanced protocols (C3) was performed. The results, though, were limited to small number ($N \leq 5$) of entangled probes.

Metrological scenarios with *spatial correlations* have been investigated for various noise models, including correlated dephasing of entangled input probes [Jes14] (a particular instance of B3) and two-body losses in Bose-Einstein condensate magnetometry [Cza19]. Quantum error correction codes addressing spatially correlated noise were developed in Refs. [Lay18; Lay19].

Both *spatial and temporal* noise correlations have been studied in the context of Gaussian systems and squeezing-based quantum metrology [Bea18; Rib22]—the Gaussianity of the setup allowed to explore large-N limits.

Correlations can serve as an information resource, not just noise. In thermometry [Sea19; Shu20; Men25], correlations between initially independent thermometers arising from their interaction with a common thermal bath (a particular case of scheme A3) were shown to be an important source of information [Pla22].

In general, the schemes belonging to class A3 are often called collisional models or quantum Markov chains, and have been studied systematically in Refs. [Guţ11; Cat15; Cic22; God23]. In all these studies, initially independent particles sequentially probe a measured system, and their post-interaction correlated state is used for estimation. Therefore, the measured system serves as correlation-mediating environment. For collisional models, asymptotic upper bounds for the QFI have been derived, and some specific measurement protocols have been studied. The crucial assumption that has been made was that probe particles are initially identical and independent—this allowed to use *local asymptotic normality* phenomenon to derive asymptotic bounds.

The general theoretical framework to deal with QFI optimization in the presence of correlations was introduced in Ref. [Yan19], where the most general scheme depicted in Fig. 1.1, C3, was explicitly introduced—it was proposed to represent both the signal encoding and the adaptive estimation strategy as quantum combs. This work derived the first upper bound for the QFI of such schemes. While the approach was general, its practical utility was limited: the numerical value of the derived bound scaled exponentially with N, making it loose and uninformative, especially for large systems. More practical theoretical approaches were presented in Refs. [Alt21; Liu23], demonstrating how to find the *exact* optimal QFI value for scheme C3 using a single SDP. The SDP formulation in Ref. [Alt21] correctly maximized the QFI, but contained an error in constructing the corresponding optimal estimation strategy, which was subsequently corrected in Ref. [Liu23], in which, additionally, symmetry-based reduction of the algorithm-complexity was proposed. Despite this progress, the time and memory of the proposed methods scaled exponentially with N due to the curse of dimensionality. This limited the practical applicability of the algorithms to $N \lesssim 5$.

As mentioned earlier, the curse of dimensionality can be circumvented by representing both the estimation strategy and signal-encoding combs as tensor networks (Chapter 2, Ref. [Kur25c]). The efficiency of this procedure, however, relies on the dimensions of

the correlation-mediating environment and the ancillary system being sufficiently small. As a result, the obtained QFI is not guaranteed to be strictly optimal and should be interpreted as a lower bound on the true optimum of (1.20).

To determine how close the tensor-network-optimized QFI is to the true optimum, one needs informative fundamental QFI upper bounds. Unfortunately, the bounds presented in Chapter 3 [Kur23b] are valid only for uncorrelated noise (scheme C2), while those derived in Ref. [Yan19] are too loose to be useful.

To address this gap, we extended the methods from Ref. [Kur23b] to correlated scenarios, deriving general upper bounds for scheme C3. This result appears in **Chapter 4**, which presents Ref. [Kur25b]. Since A3 and B3 are special cases of C3, the bounds apply to these schemes as well. Importantly, our results prove that standard and Heisenberg scalings of the QFI remain fundamental even in the presence of correlations, assuming that subsequently probed channels are the same. This is in stark contrast to Ref. [Yan19], in which the bounds scaled exponentially with N.

While our bounds are not guaranteed to be tight, their tightness can be systematically improved at the cost of increased computational complexity. Comparison with tensor-network results [Kur25c] for correlated dephasing and collisional quantum thermometry models, showed close agreement between upper and lower bounds, validating both approaches.

At this point, let us summarize the theoretical landscape within which the Chapters 2–4 of this thesis are situated.

Early quantum metrology focused on concrete protocols, but as the field matured, more universal theoretical frameworks emerged. In particular, general tools for solving the QFI optimization problem (1.20) were developed to handle metrological scenarios of progressively increasing complexity.

For the simplest scenarios involving only unitary encoding (A1, B1, C1), an exact analytical solution to (1.20) was provided in Ref. [Gio06] for arbitrary N. However, for more complex scenarios involving noise, the problem becomes significantly more challenging. Exact solutions can be found in practice only for relatively small numbers of channels ($N \lesssim 5$)—even this regime is highly nontrivial and was fully solved only recently. For larger N, rather than seeking exact solutions, one can derive fundamental upper bounds or constructive lower bounds applicable up to $N \lesssim 100$ through the use of tensor-network methods.

Table 1.1 summarizes key references for the QFI optimization proble (1.20) across the different scheme classes from Fig. 1.1, and indicates how Chapters 2–4 of this thesis advance the state of the art. For each scheme class, we distinguish three types of results: (i) exact solutions to (1.20) for $N \lesssim 5$, (ii) constructive lower bounds valid for larger systems ($N \lesssim 100$), and (iii) fundamental upper bounds valid for arbitrary N.

All algorithms for QFI optimization and upper bound computation listed in Table 1.1, including those presented in Chapters 2–4, are available in our Qmetro++ Python package [Dul26]. This package not only implements previously developed techniques

but also provides a unified framework encompassing different types of tensor networks.

<table>
<tr><th></th><th>Exact ($N \approx 5$)</th><th>Constructive lower bound ($N \approx 100$)</th><th>Upper bound (any N)</th></tr>
<tr><td>A1, B1, C1</td><td colspan="2">[Gio06] [1]</td><td></td></tr>
<tr><td>A2</td><td colspan="2">[Mac13] [2]</td><td>[Sar06] [2], [Fuj08] [2], [Dem12] [2]</td></tr>
<tr><td>B2</td><td>[Fuj08][3], [Dem12][3] [Zho21][3] [4]</td><td>[Cha20]</td><td>[Fuj08], [Dem12]</td></tr>
<tr><td>C2</td><td>[Alt21], [Liu23][4]</td><td>Chapter 2</td><td>[Dem14], Chapter 3</td></tr>
<tr><td>A3</td><td colspan="2">[Cha20][5], [Dul26]</td><td rowspan="3">Chapter 4</td></tr>
<tr><td>B3</td><td rowspan="2">[Alt21], [Liu23][4]</td><td>[Cha20][5], [Dul26]</td></tr>
<tr><td>C3</td><td>Chapter 2</td></tr>
</table>

[1] Exact solution for any N
[2] Solution for $N = 1$ generalizes to any N due to QFI additivity for independent quantum states.
[3] Solution for $N = 1$ generalizes to small N when one replaces $\Lambda_\varphi \to \Lambda_\varphi^{\otimes N}$.
[4] Correct construction of optimal protocol using saddle point method.
[5] For limited class of correlated models only.

Table 1.1: Overview of theoretical works addressing the QFI optimization problem (1.20) for protocol classes (A1–C3) depicted in Fig. 1.1. For each class, we distinguish: (i) exact solutions for $N \lesssim 5$; (ii) constructive lower bounds for $N \lesssim 100$; (iii) upper bounds for arbitrary N. The contributions presented in Chapters 2–4 of this thesis are highlighted.

1.5 Quantum-inspired superresolution

In the previous section, we focused on the results relevant to the broader field of quantum metrology, with particular emphasis on QFI optimization over probe states and control operations. This provided context for the results presented in Chapters 2–4 of this thesis. The results in Chapters 5–6 are motivated by a specific application of quantum metrology: quantum-inspired superresolution imaging. We therefore now introduce its basic concepts.

Ideally, an optical imaging system would produce a point image—a Dirac delta intensity distribution—from a pointlike emitter in the object plane. In reality, no such ideal systems exist. Every imaging apparatus has finite apertures, causing light diffraction that blurs the resulting image. This limits the imaging *resolution*—the ability to resolve small details of an object. The classical resolution limit, formulated by Rayleigh [Ray79], states that the minimal distance between two resolvable object is $d = 1.22\lambda/2\mathrm{NA}$, λ is a wavelength, and NA is the numerical aperture of a considered system, which is typically close to unity.

However, this limitation is valid only under certain assumptions, such as uniform and classical illumination of the imaged sample, non-fluctuating power of light emitters, standard measurement of light in the image plane, etc. Breaking one or more of these assumption allows to overcome the Rayleigh limit, and the techniques allowing to do so are known under the common name of *superresolution imaging*[Rus06; Bet06; Der09; Leu11; Sch13; Roz14; Gat14; Moe15; Tsa16; Par18; Mor19]. The brief discussion and classification of superresolution techniques will be given in Chapter 6. Here, we are going to focus on the description of optical imaging from the perspective of estimation theory, and show how such description inspired new approach to superresolution.

Let us examine the optical imaging theory more closely by considering a linear, translationally invariant optical imaging system. When the object consists of N incoherent point light emitters with powers q_i and positions $\boldsymbol{r}_i$ in the object plane, the intensity distribution in the image plane is

$$I(\boldsymbol{r}') = \sum_{i=1}^{N} q_i \eta U(\boldsymbol{r}' - \boldsymbol{r}_i), \tag{1.21}$$

where η is an attenuation factor and U is the point spread function (PSF), describing the normalized intensity distribution from a single point emitter. We assume the system's magnification factor is unity; relaxing this assumption amounts to coordinate rescaling.

The PSF width dictates the imaging system's resolving power. When the PSF is too wide, images from different emitters overlap, making them difficult to distinguish—this underlies the classical Rayleigh resolution limit. For systems with circular apertures, the PSF has the form of an *Airy disk* [Bor13]. In theoretical studies, this is commonly approximated by a Gaussian shape of width σ, which characterizes the system's resolution scale. Furhtermore, we will assume that both object and image are 1D, so we can replace 2D vectors $\boldsymbol{r}, \boldsymbol{r}'$ with single coordinates x, x'. All the results can be easily generalized to 2D case and more general PSFs, but this approximation allows to focus on the essence of presented phenomena. The Gaussian 1D PSF is

$$U(x') = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x'^2}{2\sigma^2}}. \tag{1.22}$$

Let us formulate optical imaging as an estimation problem. To make it as simple as possible, consider an object consisting of two equally bright point sources with a known centroid, placed at $x_+ = s/2$, $x_- = -s/2$. Then, the separation s is the only unknown parameter, which we want to estimate. The observed intensity distribution is

$$I(x') = \frac{P_0}{2} \left(U(x' - x_+) + U(x' - x_-) \right), \tag{1.23}$$

where P_0 is the total power observed in the image plane.

When accounting for the quantum nature of light, the normalized intensity distribution should be interpreted as the probability density function for detecting a photon

at a given position in the image plane. For weak Poissonian sources, subsequent photon detections are independent, and the probability of detecting a single photon at position x' is

$$p(x'|s) = \frac{1}{2}\left(U(x' - x_+) + U(x' - x_-)\right). \tag{1.24}$$

For k detected photons, the CRB (1.13) implies that the MSE of any estimator of the separation between two sources satisfies

$$\Delta^2 \tilde{s} \geq \frac{1}{kF(s)}, \tag{1.25}$$

where $F(s)$ is a separation-dependent single-photon FI calculated for $p(x'|s)$. For a Gaussian PSF (1.22), performing an integral from (1.11), yields the FI $F(s)$ plotted in Fig. 1.2. Crucially, in the subdiffraction regime $s \ll \sigma$, we have

$$F(s) = \frac{1}{\sigma^2}\left(\frac{s^2}{8\sigma^2} + \mathcal{O}\left(\frac{s^4}{\sigma^4}\right)\right), \tag{1.26}$$

so the FI vanishes as $s \to 0$. Consequently, estimating s becomes significantly harder for $s \ll \sigma$ because, according to (1.25), the MSE of any locally unbiased estimator diverges. In other words, one must substantially increase the number of collected photons k to achieve a desired precision. This behavior is called the *Rayleigh curse*, as it provides an estimation-theoretic explanation for the difficulty of measuring object features at length scales below the Rayleigh limit.

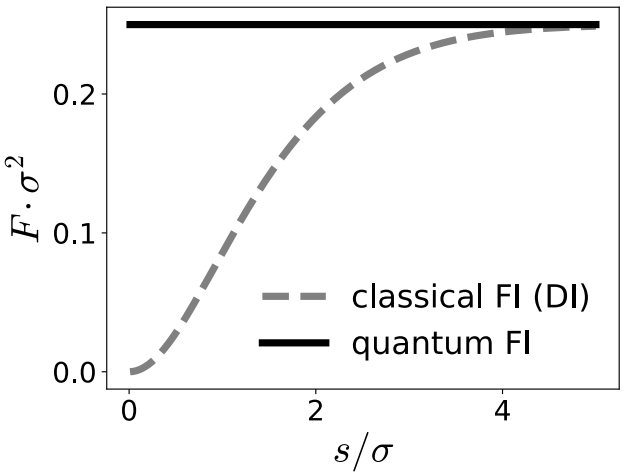


Figure 1.2: Fisher information (FI) for separation (s) estimation of two incoherent light emitters imaged through an appartus with a Gaussian point spread function of width σ. Classical FI for direct imaging (gray dashed) and quantum FI (black solid) are shown.

The above reasoning assumed position-basis measurements of light in the image plane, known as direct imaging (DI). Quantum mechanics, however, permits a much broader class of measurements. Following the seminal work of Tsang [Tsa16], we now explore the consequences of going beyond the DI paradigm.

For the considered two-point object, the density matrix of a single photon in an image plane is

$$\rho_s = \frac{1}{2}\left(|u_+\rangle\langle u_+| + |u_-\rangle\langle u_-|\right), \tag{1.27}$$

where

$$|u_\pm\rangle = \int \mathrm{d}x' u(x' - x_\pm)\,|x'\rangle \tag{1.28}$$

are pure states of photons emitted from right/left source respectively, with $u(x')$ satisfying $|u(x')|^2 = U(x')$—it is also typically assumed that $u(x') \in \mathbb{R}_+$, so $u(x') = \sqrt{U(x')}$, $|x'\rangle$ are eigenstates of a position operator. Note that density matrices, not state vectors, are added to represent the contributions from two sources because we assumed the sources are incoherent.

We can now apply *quantum estimation theory* tools and calculate the QFI of the parameter-dependent state ρ_s, which equals the classical FI maximized over all measurements $\boldsymbol{M}$. This yields the surprising result

$$F_Q(\rho_s) = \frac{1}{4\sigma^2}, \tag{1.29}$$

which is independent of s! It means that optimal quantum measurement may, in principle completely circumvent the Rayleigh curse, and allow for precise estimation of the separation between two points with highly overlapping PSFs.

As shown in Ref. [Tsa16], one possible QCRB-saturating measurement, for which the classical FI equals QFI, is the measurement of the optical field in the basis of Hermite-Gaussian (H-G) modes

$$|\phi_j\rangle = \int \mathrm{d}x' \phi_j(x')\,|x'\rangle, \quad \phi_j(x') = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \frac{1}{\sqrt{2^j j!}} H_j\left(\frac{x'}{\sigma\sqrt{2}}\right) \exp\left(-\frac{x'^2}{4\sigma^2}\right), \tag{1.30}$$

where H_j are Hermite polynomials. This choice is not unique—the general construction of the family of optimal measurements will be given in Chapter 5.

Interestingly, the optimal measurement proposed by Tsang can be implemented using classical linear optics. Therefore, the resulting superresolution technique does not rely on inherently quantum effects. Nevertheless, quantum estimation theory provides the appropriate tools for finding the optimal superresolving measurement. Consequently, Tsang's technique is often called *quantum-inspired superresolution*.

The technique has been extended to more complex, 2D objects with multiple sources [Tsa17], non-Gaussian PSFs [Tsa18], and its usefulness has been demonstrated in proof-of-principle experiments [Pus21; Fra23; Dup25].

Importantly, the impact of various noise sources has been studied, including detector dark counts [Lup20], crosstalk [Ges20], and device misalignment [Alm21]. These studies show that, unfortunately, the FI drops to zero as $s \to 0$ in the presence of even arbitrarily small noise of almost any type [Oh21]. Nevertheless, this does not mean that the proposed techniqe is useless—it still achieves larger FI than DI for small but nonzero values of s. In general, noise effects are most severe as $s \to 0$ but

have only a minor effect on estimation precision for larger s. This aligns with the intuition that the small-s regime is more problematic due to significant PSFs overlap.

This case study illustrates a general limitation of QFI-based analysis: the QCRB is guaranteed to be saturable only when one can *precisely* implement *every* measurement allowed by quantum mechanics. In practice, the intended QCRB-saturating measurement is often disturbed by noise. This may drastically reduce the resulting estimation precision in some regimes (e.g., $s \to 0$ for our superresolution model), while in other regimes such disturbances may be less severe (e.g., when $s \gg \sigma$).

This behavior agrees with our intuition once we understand the details of this particular model. Nevertheless, when comparing different regimes (small s versus large s) based solely on the QFI of ρ_s, it appears that estimation can be performed equally well in both cases.

To capture the difference between the two regimes without referring to any particular noise model, one needs to quantify how *susceptible* the FI is to noise disturbing the measurement. In Chapter 5 [Kur23a], we introduce a general quantity called *Fisher information measurement noise susceptibility* (FI MeNoS), which addresses this challenge—it quantifies the maximal rate of FI decrease caused by infinitesimal measurement noise of any type. We derive a general formula for FI MeNoS that can be applied to different physical models beyond superresolution imaging, such as quantum interferometry. For quantum-inspired superresolution, we show that the noise susceptibility of all QCRB-saturating measurements diverges as $s \to 0$. This clearly explains why this regime is problematic to handle despite the non-decreasing value of the QFI.

1.5.1 The role of photon correlations

So far, we assumed weak, Poissonian light sources, for which photon detections are independent events, so a single-photon density matrix captures all relevant effects. However, light emitters used in optical imaging often exhibit non-Poissonian behavior, which is either caused by *blinking* phenomenon [Rus06; Der09], leading to super-Poissonian statistics, or by inherently quantum *antibunching* effect [Sch12; Sch13], leading to sub-Poissonian statistics. In both cases, it was demonstrated that non-Poissonianity leads to increase of the DI resolution, which is reflected by the increased value of the classical FI [Kur21]. However, as we demonstrate below, this is no longer true at the level of the QFI—when optimal measurement is performed, there is no gain from non-Poissonian light statistics for estimation of separation between two equally bright incoherent sources with a known centroid.

Let us assume that we measure k photons in total. Let $r_i \in \{+, -\}$ denote the position of the emitter from which the ith photon was emitted ($+$ for x_+, $-$ for x_-). We treat photons as distinguishable, which is justified in the weak-source regime where temporal modes of different photons are orthogonal—this assumption breaks down, for example, for strong thermal sources [Nai16]. In the absence of correlations, each photon comes from the x_+ or x_- source with equal probability $1/2$, so the total

probability that the ith photon comes from source r_i for $i \in \{1, 2, \ldots, k\}$ is

$$P(r_1, \ldots, r_k) = \frac{1}{2^k}$$

for any $r_1, r_2, \ldots, r_k$.

Then, the density matrix of k photons is simply $\rho_s^{(k)} = \rho_s^{\otimes k}$, where ρ_s is the single-photon state (1.27), and the corresponding QFI is, due to QFI additivity (1.18),

$$F_Q(\rho_s^{(k)}) = kF_Q(\rho_s) = \frac{k}{4\sigma^2}. \tag{1.31}$$

Let us now consider a general correlated model for which the multi-variable probability distribution $P(r_1, \ldots, r_k)$ can be arbitrary. The density matrix of k photons is then

$$\rho_s^{(k)} = \sum_{r_1,\ldots,r_k \in \{+,-\}} P(r_1, r_2, \ldots, r_k) \bigotimes_{i=1}^{k} |u_{r_i}\rangle\langle u_{r_i}| \,. \tag{1.32}$$

It is challenging to compute the QFI of $\rho_s^{(k)}$, as it cannot generally be written as a tensor product of single-photon density matrices. However, an upper bound can be found using QFI convexity [Hol11], from which we obtain

$$F_Q\left(\rho_s^{(k)}\right) \leq \sum_{r_1,\ldots,r_k \in \{+,-\}} P(r_1, r_2, \ldots, r_k) F_Q\left(\bigotimes_{i=1}^{k} |u_{r_i}\rangle\langle u_{r_i}|\right). \tag{1.33}$$

The QFI of a pure state $|u_{r_i}\rangle\langle u_{r_i}|$ can be calculated directly using (1.17) and (1.28), yielding

$$F_Q\left(|u_{r_i}\rangle\langle u_{r_i}|\right) = \frac{1}{4\sigma^2}, \tag{1.34}$$

which is the QFI of a single source localization. Using QFI additivity (1.18), inserting (1.34) into (1.33), and using the fact that probabilities $P(r_1, r_2, \ldots, r_k)$ sum to 1, we obtain the QFI upper bound

$$F_Q\left(\rho_s^{(k)}\right) \leq \frac{k}{4\sigma^2}, \tag{1.35}$$

which is sufficient to show that correlations do not increase the QFI. Interestingly, they also do not *decrease* it. To prove this, note that for a measurement in H-G modes (1.30) we have

$$|\langle \phi_j | u_+ \rangle|^2 = |\langle \phi_j | u_- \rangle|^2 \tag{1.36}$$

for any j because $\phi_j(x')$ are either even or odd functions of x'. Consequently, if we measure each photon independently in the H-G basis, the statistics of the measurement results are not affected by the choice of $P(r_1, r_2, \ldots, r_k)$, since it does not matter from which source a photon was emitted. This means that the classical FI associated with this measurement applied to k photons is $F = \frac{k}{4\sigma^2}$, which is the result obtained for the uncorrelated case, where the H-G measurement is known to

be QCRB-saturating. However, the classical FI of any measurement always provides a *lower bound* for the QFI, so we can conclude that

$$F_Q\left(\rho_s^{(k)}\right) \geq \frac{k}{4\sigma^2},$$

which together with (1.35) leads to the conclusion that

$$F_Q\left(\rho_s^{(k)}\right) = \frac{k}{4\sigma^2} \tag{1.37}$$

for any distribution $P(r_1, \ldots, r_k)$, and that H-G mode measurement is optimal regardless of the correlation type.

For the simple model above, photon emission statistics do not affect fundamental resolution limits when all measurements allowed by quantum mechanics are possible. This does not mean, however, that signal correlations are useless in general. Indeed, Chapters 5 and 6 demonstrate the *usefulness of signal correlations in quantum-inspired imaging*.

In Chapter 5, we show that correlations resulting from sub-Poissonian statistics caused by antibunching enable construction of measurements less susceptible to noise. We prove this for the separation estimation of a binary source (as in the model considered above) using the introduced FI MeNoS metric.

In Chapter 6, we go beyond the binary source toy model and consider estimation of spatial moments for sub-diffraction objects of arbitrarily complex shape. We show that super-Poissonian light statistics caused by blinking significantly enhances the precision of higher spatial moment estimation. We demonstrate a practical method to construct an efficient estimator and verify its usefulness through numerical simulation. Additionally, we show how temporal correlations enable measurement simplification while preserving the same information content.

Chapter 2

Quantum metrology using quantum combs and tensor network formalism

Commentary

The main goal of the presented work is to introduce an efficient algorithm for QFI optimization over adaptive protocols for uncorrelated and correlated channels (schemes C2, C3 in Fig. 1.1). Unlike previous approaches [Alt21; Liu23], our technique remains applicable even for large numbers of probed channels. This is achieved through tensor-network decomposition of quantum combs representing both the estimation strategy and signal encoding with correlations. Prior works [Cha20; Cha22] had employed tensor networks for metrological optimization in parallel schemes (B2 and partially B3). Extending this approach to adaptive schemes requires a different type of tensor network—one representing quantum processes [Whi22; But24] rather than quantum states.

The tensor-network based optimization relies on the ISS technique [Mac13], which is reviewed in Section 2.2 of the presented paper. As a novel result, the ISS technique is extended to optimization over combs in Section 3.2, and then to optimization over tensor networks representing combs in Section 4.

Additionally, the complementary MOP approach is reviewed in Sections 2.1 (for single channels) and 3.1 (for quantum combs). Consequently, the first three sections provide a detailed overview of QFI optimization techniques that were only briefly mentioned in Chapter 1. The MOP technique is not further developed in this chapter, but it serves as a crucial tool for deriving the bounds presented in Chapters 3 and 4.

Note that the work in Chapter 3 was published before the work in Chapter 2. The chapters are presented in this non-chronological order because Ref. [Kur25c] contains a thorough discussion of foundational concepts used throughout Chapters 2–4. Consequently, the bounds for adaptive schemes with uncorrelated noise (C2) derived in Chapter 3 are already employed in Chapter 2 to assess the optimality of protocols found by tensor-network optimization.

The results similar to those presented here, appeared independently in Ref. [Liu24]

(the preprint appeared on arXiv a week after our preprint).

My contribution

I conceived—through extensive discussions with RDD—and developed two principal new ideas: (i) how to apply the ISS technique to comb optimization, and (ii) how to represent combs as tensor networks to enable this optimization for large N. I implemented the optimization procedures in Python (the code was subsequently further developed by PD). I generated and analyzed the majority of numerical results reported in the paper and devised an explicit optimal adaptive protocol for perpendicular damping noise. I wrote Section 4, the majority of Section 5, and Appendices B, C, and D3. I read and coedited all remaining sections and serve as corresponding author.

IOP Publishing New J. Phys. 27 (2025) 013019 https://doi.org/10.1088/1367-2630/ada8d1

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
28 November 2024

REVISED
31 December 2024

ACCEPTED FOR PUBLICATION
10 January 2025

PUBLISHED
27 January 2025

PAPER

Quantum metrology using quantum combs and tensor network formalism

Stanisław Kurdziałek[1,*], **Piotr Dulian**[1,2,5], **Joanna Majsak**[1,3,5], **Sagnik Chakraborty**[1,4] and **Rafał Demkowicz-Dobrzański**[1]

[1] Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland
[2] Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland
[3] Quantum Research Center, Technology Innovation Institute, Abu Dhabi, United Arab Emirates
[4] Departamento de Física Teórica, Facultad de Ciencias Físicas, Universidad Complutense, 28040 Madrid, Spain
[5] These two authors contributed equally to the project.
[*] Author to whom any correspondence should be addressed.

E-mail: s.kurdzialek@student.uw.edu.pl

Keywords: quantum metrology, quantum Fisher information, quantum combs, quantum processes, tensor networks, correlated noise

Abstract

We develop an efficient algorithm for determining optimal adaptive quantum estimation protocols with arbitrary quantum control operations between subsequent uses of a probed channel.We introduce a tensor network representation of an estimation strategy, which drastically reduces the time and memory consumption of the algorithm, and allows us to analyze metrological protocols involving up to $N = 50$ qubit channel uses, whereas the state-of-the-art approaches are limited to $N < 5$. The method is applied to study the performance of the optimal adaptive metrological protocols in presence of various noise types, including correlated noise.

1. Introduction

One of the main lines of research in theoretical quantum metrology is development of efficient analytical and numerical tools to assess the potential of quantum probes in practical sensing scenarios [1–4]. On the one hand this involves derivation of fundamental bounds on achievable sensitivity in presence of decoherence [5–16], while on the other hand, development of methods that allow to directly identify the optimal metrological protocol in a particular scenario. In this paper we will focus on the latter goal.

The pursuit of identification of the optimal metrological schemes may be carried out on different levels of generality. The less fundamental, but at the same time most experimentally relevant approach, is to consider a particular physical system (light, cold atoms, etc), consider all experimentally available degrees of freedom and resources (number of atoms, energy, time, squeezing, entanglement, ancillary systems, detectors, etc) and come up with the scheme that yields the best sensitivity for the parameter(s) of interest. In doing so, one may simply follow an educated-guess path, e.g. utilize squeezed states which provide reduced noise and hence better sensitivity [17–21], or perform variational and control optimization procedures [22–24].

In this paper, we take a more fundamental approach and focus on identifying metrological protocols that lead to the optimal sensing performance, irrespective of practical aspects of their implementations. This is not to say that we consider idealized, e.g. noiseless scenarios. On the contrary, we take into account imperfections and decoherence that quantum probes experience during the sensing process, but look for the optimal ways, limited only by the laws of quantum mechanics, to exploit their full quantum sensing potential—employing entanglement, quantum error-correction, active feedback, etc. This approach helps to understand the full sensing potential of quantum systems, and indicates space for possible improvements of existing experimental realizations.

From this perspective, a quantum metrological problem is a quantum channel estimation problem, where the sensing process is represented by an action of a parameter(s) dependent quantum channel Λ_{φ}. The goal is to identify the optimal states of quantum probes as well as the measurements. This may be a

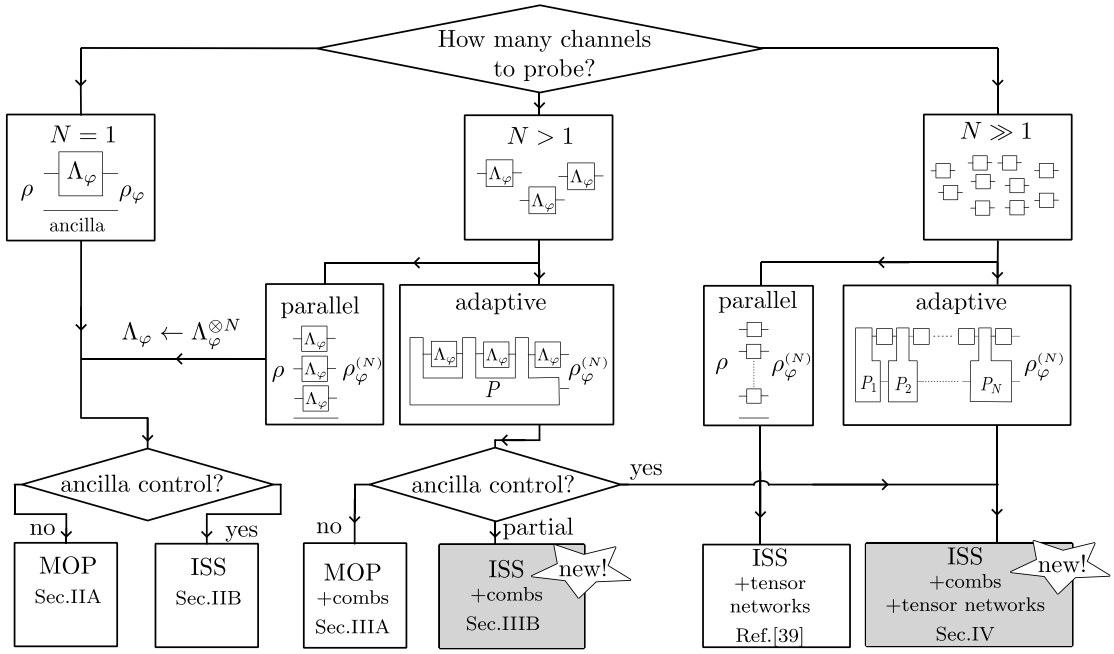


Figure 1. A big-picture view of the main results of the paper. Depending on the number of channels probed (N), different methods of increasing complexity should be applied. For the single channel estimation case ($N = 1$), both the minimization over purifications (MOP) method as well as the iterative see-saw (ISS) method may be directly applied. In case of relatively small N (in practice $N \leqslant 5$ for qubit channels) both methods may be generalized to allow for the search of the optimal adaptive strategies utilizing the concept of quantum combs. For larger N one is forced to employ the tensor networks techniques in order to avoid the curse of dimensionality problem, in which case the ISS method is ideally suited. Note that in all cases, the ISS method additionally allows for an explicit control of the size of ancillary system. The bottom row of boxes serves as a guideline, indicating in which section of the paper a given method is described, with items representing the original contribution of this paper highlighted.

reasonably easy task in case of idealized noiseless models [25–27], but becomes challenging in case of more realistic models that take into account noise and experimental imperfections [28–31].

Fortunately, effective iterative see-saw (ISS) algorithms have been proposed that work both in the quantum Fisher information (QFI) optimization paradigm [32, 33] as well as in the Bayesian one [10, 34, 35]. Furthermore, an approach of computing QFI in noisy models via minimization over purifcations (MOP) of quantum states provides an alternative method to find the optimal protocols in the form of a single semi-definite programme (SDP) [5, 8, 9], see section 2 for extensive discussion.

These approaches are effective provided the dimensionality of the quantum probe Hilbert space is small. This, in particular, makes them inefficient to use when optimizing protocols involving entangled probes that sense multiple (N) channels in parallel. One way this 'curse of dimensionality' may be overcome is via an analysis of a particular educated-guess protocol and proving that it is the optimal one by comparing its performance with the fundamental bounds. This was the way in which the use of squeezed light in lossy optical interferometry has been demonstrated to be asymptotically optimal [6, 8, 36], as well as the use of spin-squeezed states in Ramsey interfereomtry in presence of dephasing [6, 18, 37].

The alternative way to face the problem in the moderate/large N limit is to resort to the tensor network framework [38], where description of many-particle states in the form of matrix product states (MPSs) has been shown to be effective in identifying the optimal states of input probes in quantum metrology [10]—the method to do this is based on an appropriate reformulation of the ISS algorithm in the language of MPS and matrix product operators [39, 40].

From the fundamental point of view, however, the parallel sensing scheme is not the most general way to estimate a parameter encoded in a quantum channel that may be accessed a given number of times (N). In fact, one may consider more general adaptive protocols (e.g. quantum error-correction protocols, etc) that admit all quantum preparation, control and measurement operations that lead to the optimal extraction of information on the parameters encoded in the quantum channels [9, 12, 13, 26, 41]. The mathematical language to describe these protocols is the theory of quantum combs [42], within which a number of numerical methods to find optimal adaptive protocols for small scale problems have been proposed [43–45]. These approaches were primarily based on the MOP approach [44, 45] and as such did not admit a natural way to incorporate efficient tensor network description of the protocols as well as the Bayesian approach.

This paper focuses on the development of efficient methods to identify optimal adaptive protocols in the limit of large/moderate number of channel uses N. In the first step we reconcile the ISS numerical approach with the quantum comb theory, see section 3.2. In the next step, we develop a tensor network approach that allows for an efficient identification of optimal adaptive protocols in the limit of large/moderate N, see section 4. The added benefits of our methods is the ability to control the effective size of ancillary systems, which is not possible in the MOP approach. As a result, the state-of-the-art methods together with the techniques developed in this paper constitute a comprehensive tool-box of numerical methods for quantum metrology, that may be applied irrespective of the type of protocols analyzed and the number of channels sensed. See figure 1 for a big-picture view, and the context of the results presented in this paper.

Apart from these original results, we decided also to provide a comprehensive review of the MOP approaches in a unified framework in sections 2.1 and 3.1 so that the paper has a self-contained character and combines in one place many techniques that were scattered over the literature and never thoroughly discussed and contrasted with each other. These sections are not indispensable for understanding of the remaining content of the paper.

2. Optimal channel estimation

Let us start with a paradigmatic quantum metrological problem of estimating a single parameter φ encoded in the action of a general quantum channel $\Lambda_\varphi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$, where $\mathcal{L}(\mathcal{H})$ represents the set of density matrices acting on Hilbert space $\mathcal{H}$, and we allow for the input and output spaces to be different. The channel is probed by an input state $\rho \in \mathcal{L}(\mathcal{H} \otimes \mathcal{A})$, where $\mathcal{A}$ represents an ancillary system with which the probing system may be entangled and on which the channel acts trivially. The resulting output state reads $\rho_\varphi = \Lambda_\varphi \otimes \mathcal{I}(\rho)$ and is measured using a generalized measurement $\{M_i\}$. The measurement yields result i with probability $p_\varphi(i) = \mathrm{Tr}(\rho_\varphi M_i)$ and the parameter is estimated via an estimator function $\tilde{\varphi}(i)$. The final goal is to identify the protocol where the estimated value is the closest to the true one. The exact form of the cost function depends on the approach taken, whether it is the Bayesian, min–max, or local approach involving unbiased estimators [46–50]. In this paper we will focus on the latter approach, where the cost function is given by the mean squared error

$$\Delta^2\tilde{\varphi} = \sum_i p_\varphi(i)\left[\tilde{\varphi}(i) - \varphi\right]^2 \geqslant \frac{1}{F_{\mathrm{Q}}(\rho_\varphi)}, \tag{1}$$

computed at a certain operating point φ, where the estimator is assumed to be locally unbiased. The advantage of the local approach stems from the fact that, as indicated in (1), the cost, according to the Cramér–Rao bound, may be lower bounded via the inverse of the QFI $F_{\mathrm{Q}}(\rho_\varphi)$, which is only a function of the output state itself [51, 52]. As a result, the problem of identifying the optimal estimation protocol may be reformulated as a problem of maximization of the output state QFI:

$$F_{\mathrm{Q}}(\Lambda_\varphi) = \max_\rho F_{\mathrm{Q}}\left[\Lambda_\varphi \otimes \mathcal{I}(\rho)\right], \tag{2}$$

where $F_{\mathrm{Q}}(\Lambda_\varphi)$ is referred to as the quantum channel QFI [5, 53]. In what follows we will stick to the QFI-based approach as outlined above and comment on the potential extensions to other approaches in the concluding section of the paper.

In case of noisy channels Λ_φ, brute force optimization as specified in equation (2) via general purpose methods quickly becomes extremely inefficient even for low-dimensional systems [28, 29]. This is due to a relatively involved formula for the QFI in case of mixed states

$$F_{\mathrm{Q}}(\rho_\varphi) = \mathrm{Tr}\left(\rho_\varphi L_\varphi^2\right), \quad \dot{\rho}_\varphi = \frac{1}{2}\left\{\rho_\varphi, L_\varphi\right\}, \tag{3}$$

where $\dot{\rho}_\varphi = \partial_\varphi \rho_\varphi$, $\{,\}$ is the anticomutator and L_φ, implicitly defined by the right-hand-side equation above, is the symmetric logarithmic derivative (SLD).

Over the years, two approaches proved particularly effective in solving equation (2) and allowed for identification of the optimal probe states in quantum metrology in regimes out of reach for general purpose optimization methods. For the sake of completeness we provide a concise review of each of these approaches below.

2.1. Minimization over purifications (MOP) method

This method is based on the key observation that the QFI for a mixed state ρ_φ acting on some Hilbert space $\mathcal{H}$ may be equivalently expressed as a minimization of QFI for all admissible purifcations of ρ_φ [5, 6, 8]:

$$F_{\mathrm{Q}}(\rho_\varphi) = \min_{|\Psi_\varphi\rangle} F_{\mathrm{Q}}(|\Psi_\varphi\rangle) = 4 \min_{|\Psi_\varphi\rangle} \langle \dot{\Psi}_\varphi | \dot{\Psi}_\varphi \rangle, \tag{4}$$

where $|\Psi_\varphi\rangle \in \mathcal{H}\otimes\mathcal{R}$ is a purification of ρ_φ, $\rho_\varphi = \mathrm{Tr}_{\mathcal{R}}(|\Psi_\varphi\rangle\langle\Psi_\varphi|)$. The last equality in (4) is due to the fact that the explicit formula for the QFI of a pure state reads: $F_{\mathrm{Q}}(|\Psi_\varphi\rangle) = 4\left(\langle\dot{\Psi}_\varphi|\dot{\Psi}_\varphi\rangle - |\langle\Psi_\varphi|\dot{\Psi}_\varphi\rangle|^2\right)$, and that given a particular purification $|\tilde{\Psi}_\varphi\rangle$ we can always find another one $|\Psi_\varphi\rangle = \mathrm{e}^{\mathrm{i}\xi\varphi}|\tilde{\Psi}_\varphi\rangle$ (where $\xi = -\mathrm{i}\langle\dot{\tilde{\Psi}}_\varphi|\tilde{\Psi}_\varphi\rangle$) yielding the same QFI and additionally satisfying $\langle\Psi_\varphi|\dot{\Psi}_\varphi\rangle = 0$.

Utilizing this fact, we may now rewrite problem (2) in the following equivalent ways [5]:

$$\begin{aligned} F_{\mathrm{Q}}(\Lambda_\varphi) &\overset{\text{(i)}}{=} \max_{|\psi\rangle_{\mathcal{HA}}} F_{\mathrm{Q}}[\Lambda_\varphi \otimes \mathcal{I}(|\psi\rangle\langle\psi|)] \overset{\text{(ii)}}{=} 4 \max_{|\psi\rangle_{\mathcal{HA}}} \min_{|\Psi_\varphi\rangle_{\mathcal{HAR}}} \langle\dot{\Psi}_\varphi|\dot{\Psi}_\varphi\rangle \\ &\overset{\text{(iii)}}{=} 4 \max_{|\psi\rangle_{\mathcal{HA}}} \min_{\{K_{\varphi,k}\}} \langle\psi| \sum_k \dot{K}^\dagger_{\varphi,k}\dot{K}_{\varphi,k} \otimes \mathbb{1}_{\mathcal{A}} |\psi\rangle \overset{\text{(iv)}}{=} 4 \max_{\rho_{\mathcal{H}}} \min_h \mathrm{Tr}\left(\rho_{\mathcal{H}} \sum_k \dot{K}^\dagger_{\varphi,k}(h)\dot{K}_{\varphi,k}(h)\right) \\ &\overset{\text{(v)}}{=} 4 \min_h \max_{\rho_{\mathcal{H}}} \mathrm{Tr}\left(\rho_{\mathcal{H}} \sum_k \dot{K}^\dagger_{\varphi,k}(h)\dot{K}_{\varphi,k}(h)\right) \overset{\text{(vi)}}{=} 4 \min_h \|\alpha(h)\|, \quad \alpha(h) = \sum_k \dot{K}^\dagger_{\varphi,k}(h)\dot{K}_{\varphi,k}(h), \end{aligned} \tag{5}$$

where $\|\cdot\|$ in the final formula is the operator norm. In step (i) we make use of the fact that one may always restrict to pure input states due to convexity of the QFI. In (ii) $|\Psi_\varphi\rangle_{\mathcal{HAR}}$ represents a purification of the channel output state $\rho_\varphi = \Lambda_\varphi \otimes \mathcal{I}(|\psi\rangle\langle\psi|)$ (notice the different roles played by the ancillary system $\mathcal{A}$ and the reference system $\mathcal{R}$). In (iii) we note that an arbitrary purifcation $|\Psi\rangle$ of the state that is obtained by an action of the channel on a pure state, may be written in terms of a certain purification of the quantum channel itself, determined by a particular choice of a Kraus representation of the channel Λ_φ, $|\Psi_\varphi\rangle_{\mathcal{HAR}} = \sum_k K_{\varphi,k}\otimes\mathbb{1}_{\mathcal{A}}|\psi\rangle_{\mathcal{HA}}\otimes|k\rangle_{\mathcal{R}}$, where $|k\rangle_{\mathcal{R}}$ represents some orthonormal basis in $\mathcal{R}$. In fact, since the quantity of interest is local (derivatives are taken at some fixed point φ), it is enough to consider a class of Kraus representations that lead to derivatives of Kraus operators of the form [6, 8]

$$\dot{K}_{\varphi,k}(h) = \dot{K}^{(0)}_{\varphi,k} - \mathrm{i}\sum_l h_{kl} K^{(0)}_{\varphi,l}, \tag{6}$$

where $K^{(0)}_{\varphi,l}$ is some fixed Kraus representation (e.g. canonical) and h_{kl} is an arbitrary hermitian matrix. As a result, the minimization over Kraus representations effectively amounts to a minimization over a single hermitian matrix h_{kl}. In (iv) we observe that effectively the expression depends only on the reduced density matrix $\rho_{\mathcal{H}} = \mathrm{Tr}_{\mathcal{A}}(|\psi\rangle\langle\psi|)$ and not on the whole input state $|\psi\rangle_{\mathcal{HA}}$. This allows us to switch to maximization over arbitrary density matrices $\rho_{\mathcal{H}}$, which unlike pure states, form a convex set. Because of this, we can apply the minimax theorem, and switch the order of minimization and maximization (v), as both sets over which we optimize are convex (h belongs to a linear space, $\rho_{\mathcal{H}}$ belongs to a convex set of density matrices) and the optimized function is convex in h (in fact convex quadratic) and concave in $\rho_{\mathcal{H}}$ (in fact linear). Finally (vi) reflects a property of the operator norm.

Channel QFI as a semi-definite programme

Interestingly, the final variant for the channel QFI optimization problem can be cast as a simple SDP [8]:

$$F_{\mathrm{Q}}(\Lambda_\varphi) = 4 \min_{\lambda,h} \lambda, \text{ subject to } A \succeq 0, \tag{7}$$

where

$$A = \left(\begin{array}{c|ccc} \lambda\mathbb{1}_d & \dot{K}^\dagger_{\varphi,1}(h) & \dots & \dot{K}^\dagger_{\varphi,r}(h) \\ \hline \dot{K}_{\varphi,1}(h) & & & \\ \vdots & & \mathbb{1}_{d\cdot r} & \\ \dot{K}_{\varphi,r}(h) & & & \end{array}\right), \tag{8}$$

where d is the dimension of the space $\mathcal{H}$ and r is the number of canonical Kraus operators of the channel and λ is a real optimization variable. This makes this approach so appealing, as there are many good SDP solvers on the market, providing solutions accompanied by optimality benchmarks [54], see also [55] for introduction to convex optimization and SDP in particular.

Identifying the optimal input probe state

Interestingly, a solution of the above programme yields the desired channel QFI, but in general does not explicitly provide the form of the optimal input probe state $|\psi\rangle_{\mathcal{HA}}$. Inspecting the sequence of equalities in (5), one may only conclude that the reduced density matrix $\rho_{\mathcal{H}}$ corresponding to the optimal input probe state $|\psi\rangle_{\mathcal{HA}}$ should be supported on the subspace spanned by eigenvectors of the optimal $\alpha(h)$ corresponding to the largest absolute eigenvalues. Only if this subspace is one-dimensional this uniquely singles out the optimal input probe state—in this case it also implies that entanglement between $\mathcal{H}$ and $\mathcal{A}$ is not required to obtain the optimal QFI.

The potential loss of information about the optimal input probe state is due to the min–max order change in step (v) in (5). Let $(\rho^{\diamond}_{\mathcal{H}}, h^{\diamond})$ be the optimal solution in (iv). On the other hand, let $(\rho_{\mathcal{H}}, h)$ be the solution after changing the min–max order in (v). Even though, by minimax theorem $\mathrm{Tr}[\rho_{\mathcal{H}}\alpha(h)] = \mathrm{Tr}[\rho^{\diamond}_{\mathcal{H}}\alpha(h^{\diamond})]$, the resulting 'optimal' $\rho_{\mathcal{H}}$ will not necessarily correspond to the reduced density matrix of the optimal input state for the actual metrological task $\rho^{\diamond}_{\mathcal{H}}$ [13]. The problem is due to the fact that in the optimal solution of the min–max problem the figure of merit $\mathrm{Tr}[\rho_{\mathcal{H}}\alpha(h)]$ can no longer in general be interpreted as the QFI of the corresponding state $\rho_{\mathcal{H}}$, as for a given $\rho_{\mathcal{H}}$ this quantity is not minimized over h. In order to remedy this, one needs to make sure that one is exactly at the saddle point of the function that appears under min–max, as then we are sure that for the optimal state found the corresponding $h^{\diamond}$ minimizes the actual figure of merit and can be regarded as the QFI of the purification of $\rho_{\mathcal{H}}$. Such a point always exists, and can be identified by solving the following problem: find $\rho_{\mathcal{H}} \geqslant 0$, $\mathrm{Tr}(\rho_{\mathcal{H}}) = 1$ such that [13]:

$$\mathrm{Tr}\left[\rho_{\mathcal{H}}\alpha\left(h^{\diamond}\right)\right] = \left\|\alpha\left(h^{\diamond}\right)\right\|, \quad \nabla_{h=h^{\diamond}}\mathrm{Tr}\left[\rho_{\mathcal{H}}\alpha\left(h\right)\right] = 0, \tag{9}$$

where $h^{\diamond}$ is obtained from solving (7). This is again a simple SDP programme and any purification $|\psi\rangle_{\mathcal{HA}}$ of such a $\rho_{\mathcal{H}}$ will correspond to the optimal input state.

2.2. Iterative see-saw (ISS) approach

The MOP method is very powerful in identifying the optimal QFI and the corresponding optimal input probe state, but has at least two drawbacks, which prompt to look for alternative approaches. The first drawback is that it is designed only for the optimization of the QFI as a figure of merit and hence is not applicable in Bayesian [35] or minimax [47] (sic!) analysis. Hence, it may not be sufficient to identify optimal protocols in single-shot or finite resources regimes [14, 48, 50, 56]. The second drawback stems from the generality of the approach, which makes it inefficient when analyzing large dimensional systems, or protocols involving multiple uses of the channels—see section 3. In this approach it is in particular not possible to impose restrictions on the dimension of the ancillary system used. It is also not suitable for implementation of tensor network based protocols that allow for an efficient modelling of multiple-probe systems, as well as multiple-round adaptive strategies that avoid the 'curse of dimensionality issue'—see section 4.

The alternative is the ISS procedure, that was first proposed to be used in the Bayesian phase estimation problem [34, 35], then was generalized to deal with QFI optimization problems [32] and now is becoming popular in a broad range of metrological optimization tasks [33, 57]. We will present it first in the context of QFI optimization, as this will be in fact the main focus of this paper, and only then state its Bayesian variant for completeness. For conciseness we will write Λ_{φ} instead of more general $\Lambda_{\varphi} \otimes \mathcal{I}$, which does not mean that ancilla is in general not relevant in the optimization of metrological protocols, but rather than it may always be explicitly included in the definition of the channel—this is in fact the reason why this approach, unlike the purification based method, allows to explicitly analyze the role of the ancilla and in particular its dimensionality.

Let us start by considering the following 'pre-QFI' function:

$$F(\rho, L) = 2\mathrm{Tr}\left(\dot{\rho}_{\varphi} L\right) - \mathrm{Tr}\left(\rho_{\varphi} L^2\right), \tag{10}$$

where $\rho_{\varphi} = \Lambda_{\varphi}(\rho)$. Maximization of the above function over Hermitian operators L (since the SLD is necesserily Hermitian) yields the QFI for a given input ρ [32], where the corresponding optimal $L^{\diamond}$ is in fact the SLD operator as given in (3). This implies that we can write the channel QFI in the form of a double maximization problem:

$$F_{\mathrm{Q}}\left(\Lambda_{\varphi}\right) = \max_{\rho, L} F(\rho, L)\,. \tag{11}$$

This form prompts a very efficient iterative approach. We start with some random input state $\rho^{[0]}$, for which we maximize $F(\rho^{[0]}, L)$ over L to obtain $L^{[0]}$. Then fixing L we optimize $F(\rho, L^{[0]})$ over ρ to obtain $\rho^{[1]}$. We repeat this iterative procedure until $F(\rho^{[i]}, L^{[i]})$ converges (e.g. does not increase by more than 0.01% over five subsequent iteration steps)—the convergence to the optimal QFI value is guaranteed in generic cases , see

[32] for the argument (to be on a safe side one should avoid choosing non-generic input states in the first step—states that are restricted to some subspace, or have certain symmetry).

The step where we find the optimal $L^{[i]}$ given $\rho^{[i]}$ amounts just to solving the linear equation for the SLD (3), where $\rho_\varphi = \Lambda_\varphi(\rho^{[i]})$—note that it is not advisable here to use a formula involving eigendecomposition of ρ_φ, but rather directly solve the linear equation or find L as a hermitian matrix maximizing (10), which can be formulated as an SDP.

In order to perform the complementary step, note that we can rewrite the pre-QFI function as:

$$F(\rho, L) = 2\mathrm{Tr}\left(\rho\dot{\Lambda}^*_\varphi(L)\right) - \mathrm{Tr}\left(\rho\Lambda^*_\varphi\left(L^2\right)\right), \tag{12}$$

where $\Lambda^*_\varphi(\cdot) = \sum_k K_k^\dagger \cdot K_k$ represents the dual map to Λ_φ. This implies that maximizing $F(\rho, L)$ over ρ for a fixed L amounts to solving the following problem:

$$\max_\rho F(\rho, L) = \max_\rho \mathrm{Tr}(\rho M), \; M = 2\dot{\Lambda}^*_\varphi(L) - \Lambda^*_\varphi\left(L^2\right), \tag{13}$$

with standard constraints on the input state $\rho \geqslant 0$, $\mathrm{Tr}(\rho) = 1$. This is a simple SDP programme, for which the solution can be written explicitly as $\rho = |\psi^+\rangle\langle\psi^+|$, where $|\psi^+\rangle$ is the eigenvector of M corresponding to its largest eigenvalue.

There are a number of variations of the algorithm, where one may restrict the set of allowed input states ρ to some convex set, or fix the measurement, in which case the optimized quantity is the classical Fisher information (FI) [32]. Most importantly, the procedure may be easily adapted to minimize the Bayesian quadratic cost with arbitrary prior distribution for the estimated parameter $p(\varphi)$, which we briefly review below.

The average Bayesian quadratic cost for estimating parameter φ is defined as

$$\overline{\Delta^2\tilde{\varphi}} = \int \mathrm{d}\varphi\, p(\varphi) \sum_i p_\varphi(i) \left[\tilde{\varphi}(i) - \varphi\right]^2 \tag{14}$$

and corresponds to (1) averaged over the prior. Minimization of the cost over measurements, estimators and input states results in the formula for the minimal Bayesian cost of estimating a parameter of the channel in the form [10, 35]:

$$\overline{\Delta^2\tilde{\varphi}}(\Lambda_\varphi) = \Delta^2\varphi - \max_{\rho,\bar{L}} \mathrm{Tr}\left(2\bar{\rho}'\bar{L} - \bar{\rho}\bar{L}^2\right), \tag{15}$$

where $\Delta^2\varphi$ represents variance of the prior distribution $p(\varphi)$, $\bar{\rho} = \int \mathrm{d}\varphi\, p(\varphi)\rho_\varphi$ is the output state averaged with respect to the prior, whereas $\bar{\rho}' = \int \mathrm{d}\varphi\, p(\varphi)(\varphi - \bar{\varphi})\rho_\varphi$ with $\bar{\varphi}$ being the prior expectation value of φ. Comparing the above formula with (10), (11) we see that an analogous iterative optimization procedure may now be applied. With fixed ρ, the search for the corresponding optimal $\bar{L}$ amounts to solving the SLD-like equation, (3), with ρ_φ replaced by $\bar{\rho}$ and $\dot{\rho}_\varphi$ replaced by $\bar{\rho}'$. On the other hand, with fixed $\bar{L}$ the search for the optimal input state ρ amounts to the search for the eigenstate corresponding to the largest eigenvalue of

$$\bar{M} = \int \mathrm{d}\varphi\, p(\varphi)\, \Lambda^*_\varphi\left[2(\varphi - \bar{\varphi})\bar{L} - \bar{L}^2\right]. \tag{16}$$

This shows how versatile the see-saw method is, as it can equally well address two conceptually different estimation problems.

3. Optimal channel estimation with multiple coherent uses

In the previous section we have presented efficient methods to identify optimal metrological protocols focusing on a single use of a quantum channel. In quantum metrology, however, we most typically face a situation where quantum channels may be utilized multiple times. Physically, this represents situations when we are able to utilize many quantum probing systems simultaneously (e.g. multiple atoms sensing common magnetic field, multiple photons travelling through the same interferometer, etc), or utilize a single quantum system to perform a coherent sensing of the same environment over extended periods of time (single photon bouncing multiple-times in a cavity, single spin experiencing the same magnetic field over long times, etc) or a combination of both. In order to understand the fundamental potential of quantum metrology, one should be able to identify the optimal protocols which lead to the best estimation of quantum channel parameters, for a given number of channel uses, assuming any kind of quantum control is allowed.

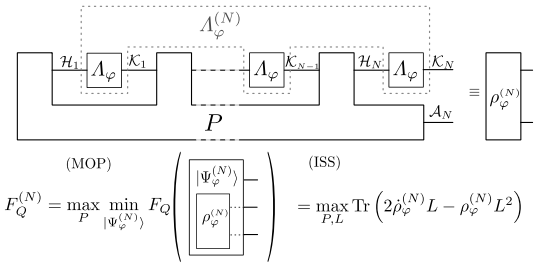


Figure 2. A general scheme of coherent probing of N independent quantum channels Λ_φ (or a general global channel $\Lambda_\varphi^{(N)}$—gray) via an adaptive strategy represented by a quantum comb P. Quantum Fisher information of the output state is optimized over P either via minimization over purifications (MOP) method or the iterative see-saw approach (ISS).

One option is to probe quantum channels in parallel, sending (possibly) entangled states of N probes into $\Lambda_\varphi^{\otimes N}$. Note that this situation is formally equivalent to the one discussed in the previous section, where Λ_φ needs to be replaced by $\Lambda_\varphi^{\otimes N}$. Hence, the methods presented above are valid, and can be applied provided N is not too large, as in this case the curse of dimensionality makes the problem numerically intractable. If this is actually the case, however, one may no longer optimize over arbitrary input states but instead needs to restrict to some reasonable classes of states and compute their performance without invoking the formalism of the full Hilbert space. This can be done using tensor networks methods [39, 40], or by direct Heisenberg picture computations of the performance of particular input probe states and measurement observables [31, 37]. When the performance of the protocols is shown to coincide with some fundamental bounds, this proves the protocols are indeed optimal.

Still, the parallel schemes do not cover the most general quantum adaptive strategies, including quantum control, active quantum feedback, quantum error correction, etc. The problem of the search for the most general quantum adaptive strategy may be formulated as follows. Given N uses of a quantum channel $\Lambda_\varphi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ find the optimal superoperator P, see figure 2, that yields the output state $\rho_\varphi^{(N)} \in \mathcal{L}(\mathcal{K}_N \otimes \mathcal{A}_N)$ with maximal possible QFI (or minimizing the corresponding Bayesian cost if one follows the Bayesian approach). In what follows, we will denote superoperators with non-italic font P, while its italic variant P will represent the corresponding Choi–Jamiołkowski (CJ) operator [58], i.e.:

$$P = \mathrm{P} \otimes \mathcal{I}_{\mathcal{H}_{\mathrm{in}}}\left(|\Phi\rangle\langle\Phi|\right), \tag{17}$$

where $\mathcal{I}_{\mathcal{H}_{\mathrm{in}}}$ is the identity operator on tensor product of all input spaces P (in our case $\mathcal{H}_{\mathrm{in}} = \mathcal{K}_1 \otimes \ldots \mathcal{K}_{N-1}$) while $|\Phi\rangle = \sum_i |i\rangle \otimes |i\rangle$ is a non-normalized maximally entangled state on $\mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{\mathrm{in}}$, where $|i\rangle$ is an orthonormal basis in $\mathcal{H}_{\mathrm{in}}$. In particular, $\Lambda_\varphi \in \mathcal{L}(\mathcal{K} \otimes \mathcal{H})$ will represent the CJ operator of Λ_φ.

The superoperator P represents all possible interaction of the sensing system with arbitrarily large ancillary systems, and allows for any control operations in between subsequent channel uses. Mathematically, P is a linear operator $\mathrm{P} : \mathcal{L}(\mathcal{K}_1 \otimes \ldots \mathcal{K}_{N-1}) \to \mathcal{L}(\mathcal{H}_1 \otimes \ldots \mathcal{H}_N \otimes \mathcal{A}_N)$ that satisfies conditions for being a quantum comb [42] $\mathrm{P} \in \mathrm{Comb}[(\emptyset, \mathcal{H}_1), (\mathcal{K}_1, \mathcal{H}_2), \ldots, (\mathcal{K}_{N-1}, \mathcal{H}_N \otimes \mathcal{A}_N)]$, where pairs of spaces represent respective input/output spaces of each 'tooth' of the comb. In terms of the corresponding CJ matrix $P \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{K}_1 \otimes \ldots \mathcal{H}_{N-1} \otimes \mathcal{K}_{N-1} \otimes \mathcal{H}_N \otimes \mathcal{A}_N)$ (the ordering of spaces is chosen for notational convenience) the conditions for P being a comb read:

$$\begin{aligned} &P \geqslant 0,\ \mathrm{Tr}_{\mathcal{A}_N \otimes \mathcal{H}_N} P = P^{(N-1)} \otimes \mathbb{1}_{\mathcal{K}_{N-1}}, \\ &\underset{1<k<N}{\forall} \mathrm{Tr}_{\mathcal{H}_k} P^{(k)} = P^{(k-1)} \otimes \mathbb{1}_{\mathcal{K}_{k-1}},\ \mathrm{Tr}_{\mathcal{H}_1} P^{(1)} = 1, \end{aligned} \tag{18}$$

where $P^{(k)}$ represents P traced out over $\mathcal{K}_k \otimes \mathcal{H}_{k+1} \otimes \cdots \otimes \mathcal{K}_{N-1} \otimes \mathcal{H}_N \otimes \mathcal{A}_N$. Intuitively, these conditions are related to the causal structure of adaptive protocols—input $\mathcal{K}_i$ may affect only outputs $\mathcal{H}_{i+1}, \ldots, \mathcal{H}_N$, all previous outputs cannot depend on $\mathcal{K}_i$, see [42] for a further discussion. The final output state $\rho_\varphi^{(N)}$ is now obtained by concatenating P operation with N-fold use of the channel Λ_φ, which mathematically corresponds to application of the link product operation to the corresponding CJ operators [42], defined as follows:

$$\rho_\varphi^{(N)} = \Lambda_\varphi^{\otimes N} \star P. \tag{19}$$

Given two operators $A \in \mathcal{L}(\mathcal{A} \otimes \mathcal{C})$, $B \in \mathcal{L}(\mathcal{C} \otimes \mathcal{B})$, where the common subsystem on which they act is denoted by $\mathcal{C}$, the link product is defined as $A \star B = \mathrm{Tr}_{\mathcal{C}}\left[(A \otimes \mathbb{1}_{\mathcal{B}})(\mathbb{1}_{\mathcal{A}} \otimes B^{T_{\mathcal{C}}})\right]$, where $T_{\mathcal{C}}$ denotes

transposition with respect to subsystem $\mathcal{C}$. Note that when $A = |A\rangle\langle A|$, $B = |B\rangle\langle B|$ are rank-1 operators, so is their link product, as $A \star B = |A \star B\rangle\langle A \star B|$, where $|A \star B\rangle := \sum_c \langle c|A\rangle \otimes \langle c|B\rangle$, with $\{|c\rangle \in \mathcal{C}\}$ representing the basis in $\mathcal{C}$ (the distinguished basis, with respect to which the partial transposition is defined).

The problem of identifying the optimal metrological protocol now amounts to the following optimization task:

$$F_{\mathrm{Q}}^{(N)}(\Lambda_\varphi) = \max_P F_{\mathrm{Q}}\left[\Lambda_\varphi^{\otimes N} \star P\right], \tag{20}$$

with constraints on P given in (18). When compared with the single channel estimation problem, (2), the only difference amounts to the replacement of the input state ρ with the quantum comb P. This problem can again be approached using either the MOP or the ISS method. In the discussion below, we will consider an even more general scenario, where we replace N independent uses of a channel Λ_φ by an arbitrary N-teeth quantum comb $\Lambda_\varphi^{(N)}$—mathematically this amounts to replacing $\Lambda_\varphi^{\otimes N}$ by $\Lambda_\varphi^{(N)}$. This will allow us to also discuss the models where the probed channels are subject to correlated noise, see figure 2 and section 5.5.

3.1. Minimization over purifications (MOP) method

Since the set of quantum combs is convex and the output state $\rho_\varphi^{(N)}$ is a linear function of P, the convexity of QFI implies that we may restrict ourselves to extremal quantum combs at the input. Unlike in the single-channel estimation case, an extremal comb is not necessarily pure, i.e. a CJ operator of rank one $P = |P\rangle\langle P|$, due to non-trivial constraints (18). Still, we may always purify it at the expense of possibly increasing the dimension of the ancillary system $\mathcal{A}_N$. Hence, we may now follow an analogous procedure as in (5), by noticing that the minimization of the QFI formula over different purifications of $\rho_\varphi^{(N)}$ can again be understood in terms of minimization over different Kraus representations of $\Lambda_\varphi^{(N)}$. Let $\{K_{\varphi,k}^{(N)}\}$ be a Kraus representation of $\Lambda_\varphi^{(N)}$, which can be written in terms of the decomposition of the corresponding CJ matrix $\Lambda_\varphi^{(N)} = \sum_k |K_{\varphi,k}^{(N)}\rangle\langle K_{\varphi,k}^{(N)}|$, where $|K_{\varphi,k}^{(N)}\rangle \in \mathcal{K}_1 \otimes \mathcal{H}_1 \otimes \cdots \otimes \mathcal{K}_N \otimes \mathcal{H}_N$ represents a vectorized Kraus operator. Each Kraus representation may be associated with the following purification $|\Psi_\varphi^{(N)}\rangle_{\mathcal{K}_N\mathcal{A}_N\mathcal{R}} = \sum_k |K_{\varphi,k}^{(N)} \star P\rangle \otimes |k\rangle_{\mathcal{R}} \in \mathcal{K}_N \otimes \mathcal{A}_N \otimes \mathcal{R}$, where we have used the notation for the link product of rank-1 operators. With this notation, we can now adapt (5) in order to derive the formula for the optimal QFI:

$$\begin{aligned} F_{\mathrm{Q}}\left(\Lambda_\varphi^{(N)}\right) &= 4\max_P \min_{|\Psi_\varphi^{(N)}\rangle_{\mathcal{K}_N\mathcal{A}_N\mathcal{R}}} \langle \dot{\Psi}_\varphi^{(N)}|\dot{\Psi}_\varphi^{(N)}\rangle = 4\max_P \min_{\left\{K_{\varphi,k}^{(N)}\right\}} \sum_k \langle \dot{K}_{\varphi,k}^{(N)} \star P|\dot{K}_{\varphi,k}^{(N)} \star P\rangle \\ &= 4\max_P \min_{\left\{K_{\varphi,k}^{(N)}\right\}} \mathrm{Tr}\left(\sum_k |\dot{K}_{\varphi,k}^{(N)}\rangle\langle \dot{K}_{\varphi,k}^{(N)}| \star P\right) = 4\max_P \min_h \mathrm{Tr}\left[\left(\Omega(h) \otimes \mathbb{1}_{\mathcal{A}_N}\right)\left(P \otimes \mathbb{1}_{\mathcal{K}_N}\right)\right] \\ &\overset{\text{(i)}}{=} 4\max_{\tilde{P}} \min_h \mathrm{Tr}\left[\tilde{\Omega}(h)\tilde{P}\right] \overset{\text{(ii)}}{=} 4\min_h \max_{\tilde{P}} \mathrm{Tr}\left[\tilde{\Omega}(h)\tilde{P}\right], \end{aligned} \tag{21}$$

where $\Omega(h) = \sum_k |\dot{K}_{\varphi,k}^{(N)}(h)\rangle\langle \dot{K}_{\varphi,k}^{(N)}(h)|^{\mathrm{T}}$, while $|\dot{K}_{\varphi,k}^{(N)}(h)\rangle$ is defined analogously as in (6), but this time h is in principle a huge matrix, as the number of Kraus operators $K_{\varphi,k}^{(N)}$ will typically grow exponentially with N. In step (i) we performed the partial trace over spaces $\mathcal{A}_N$ and $\mathcal{K}_N$ and introduced $\tilde{P} = \mathrm{Tr}_{\mathcal{A}_N} P$, $\tilde{\Omega}(h) = \mathrm{Tr}_{\mathcal{K}_N}\Omega(h) \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{K}_1 \otimes \ldots \mathcal{H}_{N-1} \otimes \mathcal{K}_{N-1} \otimes \mathcal{H}_{\mathcal{N}})$, while in (ii) we again used the minimax theorem, as optimization over both h and $\tilde{P}$ is over convex spaces and the function itself is convex in h and concave (linear) in $\tilde{P}$. Note that in the special case $N = 1$, $\tilde{\Omega}(h) = \alpha(h)$, $\tilde{P} = \rho_{\mathcal{H}}$ and we recover the formula from (5). Unlike in (5), however, we cannot replace the maximization over the comb by the operator norm due to nontrivial constraints on $\tilde{P}$. Nevertheless, if maximization over $\tilde{P}$ is replaced by an appropriate minimization of the dual problem, one may in the end write the double minimization as a single SDP in a form resembling that of the single channel optimization (7) [45]:

$$\begin{aligned} &F_{\mathrm{Q}}\left(\Lambda_\varphi^{(N)}\right) = 4 \min_{\lambda,h,Q^{(k)}} \lambda, \text{ subject to } A \succeq 0, \\ &\forall_{2\leqslant k\leqslant N-1} \mathrm{Tr}_{\mathcal{K}_k} Q^{(k)} = \mathbb{1}_{\mathcal{H}_k} \otimes Q^{(k-1)}, \mathrm{Tr}_{\mathcal{K}_1} Q^{(1)} = \mathbb{1}_{\mathcal{H}_1} \end{aligned} \tag{22}$$

where

$$A = \left(\begin{array}{c|ccc} \mathbb{1}_{\mathcal{H}_N} \otimes Q^{(N-1)} & |\dot{\tilde{K}}_{1,1}(h)\rangle & \ldots & |\dot{\tilde{K}}_{r,d}(h)\rangle \\ \hline \langle \dot{\tilde{K}}_{1,1}(h)| & & & \\ \vdots & & \lambda\mathbb{1}_{dr} & \\ \langle \dot{\tilde{K}}_{r,d}(h)| & & & \end{array}\right). \tag{23}$$

In the above expression, $d = \dim(\mathcal{K}_N)$, $|\dot{\tilde{K}}_{i,k}(h)\rangle = {}_{\mathcal{K}_N}\langle i|\dot{K}^{(N)}_{\varphi,k}(h)\rangle \in \mathcal{K}_1 \otimes \mathcal{H}_1 \otimes \ldots \mathcal{K}_{N-1} \otimes \mathcal{H}_{N-1} \otimes \mathcal{H}_N$, while dual problem variables $Q^{(k)} \in \mathcal{L}(\mathcal{K}_1 \otimes \mathcal{H}_1 \otimes \cdots \otimes \mathcal{K}_k \otimes \mathcal{H}_k)$ are subject to the same quantum-comb constraints, apart from the positivity requirement.

The above optimization problem will correctly yield the optimal QFI $F^{\diamond}_{\mathrm{Q}}$ of the channel and the corresponding $h^{\diamond}$ matrix. If, however, one wants to identify the corresponding quantum comb $P^{\diamond}$ that represents the optimal strategy, one needs to step back to the original primal problem formulation, and find $\tilde{P}$ satisfying the quantum comb constraints such that [45]:

$$4\mathrm{Tr}\left[\tilde{\Omega}\left(h^{\diamond}\right)\tilde{P}\right] = F^{\diamond}_{\mathrm{Q}}, \quad \nabla_{h=h^{\diamond}}\mathrm{Tr}\left[\tilde{\Omega}\left(h\right)\tilde{P}\right] = 0. \tag{24}$$

Similarly as in (9) the second condition is necessary to make sure the solution is at the saddle point and the resulting $\tilde{P}^{\diamond}$ corresponds indeed to the optimal protocol. Note that in the original paper [44] this second condition was not included and hence the procedure described there might not lead to the actual optimal protocol—it has only been remedied in the follow-up paper [45].

3.2. Iterative see-saw (ISS) approach

As a completely new result, we demonstrate how the ISS optimization can be generalized to the multiple coherent uses regime—it turns out to be more straightforward than generalization of the MOP method. The generalization of the 'pre-QFI' function from (10), to be maximized, now takes the form:

$$F^{(N)}\left(P, L\right) = 2\mathrm{Tr}\left(\dot{\rho}^{(N)}_{\varphi} L\right) - \mathrm{Tr}\left(\rho^{(N)}_{\varphi} L^2\right), \tag{25}$$

where $\rho^{(N)}_{\varphi} = \Lambda^{(N)}_{\varphi} \star P$. The optimal QFI is again obtained as a result of double maximization:

$$F_{\mathrm{Q}}\left(\Lambda^{(N)}_{\varphi}\right) = \max_{P,L} F^{(N)}\left(P, L\right), \tag{26}$$

with constraints on P to be a quantum comb as given in (18). We start the iteration procedure with some randomly chosen initial $P^{[0]}$ and find the corresponding $L^{[0]}$ maximizing $F^{(N)}(P^{[0]}, L)$. This step is identical to the one in the single channel estimation approach and amounts to finding the SLD for the state $\rho^{(N)}_{\varphi}$. Then fixing $L^{[0]}$ we identify the optimal $P^{[1]}$ and so on. This second step, while slightly more complex than in the single channel case (13), can nevertheless again be written as a relatively simple SDP programme:

$$\begin{aligned} &\max_{P} F^{(N)}\left(P, L\right) = \max_{P} \mathrm{Tr}\left(P M^{(N)}\right), \\ &M^{(N)} = 2L \star \left(\dot{\Lambda}^{(N)}_{\varphi}\right)^{\mathrm{T}} - L^2 \star \left(\Lambda^{(N)}_{\varphi}\right)^{\mathrm{T}}, \end{aligned} \tag{27}$$

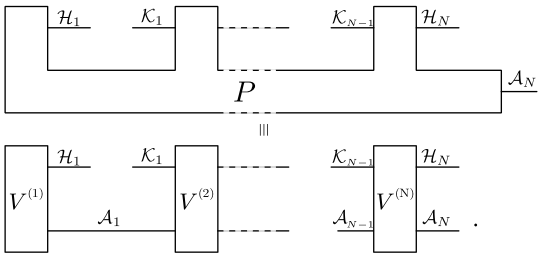
where we used straightforward properties of the link product operation to rewrite (25) in the above form. This is a linear optimization problem in P with convex constraints (18), hence an SDP. In each step of the iteration $F^{(N)}$ will not decrease, and we terminate the procedure when it converges up to the desired accuracy.

Similarly as in the MOP method, this procedure is efficient provided the dimensions of spaces on which channel $\Lambda^{(N)}_{\varphi}$ acts are reasonable—for multiple uses of a single qubit channel this usually means $N \leqslant 5$. One of advantages of the ISS approach over the MOP is that we can control the size of the final ancillary space $\mathcal{A}_N$ by restricting optimization over P to combs with a fixed dimension of $\mathcal{A}_N$ (but without the control of the size of ancillary systems required for the inner action of the comb itself). The main advantage, however, is that the approach can be naturally adapted to the tensor network formalism, as described in section 4 and in many cases allows to go around the curse of dimensionality problem and study optimal metrological protocols in the limit of large number of coherent channel uses. In this approach it will also be possible to control the size of ancilla at every step of the protocol.

3.3. Decomposition of a quantum comb into elementary operations

Given the CJ operator P of the optimal quantum comb, obtained from either the MOP or ISS method, it may be non-trivial to obtain a deeper insight into the character of the metrological protocol it actually represents and how to implement it in practice using a set of simple quantum gates. For this purpose procedures of decomposition of a quantum comb into a sequence of unitary operations (isometries) has been proposed in [42, 59]. A quantum comb P can be always rewritten as a concatenation of isometries $\{V^{(k)}\}_{k=1}^{N}$ corresponding to the comb's consecutive teeth, s.t. $V^{(k)} : \mathcal{K}_{k-1} \otimes \mathcal{A}_{k-1} \to \mathcal{H}_k \otimes \mathcal{A}_k$ (where $\mathcal{A}_{k-1}, \mathcal{A}_k$ are the

ancillae on the input and output of the kth tooth, respectively and $\mathcal{K}_0, \mathcal{A}_0$ are trivially $\mathbb{C}$). The input and output spaces of the isometries can schematically be pictured as follows:



(28)

The details on how to find the corresponding isometries are presented in appendix A.

However, constructing these isometries in general will require ancillary systems of dimensions that may grow exponentially with k. The minimal possible dimension of ancilla $\mathcal{A}_k$ necessary to represent $V^{(k)}$ is $\dim(\mathcal{A}_k) = \mathrm{rank}(P^{(k)})$ [59]. Rank of $P^{(k)}$ may be as large as the dimension of the space on which the operator acts, which in this case is the product of dimensions of all its input and output spaces $\mathcal{K}_{i-1}, \mathcal{H}_i$, up to $i = k$. From our numerical experience it appears that in typical metrological scenarios, the isometries corresponding to optimal metrological strategies in noisy models obtained from comb optimization methods described above are very complex, and the size of ancillary systems indeed tends to grow significantly with k—see the discussion of an example in section 5.4.

Another challenge is that even if the optimal strategy may not be that complex, it may be returned by the optimization procedure in a basis which is not easy to interpret. In principle one could use automated methods, implemented in e.g. BQSKit [60] , UniversalQCompiler [61] of decomposition of isometries into gates from standard gate sets to obtain a practical implementation of the optimal protocols. This approach to describing optimal metrological protocols was taken e.g. in [45]. However, the challenge of interpreting those schemes still remains.

These difficulties in obtaining a simple and intuitive structure of the optimal protocols are related to substantial freedom when identifying optimal quantum combs. The optimal comb found is typically not unique and its form rarely allows for a direct intuitive understanding of the essence of the protocol. This is another strong argument (apart from the curse of dimensionality issue) in favour of the tensor network approach presented in the next section, where we are able to control the size of the ancillary systems at each 'tooth' of the comb, and have means to force the optimization procedure, tooth by tooth, to yield optimal strategies in a form we are able to interpret. Note that by nature of the approach, this is not possible in the MOP-type methods.

4. Tensor network approach to identify the optimal adaptive metrological protocols

The main obstacle in finding the optimal adaptive protocols is the exponential growth of complexity of algorithms presented in section 3 with increasing N. This is related to the size of the CJ operator P representing the estimation strategy comb. Even in the simplest case of qubit channels Λ_φ and no output ancilla, P acts on a Hilbert space of total dimension 2^{2N-1}. This limits the applicability of introduced methods on present-day personal computers to the cases where $N \leqslant 5$—for larger N, SDP memory and time requirements are hard to meet.

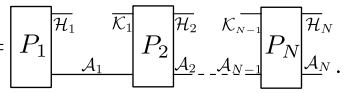
Moreover, in the approaches presented so far, the internal structure of P cannot be controlled—in particular, we cannot limit the size of the ancillary system required to implement each tooth of the optimal comb, we can only control the last ancilla $\mathcal{A}_N$. Consequently, the strategies obtained may be very complex even for small N.

To overcome both of these problems, we decompose P into teeth $P_1, P_2, \ldots, P_N$, representing simpler quantum channels whose concatenation leads to P. Formally,

$$P = P_1 \star P_2 \star \ldots \star P_N, \tag{29}$$

where $P_1 \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{A}_1)$, $P_k \in \mathcal{L}(\mathcal{A}_{k-1} \otimes \mathcal{K}_{k-1} \otimes \mathcal{H}_k \otimes \mathcal{A}_k)$ for $2 \leqslant k \leqslant N$—when the link product between P_k and P_{k+1} is performed the common subspace is $\mathcal{A}_k$. Importantly, P_i can be arbitrary CJ matrices, not necesarily isometries, contrary to the comb decomposition described in section 3.3. This allows for effective optimization over each comb tooth P_i, since CJ matrices, unlike isometries, form a convex set. This also changes the meaning of ancillary spaces $\mathcal{A}$—in section 3.3 they played dual role of a comb memory and purification, whereas in (29) we do not need to purify P_i, so ancillary system is only used as a quantum memory.

The link product is linear in both arguments, and can be represented as contraction of indices between two tensors (see appendix B for short introduction to tensor networks formalism and technical details). Consequently, we can write down the RHS of (29) as a tensor network, whose nodes (rectangles) represent quantum channels, open links represent subspaces on which P acts and closed links represent link products:

$$P = \tag{30}$$


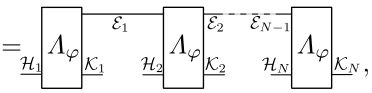
Notice that with this graphic notation it is clear which subspaces must be contracted while performing the link product.

When output and input spaces of probed channels are all of the same dimension ($\dim(\mathcal{H}_k) = \dim(\mathcal{K}_k) = d_\mathcal{H}$), and the size of ancilla is fixed during the whole protocol ($\dim(\mathcal{A}_k) = d_\mathcal{A}$), then $d_\mathcal{H}^{4N-2} d_\mathcal{A}^2$ complex variables are required to store P in the memory. We can substantially compress the information about P by storing $P_1, P_2, \ldots, P_N$ separately, which requires only $d_\mathcal{H}^2 d_\mathcal{A}^2 + (N-1) d_\mathcal{H}^4 d_\mathcal{A}^4$ variables. The latter approach is much more effective for fixed $d_\mathcal{A}$ and growing N—the used memory scales linearly, not exponentially with N. To take advantage of this, it is crucial to design an optimization algorithm that operates on $P_1, P_2, \ldots, P_N$ separately, and does not need to refer to the whole P. This approach is completely different than the one from section 3.3—instead of finding optimal P and decomposing it, we use ansatz (29) from the very beginning to overcome the curse of dimensionality.

The compression of P is only possible for combs which can be written as (29) with small $d_\mathcal{A}$. In general, the size of ancillary system required to simulate all possible combs grows exponentially with N, and so does the size of CJ operators P_k. However, it makes sense to search only through strategies P with limited $d_\mathcal{A}$ because such strategies are usually substantially easier to implement in practice. One may also optimize the QFI over P with growing $d_\mathcal{A}$—when the figure of merit no longer increases with increasing $d_\mathcal{A}$ it strongly suggests that the optimal strategy P has been found.

The analogous idea is used in other tensor networks approximations, in particular in the construction of MPS representation of entangled states of N particles [38, 62]. The size of a density matrix of such a system grows exponentially with N. However, states that are weakly entangled can be efficiently represented as a tensor network (MPS). Then, the memory required to store the information about a state grows linearly with N, but also depends on a so-called bond dimension d. The more entangled the state is, the larger d is required. In our case, $d_\mathcal{A}$ plays a role of a bond dimension, with a clear physical interpretation of the size of available ancillary system.

The CJ operator $\Lambda_\varphi^{(N)}$, containing information about the estimated signal and (possibly correlated) noise, can be also represented as a tensor network. Using the introduced graphical notation:

$$\Lambda_\varphi^{(N)} = \tag{31}$$


where $\mathcal{E}_k$ represents the environment space after the kth use of the channel to be estimated. When the signal and the noise are not correlated, then $\Lambda_\varphi^{(N)} = \Lambda_\varphi^{\otimes N}$, the state of environment does not affect the action of the subsequent channel, and consequently $\mathcal{E}_k$ links may be ignored—the tensor network representing $\Lambda_\varphi^{\otimes N}$ is a trivial network without connections. Different types of correlations can be simulated using non-trivial action of channels on $\mathcal{E}_k$—see section 5.5 for an example involving a correlated dephasing noise model. The more complicated and long-range correlations are, the larger the dimension of $\mathcal{E}_k$ required to simulate them. Again, to simulate all possible signal combs $\Lambda_\varphi^{(N)}$, one would need $\dim(\mathcal{E}_k)$ that grows exponentially with N. Luckily, for many typical correlation models, the required $\dim(\mathcal{E}_k)$ does not depend on N at all.

The problem of identifying the optimal QFI for a sequence of (possibly correlated) channels Λ_φ, with limited size of ancillary system $d_\mathcal{A}$, can be written using ISS approach as

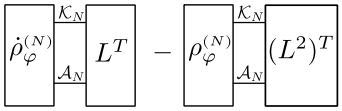
$$F_\mathrm{Q}^{(d_\mathcal{A})}\left(\Lambda_\varphi^{(N)}\right) = \max_{P_1, P_2, \ldots, P_N, L} F^{(N)}(P, L), \tag{32}$$

where P is given by (29) with links $\mathcal{A}$ of dimensions $d_\mathcal{A}$ and $F^{(N)}(P, L)$ is defined as in (25). Obviously, $F_\mathrm{Q}^{(d_\mathcal{A})}(\Lambda_\varphi^{(N)}) \leqslant F_\mathrm{Q}(\Lambda_\varphi^{(N)})$, and $F_\mathrm{Q}^{(d_\mathcal{A})}(\Lambda_\varphi^{(N)}) = F_\mathrm{Q}(\Lambda_\varphi^{(N)})$ for sufficiently large $d_\mathcal{A}$.

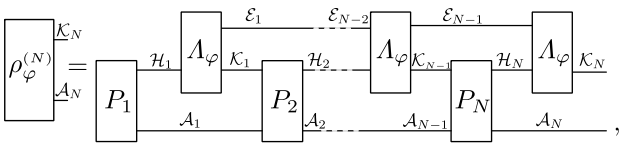
To perform the maximization (32) numerically, we proceed as follows. Initially, $P_1, P_2, \ldots, P_N$ are random CJ operators and L is a random hermitian matrix. Then, we maximize the figure of merit over P_1, fixing $P_2, \ldots, P_N, L$. In the next step, we maximize over P_2, then over P_3, etc. In the final step, the maximization over L is performed. The entire procedure is then repeated until convergence is achieved. Notice, that compared to

the algorithm from section 3.2, we need to perform $N+1$ maximization steps in one iteration instead of just 2 steps. However, the computational complexity of each step does not scale with N. It is also crucial that each maximization step is SDP (as we show below).

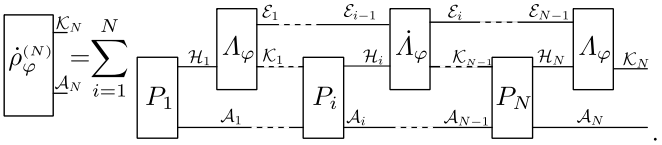
The network representing $F^{(N)}(P,L)$ can be depicted as

$$F^{(N)}(P,L) = 2$$  $$-$$, (33)

where we used the identity $A \star B = \mathrm{Tr}(AB^{\mathrm{T}})$ valid for matrices A, B acting on the same Hilbert space. The matrices $\rho_\varphi^{(N)}$, $\dot{\rho}_\varphi^{(N)}, L$ are all of size $d_{\mathcal{H}} d_{\mathcal{A}} \times d_{\mathcal{H}} d_{\mathcal{A}}$, which does not depend on N. Hence those matrices can be easily stored in memory. Given $\rho_\varphi^{(N)}$, $\dot{\rho}_\varphi^{(N)}$, the optimization over L can be performed in the same way as it was done in sections 2.2 and 3.2. The nontrivial part is the efficient computation of matrices $\rho_\varphi^{(N)}$, $\dot{\rho}_\varphi^{(N)}$. This can be done by contracting the following networks:



(34)

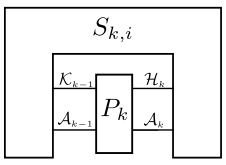


(35)

Notice, that $\dot{\rho}_\varphi^{(N)}$ is represented as a sum of N elements—in the ith element the derivative acts on CJ operator of the ith probed channel Λ_φ.

The key property of tensor networks is the freedom of choice of indices contraction order. Even though $\rho_\varphi^{(N)}$ is constructed by contraction of networks representing P and $\Lambda_\varphi^{(N)}$, it is not necessary to compute these constituent networks in order to obtain $\rho_\varphi^{(N)}$ (which would be equivalent to the procedure from section(3.2)). Instead, one should contract indices in the following order: $\mathcal{H}_1, \mathcal{A}_1, \mathcal{K}_1, \mathcal{E}_1, \mathcal{H}_2, \mathcal{A}_2$, ..., $\mathcal{E}_{N-1}, \mathcal{H}_N$. Then, the maximal size of a tensor we need to process at a time does not depend on N. This allows to compute $\rho_\varphi^{(N)}$ and its derivative efficiently even for large N, and then find optimal L in the same way as in previously described ISS procedures.

To optimize over P_k, we proceed as follows. Firstly, we represent the figure of merit using (33). Then, we replace $\rho_\varphi^{(N)}$ and $\dot{\rho}_\varphi^{(N)}$ with networks (34) and (35). The figure of merit is then represented as a sum of $N+1$ networks: N from the term $2\mathrm{Tr}(\dot{\rho}_\varphi^{(N)} L)$ and 1 from the term $\mathrm{Tr}(\rho_\varphi^{(N)} L^2)$. In each component network, we contract all the indices apart from those corresponding to subspaces linked to P_k. Then, we obtain

$$F^{(N)}(P,L) = \sum_{i=0}^{N}$$  , (36)

where $S_{k,0}$ denotes the contracted network representing term with $\rho_\varphi^{(N)}$ and $S_{k,i}$ for $i>0$ denotes the ith term from the expansion of the term with $\dot{\rho}_\varphi^{(N)}$. CJ operators $S_{k,i}$ act on the Hilbert space of dimension $d_{\mathcal{A}}^2 d_{\mathcal{H}}^2$, and their sum

$$S_k = \sum_{i=0}^{N} S_{k,i} \tag{37}$$

can be directly computed, which allows to write the figure of merit as

$$F^{(N)}(P,L) = \mathrm{Tr}\left(P_k S_k^{\mathrm{T}}\right). \tag{38}$$

Interestingly, when proper partial results of computations are saved, then the time complexity of computing $S_1, S_2, \ldots, S_N$ is $O(N)$, which is also the time complexity of the whole algorithm, see appendix C. The optimization over P_k boils down to the following SDP:

$$\begin{aligned} &\max_{P_k} F^{(N)}(P, L) = \max_{P_k} \mathrm{Tr}\left(P_k S_k^{\mathrm{T}}\right), \\ &\text{s. t. } \mathrm{Tr}_{\mathcal{H}_k \otimes \mathcal{A}_k}(P_k) = \mathbb{1}_{\mathcal{K}_{k-1} \otimes \mathcal{A}_{k-1}}, P_k \geqslant 0. \end{aligned} \tag{39}$$

The first condition for P_k corresponds to the trace preservation of a channel, and should be replaced with $\mathrm{Tr}(P_k) = 1$ for $k = 1$ because P_1 represents a density matrix of an input state.

The performance of the iterative optimization is the most stable when CJ operators Λ_φ are full-rank. In other cases, we observe that convergence to the optimal value is not always achieved. To overcome this issue, we add an artificial depolarizing noise to each channel Λ_φ. The strength of this noise decays exponentially over the course of running the algorithm, such that its role becomes negligible for final iterations, and the optimization result is unaffected. Due to this improvement, our algorithm is very stable, and converges to the same value with different random inputs.

The algorithm outputs not only the optimal QFI value, $F_{\mathrm{Q}}^{(d_{\mathcal{A}})}(\Lambda_\varphi^{(N)})$, but also the sequence of CJ operators P_k corresponding to the optimal estimation protocol. In many cases, this protocol can be further simplified. Let us consider the following transformation of two subsequent comb teeth:

$$\mathrm{P}_k \to (\mathcal{I}_{\mathcal{H}_k} \otimes \mathrm{U}_{\mathcal{A}_k}) \circ \mathrm{P}_k \tag{40}$$

$$\mathrm{P}_{k+1} \to \mathrm{P}_{k+1} \circ \left(\mathcal{I}_{\mathcal{K}_k} \otimes \mathrm{U}_{\mathcal{A}_k}^{\dagger}\right), \tag{41}$$

where $\mathcal{I}$ is an identity channel, U is an arbitrary unitary channel, $\mathrm{U}^{\dagger}$ is its inverse and subscript denotes a space on which a channel acts. Channels $\mathrm{U}_{\mathcal{A}_k}$ and $\mathrm{U}_{\mathcal{A}_k}^{\dagger}$ become identity when the link product $P_k \star P_{k+1}$ is performed—consequently, the described transformation does not change P. The proper choice of $\mathrm{U}_{\mathcal{A}_k}$ may simplify the teeth $P_1, P_2, \ldots, P_k$ of the optimal strategy—this is an analogue of the 'local-gauge' choice in the MPS description [38]. In our algorithm we can fix the initial teeth in an easy to interpret basis, rerun the optimization over the remaining ones and thus, step by step, limit some of the gauge freedom.

5. Examples

To demonstrate the efficiency of the introduced tensor network based approach, we use it to find the optimal adaptive protocols of estimation of different noisy qubit channels and the corresponding QFI in the limit of large number of channel uses. In our numerical implementation to solve SDP problems we used CVXPY package [54] for modeling and MOSEK solver [63].

The methods we have developed can be applied to channels Λ_φ with arbitrary parameter dependence. Still, for concreteness we focus on qubit channel estimation models with unitary parameter encoding and parameter independent noise, for which the Kraus operators have the following structure:

$$K_{\varphi,k} = U_\varphi K_k, \quad U_\varphi = \mathrm{e}^{-\frac{\mathrm{i}}{2}\varphi\sigma_z}, \tag{42}$$

which means that the angle of rotation of the Bloch vector around z axis is estimated, and signal comes after noise described by Kraus operators K_k. This allows us to discuss models that manifest qualitatively different behaviour both in terms of asymptotic QFI scaling, and in terms of the impact of the size of available ancillary system.

We will present results on the achievable QFI for optimal adaptive protocols utilizing given size of ancillary systems for four representative types of uncorrelated noise affecting phase estimation: perpendicular dephasing (5.1), parallel dephasing (5.2), perpendicular amplitude damping (5.3), and parallel damping (5.4)—see figure 3. We present the result for number of channel uses up to $N = 20$ as in these regime all the relevant qualitative observations can be made, but this is not a fundamental limitation of the method. As a final example, in section 5.5 we will present results for a correlated dephasing noise model (for number of channel uses up to $N = 50$), which shows how effective the tensor network method is in understanding the potential of correlated noise models, for which the fundamental metrological limitations are not yet fully understood.

For all the cases studied, we obtain the optimal adaptive QFI as a function of N for 0, 1 and 2-qubit ancillary systems (this corresponds to $d_{\mathcal{A}} \in \{1, 2, 4\}$). We compare our results with fundamental upper-bounds, which were derived in [16]. The bounds are guaranteed to be saturable when $N \to \infty$ and ancillary system is large enough [13, 16]. However, for some cases, the gap between the upper bound and the

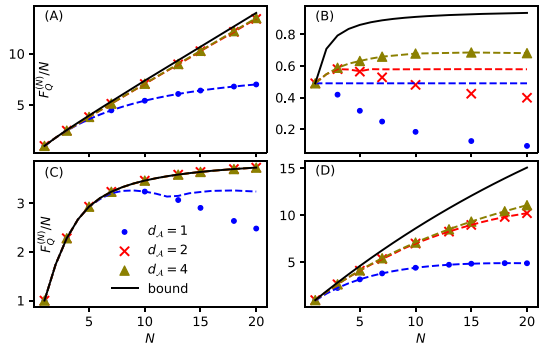


Figure 3. Optimal values of QFI normalised by N for optimal adaptive strategy, given N channel uses, for different dimensions of ancillary system: $d_{\mathcal{A}} = 1$ (blue), $d_{\mathcal{A}} = 2$ (red) and $d_{\mathcal{A}} = 4$ (yellow). Symbols depict values of QFI in a situation where all the N channels are used in a single adaptive protocol, while dashed lines allow for multiple repeated experiments with N channels used in total. Solid black lines correspond to the fundamental upper-bound computed using the methods from [16]. The four plots correspond to different metrological models: (A) perpendicular dephasing, equation (43) ($p = 0.9$); (B) parallel dephasing, equation (44)($p = 0.85$) (C) perpendicular amplitude damping, equation (45) ($p = 0.75$); (D) parallel amplitude damping, equation (48) ($p = 0.9$).

result of ISS optimization is very small. It means, that an almost optimal adaptive estimation protocol can be implemented with a small size ancillary system. To guarantee a high precision of the results obtained, we stop the algorithm only when the relative result increases less than 0.01% over 5 full iterations (optimization over P_1, ..., P_N, L is counted as one iteration). We also made sure, that for different, randomly chosen initial guesses P_1, ..., P_N, L, the final result is the same up to a numerical error.

5.1. Perpendicular dephasing

The noise Kraus operators for this model are given by

$$K_1 = \sqrt{p}\,\mathbb{1},\ K_2 = \sqrt{1-p}\sigma_x, \tag{43}$$

which means that the dephasing acts perpendicularly to the axis of the parameter encoding unitary rotation. This model has long been a paradigmatic example of the potential of application of quantum-error correction inspired protocols to recover the Heisenberg scaling of precision despite presence of noise [11, 12, 41, 64–66].

In case when signal comes before a noise, that is when $K_{\varphi,k} = K_k U_\varphi$ instead of $K_{\varphi,k} = U_\varphi K_k$ (42), the impact of noise can be completely eradicated with only one ancillary qubit [11, 12, 41, 64–66].

It is also known that when signal comes after the noise, which is the case we consider in this paper, the Heisenberg scaling may also be preserved, yet with a reduced coefficient [16]. In [16] the optimal protocol in case of $N = 2$ uses of channel has also been explicitly constructed, which required no ancillary system at all. It was not clear, however, what is the structure of the optimal protocol for larger N and in particular what size of ancillary system is required to reach the optimal performance.

In figure 3(A) we see that the ancillary system is indeed needed in order to preserve the character of the Heisenberg scaling, and that a single qubit ancilla already provides an almost optimal performance—the results fall very close to the fundamental bound, and the numerical improvements thanks to the addition of second ancillary qubit are marginal.

The exemplary almost optimal protocol for $N = 3$ that utilizes only a single ancillary qubit, that we were able to extract from the obtained numerical results, is described in appendix D.1.

5.2. Parallel dephasing

Parallel dephasing is one of the most commonly considered decoherence models in quantum metrology, as it represents the typical situation where the coherences required to sense the parameter of interest are being reduced by decoherence processes. The corresponding noise Kraus operators are

$$K_1 = \sqrt{p}\,\mathbb{1},\ K_2 = \sqrt{1-p}\sigma_z, \tag{44}$$

which means that the dephasing is defined with respect to the same axis as the parameter encoding rotation.

It is well known that this model does not admit asymptotic Heisenberg scaling, and the quantum enhancement amounts to a constant factor improvement, with the asymptotically achievable upper bound on QFI given as $F^{(N)} \leqslant N\frac{(p-1/2)^2}{p(1-p)}$ [6, 8, 9]. It is also well known, that in the parallel-scheme framework the bound can be asymptotically achieved via a Ramsey interferometry scheme and the use of weakly

spin-squeezed states [6, 18, 39] or low bond dimension MPS [39, 67]. These protocols may be practical in many-body systems, as the effecitve entanglement required between the particles is weak, but, nevertheless, large number of elementary probes need to be entangled. As such, it is not obvious if the performance of this optimal parallel many-probe protocol can be effectively simulated via adaptive protocols with small ancilla size in the limit of large N.

Figure 3(B) illustrates that increasing size of ancilla significantly improves the performance of the protocol. The largest size considered, $d_{\mathcal{A}} = 4$, shows significantly better performance than the single qubit ancilla case, but clearly, increasing the dimension of ancilla further would allow to approach the bound even closer—this would require a more dedicated numerical effort though, as three qubit ancilla $d_{\mathcal{A}} = 8$ is on the borderline of numerical complexity that a high performance PC is capable of dealing with. Note that for low dimensional ancilla, at some point QFI starts to drop down, since the decoherence is dominating, and the ancillary systems are not sufficiently large to deal with. In this case it is more advisable to stop the protocol at some $n < N$, and use the remaining resources for a fresh run of the protocol—this strategy is illustrated with dashed lines, and this allows to avoid the drop in per-channel-use performance with increasing N. Interestingly, the effect of QFI per channel descrease is barely noticeable already with a two qubit ancilla.

One might wonder, how relevant is the adaptive aspect of the protocol, and ask how the optimal entanglement based strategy with the same number of available qubits would perform. In appendix D.2 we present numerical results, that confirm that the adaptive strategy with one probe and two ancillary qubits (three qubits total) outperforms the optimal strategy utilizing three entangled qubits. This remains true even if we allow for sequential channel probing (but not full adaptiveness, since intermediate controls are not allowed) in the latter case. This shows that the adaptive strategy found is not just the imitation of an entanglement based strategy. Moreover, it clearly demonstrates that quantum adaptiveness is more than entanglement + sequentiality.

5.3. Perpendicular amplitude damping

Let us now consider a perpedicular damping model, where the corresponding Kraus operators read

$$K_1 = |-\rangle\langle -| + \sqrt{p}|+\rangle\langle +|,\ K_2 = \sqrt{1-p}|-\rangle\langle +|, \tag{45}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ are the eigenvectors of σ_x. This is variation on the standard amplitude damping model, where the damping axis is perpendicular to the phase encoding rotation axis.

Despite perpendicular character of the noise, this model, unlike the perpendicular dephasing one, does not admit asymptotic Heisenberg scaling [16]. The analysis of this model leads to a particularly interesting conclusions, as the adaptive upper-bounds for QFI derived in [16] are saturable for all values of p and N. Previously, this fact was only demonstrated for $N \leqslant 4$. Using the tensor network approach, we show numerically, that the bound is also saturable for N up to 50. Moreover, one qubit ancilla ($d_{\mathcal{A}} = 2$) is enough to saturate the bound, see figure 3(C). Inspired by the structure of the numerical solution, we have found a simple intuitively appealing analytical form of the optimal protocol for all values of p and N.

Let us consider the local parameter estimation around the value $\varphi = 0$ (for other values of φ one needs to adjust the protocol by proper rotation of a probe qubit). Initially, the probe ($\mathcal{H}_1$) and ancilla ($\mathcal{A}_1$) qubits are prepared in a product state $|-\rangle_{\mathcal{H}_1} \otimes |0\rangle_{\mathcal{A}_1}$—notice, that this state is not affected by damping noise. The protocol involves entangling operations V_i applied between the ith and $i+1$th use of the channel Λ_φ followed by the readout of the ancillary qubit. The action of V_i reads:

$$V_i|-\rangle \otimes |0\rangle = |-\rangle \otimes |0\rangle, \tag{46}$$

$$V_i|+\rangle \otimes |0\rangle = t_i|+\rangle \otimes |0\rangle + \sqrt{1-t_i^2}|-\rangle \otimes |1\rangle, \tag{47}$$

where $t_i \in [0,1]$ describes the coupling strength between the probe and the ancilla. For $t_i = 0$ the coupling is the strongest, and probe is measured in $|\pm\rangle$ basis, with readout saved on ancillary qubit; for $t_i = 1$, there is no coupling, the input state remains unaffected. All intermediate values describe a weak measurement, when information accumulated in a probe is partially transformed to ancilla.

After the action of V_i, the ancillary qubit is measured in the basis $\{|0\rangle, |1\rangle\}$. When the result is $|1\rangle$, then the probe no longer carries any information about φ, the whole protocol is restarted, and the number of a step in which it happened is saved (therefore, there is no need to specify how V_i acts on states of the form $|\psi\rangle_{\mathcal{H}} \otimes |1\rangle_{\mathcal{A}}$). When the result is $|0\rangle$, the protocol is uninterrupted. Provided $|0\rangle$ is measured in all N subsequent coherent uses of Λ_φ, then the final probe state is measured in the basis ($|\pm\rangle$). The total classical FI achieved using this scheme depends on coupling parameters t_i. When the values of t_i are chosen optimally the bound derived in [16] is saturated for all N.

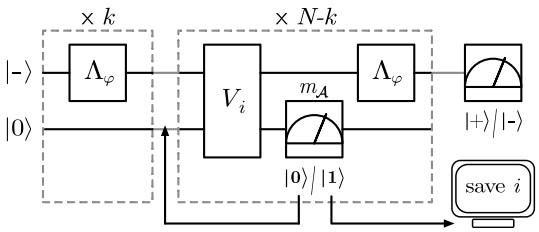


Figure 4. Schematic representation of the optimal metrological protocol for phase estimation in the presence of perpendicular damping noise. Initially, probe and ancilla qubits are prepared in a state $|-\rangle_{\mathcal{H}_1} \otimes |0\rangle_{\mathcal{A}_1}$, and for first k steps, no external control is applied—the probe state is freely evolved through channels Λ_φ. Then, the protocol consists of the interactions V_i between system and ancilla, followed by measurement of ancilla in $|0\rangle/|1\rangle$ basis. If the result of this measurement is $m_{\mathcal{A}} = 0$, the protocol is continued, if $m_{\mathcal{A}} = 1$, the protocol is terminated, and the number of step i in which it happened is saved. After N steps the system is measured in $|+\rangle/|-\rangle$ basis (provided that the protocol was not terminated earlier).

In particular we observe, that when the damping noise is weak enough, then first k optimal coupling parameters are $t_i = 1$—that means, that one should initially let the probe freely evolve through channels Λ_φ. After this stage, interactions V_i effectively keep the probe state in some optimal, fixed point, in which the effect of damping noise is not too large, and at the same time, the measurements performed on ancillary qubit are as informative as possible. See figure 4 for the sketch of the described strategy and appendix D.3 for the derivation and more details.

Apart from one ancillary qubit, the described protocol requires classical memory to store information about the step number in which $|1\rangle$ was measured. The protocol found by our algorithm does not need any extra memory by construction—this can be understood in a way that information instead of being extracted by the measurement is constantly being feed-forward in the structure of the state. However, we decided to present the measurement based protocol thanks to more intuitively appealing nature. Notice, that it is not possible to saturate the bound without ancilla (see figure 3(C), points corresponding to $d_{\mathcal{A}} = 1$).

5.4. Parallel amplitude damping

For completness of discussion of uncorrelated noise models, we consider the standard amplitude damping model representing e.g. spontaneous emission, where decoherence axis is parallel to the phase encoding rotation axis. The Kraus operators for this model are

$$K_1 = |0\rangle\langle 0| + \sqrt{p}|1\rangle\langle 1|, \; K_2 = \sqrt{1-p}|0\rangle\langle 1|. \tag{48}$$

In this model, again, the Heisenberg scaling is asymptotically unattainable [8, 16], as can be also seen in figure 3(D) where the curve representing the bound on $F^{(N)}/N$ (black solid) reveals the asymptotic convergence to a constant for large N.

Similarly to the parallel dephasing model, we see that the performance of the protocol improves significantly with increasing size of the ancillary system, and the 2-qubit ancilla results could be further improved by increasing $d_{\mathcal{A}}$. It can also be seen that while a strategy with 1 ancilla qubit offers a significant advantage over the one without ancilla for all $N > 1$, the difference between strategies with 1 or 2 ancilla qubits are negligible for small N.

We use this example to demonstrate that the tensor network approach provides benefits also in the low-N regime, by allowing for identifying a simpler structure of optimal protocols than the standard full-comb optimization procedures. More concretely, we focus on the basic $N = 2$ uses case, and $p = 0.5$, and show a significantly simpler protocol than the one obtained in [45], where the MOP optimization, combined with standard procedure of decomposing combs into isometries (discussed in section 3.3) was applied.

Because the difference between QFIs achieved with one- and two-qubit ancillae is small for $N = 2$, here we present the strategy with one-qubit ancilla, which is more intuitive and almost optimal (for $N = 2$ the QFI changes from 2.174 to 2.179 when the second ancillary qubit is added). The 2-qubit ancilla strategy achieves the maximal possible QFI and is described in appendix D.4.

The optimal strategy with one-qubit ancilla is schematically presented in figure 5. Its first tooth prepares an entangled state $|\psi_0\rangle$, which is close to the maximally entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The second tooth applies a unitary U on the subspace spanned by $|00\rangle, |11\rangle$ and prepares a different entangled state $|\psi_1\rangle$ for an input state $|01\rangle$ (notice, that input $|10\rangle$ is forbidden due the noise character). See appendix D.4 for more detailed description of this protocol.

In [45] the authors provided a decomposition of the resulting isometries into elementary gates using the package from [61]. The first tooth of their strategy's comb prepares a pure state. To represent the action of

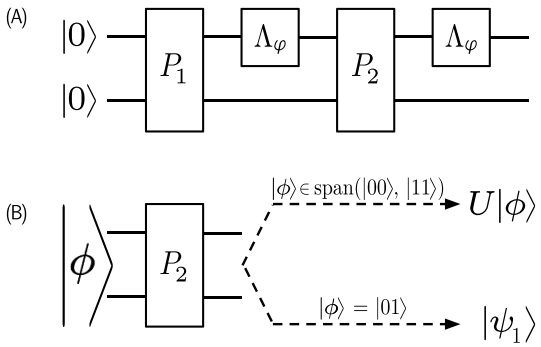


Figure 5. Optimal strategy for parallel damping noise model with two uses of channel Λ_φ, when ancilla is restricted to 1 qubit. (a) Teeth P_1, P_2 of the strategy act before each channel use. First tooth of strategy, P_1, prepares an entangled state. After channel Λ_φ acts for the first time, the 2-qubit state is $|\phi\rangle$. (b) Second tooth, P_2, acts on $|\phi\rangle$ with a unitary U acting only on subspace spanned by $|00\rangle, |11\rangle$ or prepares a new entangled state $|\psi_1\rangle$ if the input state to P_2 was $|01\rangle$.

the second tooth, 33 CNOT gates acting between 5 qubits were required. As demonstrated above, the second tooth of the strategy which we obtained from our tensor network method, restricting ancilla to one qubit, is significantly simpler and provides an intuitve understanding of the action of the protocol. Even strictly optimal strategy involving two ancillary qubits is much simpler than the one presented in [45], as we demonstrate in appendix D.4.

5.5. Correlated parallel dephasing

The examples discussed above covered four uncorrelated noise models. For such models, asymptotically tight bounds can be efficiently derived, and serve as a benchmark of actual protocols [6, 8, 9, 11, 16].

We now move on to show the potential of tensor network methods to deal with correlated noise models. In this case, there are no universal methods to derive fundamental bounds, and hence effective numerical methods to identify optimal metrological protocols are even more desired. We will focus on a simple generalization of parallel dephasing model, that would allow us to study the impact of the strength of temporal (anti-)correlations present in the dephasing process.

Parallel dephasing noise can be interpreted as random rotations of a qubit around the dephasing axis. In the uncorrelated case, these random rotations are assumed to be independent for subsequent interactions of a probe state with the probed channels. In what follows, we will consider a model where these random rotations may be (anti)correlated up to the desired degree.

The physical context for such a model is a situation where we measure the value of a constant magnetic field with a known direction using a spin $1/2$ probe. The probing spin is also affected by a randomly fluctuating field of a different origin, whose direction is parallel to the field we want to estimate. Moreover, if the time scale of the field fluctuations is comparable or slower than the system probing dynamics, then the random rotations representing dephasing in the subsequently probed channels will be correlated. Therefore, $\Lambda_\varphi^{(N)} \neq \Lambda_\varphi^{\otimes N}$, and the noise is properly described by tensor network with nontrivial Hilbert spaces $\mathcal{E}_i$—see (31).

The one-qubit dephasing of strength p, defined in (44), can be alternatively described using a different set of Kraus operators

$$K_1 = \frac{1}{\sqrt{2}} U_{+\epsilon}, \quad K_2 = \frac{1}{\sqrt{2}} U_{-\epsilon}, \tag{49}$$

where $U_{\pm\epsilon} = \mathrm{e}^{\mp \frac{\mathrm{i}}{2}\epsilon\sigma_z}$, $p = \cos^2(\epsilon/2)$. This implies, that the dephasing model can be as well interpreted as resulting from a rotation by angle ϵ around z axis in random direction (50% left, 50% right).

To represent the most basic form of dephasing correlations, we assume that the rotational directions for consecutive dephasing channels are elements of a binary Markov chain given by

$$p_{i|i-1}(+|+) = p_{i|i-1}(-|-) = \frac{1+C}{2}, \tag{50}$$

$$p_{i|i-1}(+|-) = p_{i|i-1}(-|+) = \frac{1-C}{2}, \tag{51}$$

where $p_{i|i-1}(s_i|s_{i-1})$ for $i \in \{2,3,\ldots,N\}$, is the conditional probability of rotational direction s_i in channel i assuming direction s_{i-1} in channel $i-1$, $C \in [-1,1]$ is a correlation parameter: $C = 0$ corresponds to no

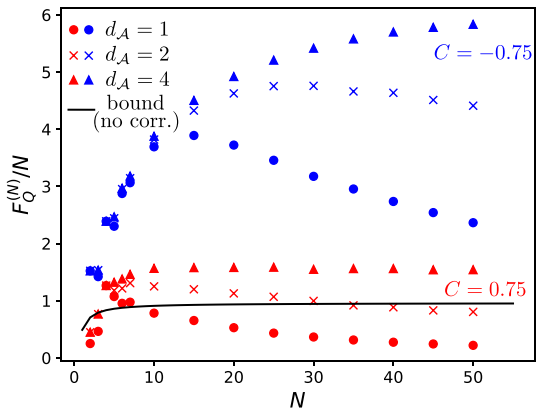


Figure 6. The QFI per channel ($F_Q^{(N)}/N$) as a function of number of channels coherently probed (N) in presence of correlated dephasing noise. The data shown were obtained for dephasing strength $p = 0.85$, for negative ($C = -0.75$, blue) and positive ($C = 0.75$, red) correlations. For comparison, we also show an upper bound for precision for uncorrelated dephasing noise (black line). Interestingly, this bound can be violated with the help of correlations. For small values of N, zero and one-qubit ancilla ($d_\mathcal{A} = 1, 2$) seem to be enough to obtain an optimal precision. However, for larger N, the strategy involving two ancillary qubits ($d_\mathcal{A} = 4$) performs substantially better. Notice, that calculations for $N = 50$ would be impossible without tensor network decomposition technique.

correlations, $C = 1$ means maximal positive correlations (all rotations are in the same direction), and $C = -1$—maximal negative correlations (directions are always different in two neighbor channels). Furthermore, we assume that directions $+$ and $-$ are equally probable in the first channel: $p_1(+) = p_1(-) = 1/2$.

To model this type of correlations using our tensor networks framework, we consider channels Λ_φ acting on a two-qubit space: first qubit ($\mathcal{K}$) is the physical probe, and the second one ($\mathcal{E}$) is a classical bit of memory. When this classical bit is in a state $|\pm\rangle$, then the unitary $U_{\pm\epsilon}$ acts on the probe. After each channel use, the register state is drawn according to the conditional probabilities described in (50) and (51). The Kraus operators describing the action of a channel on a probe and register are

$$K_1 = \sqrt{\frac{1+C}{2}} U_{+\epsilon} \otimes |+\rangle\langle +|, \quad K_2 = \sqrt{\frac{1-C}{2}} U_{+\epsilon} \otimes |-\rangle\langle +|, \tag{52}$$

$$K_3 = \sqrt{\frac{1-C}{2}} U_{-\epsilon} \otimes |+\rangle\langle -|, \quad K_4 = \sqrt{\frac{1+C}{2}} U_{-\epsilon} \otimes |-\rangle\langle -|. \tag{53}$$

The input state of a register of a 1st channel is $\mathbb{1}/2$ to satisfy the condition $p_1(+) = p_1(-1) = 1/2$. Notice, that the action of a single channel on a probe is equivalent to dephasing of strength $p = \cos^2\epsilon$ when there is no information about remaining channels and about classical register state.

We performed numerical calculations for negative ($C = -0.75$) and positive ($C = 0.75$) correlations for dephasing strength $p = 0.85$. The results are shown in figure 6. Note that we also plotted the bound corresponding to the uncorrelated dephasing noise model of the same strength. Interestingly, both negative and positive correlations allow to beat the upper bound calculated for uncorrelated case. However, negative correlations allow for significantly larger values of QFI. In fact, we observed that positive correlations may even decrease the QFI for smaller values of C.

The gain in precision related with negative spatial noise correlations fluctuation is a well known phenomena, and has been discussed in parallel quantum metrological schemes [39, 68–71]. Note, however, that here for the first time, we provided results of optimal performance of adaptive protocols for *time*-correlated dephasing models in the limit of large N and with arbitrarily tunable correlation parameter.

The gain thanks to positive correlations may be less intuitive when trying to base it on spatial-correlations analogy. In this case, one would expect that the more collective character of dephasing (positive correlations) leads to even stronger reduction of the achievable QFI—the correlated dephasing [30, 35] makes the noise more similar to the signal we are sensing. This observation requires a deeper analysis, in order to understand how much of this effect is due to discretization of the phase fluctuations model, and how much is the advantage that appears thanks to the adaptive nature of the protocols considered.

6. Conclusions and outlook

The methods presented here allow for the efficient identification of optimal adaptive protocols in the paradigm of multiple coherent channel uses. These results may be seen as complementary to the research focusing on the derivation of fundamental bounds for the performance of adaptive metrological protocols [9, 11, 12, 16]. In the latter case, the obtained results are guaranteed to be larger or equal to the largest achievable QFI—if numerical optimization is not exact, the final result can be only too large, never too small. On the contrary, in the ISS optimization, the obtained QFI can never be larger than the optimal one, since we always construct an explicit protocol that allows to achieve the QFI returned by an algorithm. Provided the QFI obtained from the two approaches coincide, we are sure that the protocol we have identified is the optimal one. Interestingly, independently of our paper, a recent study appeared where metrological bounds are discussed taking into account limited size of ancillary systems [72]. This approach may be viewed as complementing our results even better, as one may now focus both on protocols and bounds under the same assumptions regarding the limit on the ancillary system size.

The continuous-time fundamental upper-bound for quantum sensing precision [15] can be obtained by taking the limit of infinitesimally short time step in a more general discrete bound [16]. The similar approach can be used to look for optimal continuous-time quantum control—one can search for optimal discrete-time adaptive strategy with decreasing time step. The only problem is that for a fixed total measurement time, more and more steps are required to make a single step short enough. Therefore, it is impossible to efficiently generalize methods relying on optimization over a full comb to continuous-time regime. However, with our tensor-network based optimization one can deal with much larger numbers of channels—therefore, it is possible to divide the whole time evolution into many small pieces, which can be interrogated with control operations. This allows for a good approximation of continuous control assisted estimation.

The methods presented here may also be generalized to fit into the Bayesian framework. This in principle poses no conceptual difficulty within the ISS approach, provided the Bayesian cost is quadratic, as the relevant figure of merit has an analogous structure to the QFI, see (14). The potential practical difficulty may come, however, from a more greedy character of the Bayesian approach when it comes to the dimension of the ancillary systems required—since the classical information is retrieved only in the end of the protocol, and Bayesian approaches typically require larger amount of information to be retrieved in the measurement compared to the QFI based approaches [7]. Nevertheless, it is worth exploring this direction and compare it with other proposals, where methods of identifying optimal adaptive Bayesian protocols in terms of optimization over quantum combs are proposed [56].

The other, and probably more promising application is to use the tensor network approach developed here to find the optimal adaptive protocols for channel discrimination problems [73–76]. In this case, even though the problem is also Bayesian at its core, the amount of information gathered over the run of the protocol is limited by the number of alternatives to be discriminated, and protocols with small size ancillary systems should be efficient. The advantage is the possibility to study the performance of the protocols in the limit of large number of channel uses, a regime out of reach for the all state-of-the art methods [77].

Our approach may also be viewed as complementary to tensor network methods developed for identification of optimal multi-partite probe states for parallel sensing schemes [39]. While these approaches are well-suited to study many particle systems (cold atoms, solid state systems, …) with simple local entanglement structure, the newly developed approach is perfect to study small scale systems that may be interrogated coherently over many rounds of an experiment (atomic clock systems, trapped ions, spins, NV-centers, …). In other words, the present approach addresses the difficulties related with understanding long-time quantum coherence/entanglement potential that may be revealed via specific adaptive protocols, while the former one focuses on spatial aspects of distributed entanglement in multi-partite sensing systems. Interestingly, it is conceivable to combine the two approaches in a unified framework, where both long-range spatial and temporal effects may be satisfactorily analyzed for the sake of identifying optimal metrological protocols. This is an ambitious direction that may be pursued in the future.

Finally, even though we have focused our research on the standard single parameter quantum metrology problem, it should be possible to generalize the methods presented here to multiparameter models as well. To do this one needs to design an effective ISS procedure involving multiparameter figures of merit such as the exact multiparameter cost or the Holevo bound [49, 78] or combine the approach presented with that latest conic-programming based tools for multi-parameter estimation problems proposed in [79].

We are also convinced that our approach may be fruitfully combined with more general quantum process tomography paradigms, where tensor network structures naturally appear [80, 81], as well as help in optimizing quantum control operations in non-Markovian models [82].

Acknowledgments

We thank Francesco Albarelli, Jessica Bavaresco, Marco Tulio Quintino, Wojciech Górecki, David Gross, Kavan Modi, Ingo Roth and Pavel Sekatski for helpful discussions. This work was supported by the National Science Center (Poland) Grant No. 2020/37/B/ST2/02134.

Note added A week after our manuscript was published, an independent preprint with similar results appeared on arXiv [83].

Data availability statement

No new data were created or analysed in this study.

Appendix A. Decomposition of a quantum comb into isometries

Here we summarize the procedure, given in [42], of decomposing a quantum comb P into a concatenation of isometries (see (28) for an illustration) applied to our setup.

Given the CJ operator P of the comb P, we construct the subsequent isometries $V^{(1)},\ldots,V^{(N)}$, where $V^{(k)}:\mathcal{K}_{k-1}\otimes\mathcal{A}_{k-1}\to\mathcal{H}_k\otimes\mathcal{A}_k$. We use the following indices to label the respective basis elements or corresponding Kraus operators:

i - basis of $\mathcal{A}_k$, $i\in[1,\mathrm{rank}(P^{(k)})]$,
j - basis of $\mathcal{A}_{k-1}$, $j\in[1,\mathrm{rank}(P^{(k-1)})]$,
m - basis of $\mathcal{H}_k$,
n - basis of $\mathcal{K}_{k-1}$.

The procedure is as follows.

1. Obtain $\forall_{k\in[1,N-1]}$ the CJ operator $P^{(k)}$ of the comb P up to its kth tooth, using (18).
2. Starting with $k=1$ for each k construct the canonical representation $\{K_i^{(k)}\}_i$ of the channel $\mathrm{P}^{(k)}$.
3. Starting with $k=2$ (since for $k=1$ the isometry is just the input state), for each k construct two different Kraus representations of the same channel (A11) (see also (A13) and (A14)):

$$\left\{K'^{(k)}_{i,m}\right\}_{i,m}=\left\{\left(\mathbb{1}_{\mathcal{H}_{1,\ldots,k-1}}\otimes\langle m|\right)K_i{}^{(k)}\right\}_{i,m},\tag{A1}$$

$$\left\{K''^{(k)}_{j,n}\right\}_{j,n}=\left\{K_j^{(k-1)}\otimes\langle n|\right\}_{j,n}.\tag{A2}$$

4. Using the system of equations

$$K'^{(k)}_{i,m}=\sum_{j,n}V^{(k)}_{im,jn}K''^{(k)}_{j,n}\tag{A3}$$

obtain the matrix elements of the isometry $V^{(k)}$ which connects the two representations of the channel. Finally from:

$$V^{(k)}=\sum_{i,j,m,n}V^{(k)}_{im,jn}|m\rangle\langle n|\otimes|i\rangle\langle j|\tag{A4}$$

obtain the isometry. From this construction we have $\dim(\mathcal{A}_k)=\mathrm{rank}(P^{(k)})$.

To see why such $V^{(k)}$ are indeed isometries whose concatenation yields P, let us first recall from (18) the condition for P to be the CJ operator of a quantum comb P:

$$\begin{aligned}&P\geqslant 0,\ \mathrm{Tr}_{\mathcal{A}_N\otimes\mathcal{H}_N}P=P^{(N-1)}\otimes\mathbb{1}_{\mathcal{K}_{N-1}},\\&\underset{1<k<N}{\forall}\mathrm{Tr}_{\mathcal{H}_k}P^{(k)}=P^{(k-1)}\otimes\mathbb{1}_{\mathcal{K}_{k-1}},\ \mathrm{Tr}_{\mathcal{H}_1}P^{(1)}=1.\end{aligned}\tag{A5}$$

Note that those conditions imply that all $P^{(k)}\in\mathcal{L}(\bigotimes_{i=1}^{k-1}\mathcal{K}_i\otimes\bigotimes_{i=1}^{k}\mathcal{H}_i)$ are positive semidefinite and if we trace out all output spaces they act on $(\mathcal{H}_i:i=1,\ldots,k)$ we get an identity operator on all input spaces $(\mathcal{K}_i:i=1,\ldots,k-1)$. Hence each $P^{(k)}$ is a CJ operator of a CPTP map [42]

$$\mathrm{P}^{(k)}:\mathcal{L}\left(\bigotimes_{i=1}^{k-1}\mathcal{K}_i\right)\to\mathcal{L}\left(\bigotimes_{i=1}^{k}\mathcal{H}_i\right).\tag{A6}$$

Through the analysis of the input and output spaces of this map it is easy to see that it is just a quantum channel obtained from comb P by ignoring (tracing out) all spaces related to teeth $k+1,\ldots,N$ that is $\mathcal{K}_k$, $\mathcal{H}_{k+1}$, ..., $\mathcal{K}_{N-1}$, $\mathcal{H}_N$, $\mathcal{A}_N$. Then the conditions (A5) (apart form positivity) can be reformulated as

$$\mathrm{Tr}_{\mathcal{H}_{k+1}}\mathrm{P}^{(k+1)}(\rho)=\mathrm{P}^{(k)}(\mathrm{Tr}_{\mathcal{K}_k}\rho), \tag{A7}$$

for any density matrix ρ. Naturally, each channel $\mathrm{P}^{(k)}$ can be represented by a set of Kraus operators $\{K_i^{(k)}\}$ and purified to isometry $W^{(k)}$ such that

$$K_i^{(k)}=\langle i|W^{(k)}, \tag{A8}$$

where $|i\rangle$ is an o.-n. basis of auxiliary space $\mathcal{A}_k$ used for purification. In the above formula partial inner product was used, that is on the RHS $\langle i|$ formally denotes $\mathbb{1}_{\mathcal{H}_{1,\ldots,k}}\otimes\langle i|$. This notation will also be used throughout this section, when applicable.

Let us for a moment assume that P can be decomposed into isometries $V^{(1)},\ldots,V^{(N)}$ and investigate what conditions they should satisfy. First, since $V^{(k)}$ performs the action of the kth tooth, it needs to have $\mathcal{K}_{k-1}$ as one of its input spaces and $\mathcal{H}_k$ as one of its output spaces. Furthermore, among its input spaces we need a space that would connect it with all previous $k-1$ teeth. The natural choice for this space is $\mathcal{A}_{k-1}$, since it can be interpreted as a comb memory after $k-1$ teeth of a comb. Analogously, we need a space that would connect $V^{(k)}$ to the next tooth and the natural choice here is $\mathcal{A}_k$. Thus we are looking for an isometry

$$V^{(k)}:\mathcal{K}_{k-1}\otimes\mathcal{A}_{k-1}\to\mathcal{H}_k\otimes\mathcal{A}_k. \tag{A9}$$

For such a choice we can multiply the purification of $k-1$ teeth ($W^{(k-1)}$) by $V^{(k)}$ and for $V^{(k)}$ to be the kth tooth we want the result of this multiplication to be the purification of k teeth—$W^{(k)}$. In summary, we need $V^{(k)}$ to satisfy:

$$W^{(k)}=V^{(k)}W^{(k-1)}. \tag{A10}$$

To see that such a matrix indeed exists let us consider two transformations:

$$\mathcal{L}\left(\bigotimes_{i=1}^{k-1}\mathcal{K}_i\right)\ni\rho\mapsto\mathrm{Tr}_{\mathcal{H}_k}\mathrm{P}^{(k)}(\rho)\in\mathcal{L}\left(\bigotimes_{i=1}^{k-1}\mathcal{H}_i\right), \tag{A11}$$

$$\mathcal{L}\left(\bigotimes_{i=1}^{k-1}\mathcal{K}_i\right)\ni\rho\mapsto\mathrm{P}^{(k-1)}\left(\mathrm{Tr}_{\mathcal{K}_{k-1}}\rho\right)\in\mathcal{L}\left(\bigotimes_{i=1}^{k-1}\mathcal{H}_i\right). \tag{A12}$$

Clearly both of them are CPTP maps thus they have Kraus representations—here $K'(k)$ and $K''(k)$ respectively. Those representations can be constructed as follows:

$$\begin{aligned}\mathrm{Tr}_{\mathcal{H}_k}\mathrm{P}^{(k)}(\rho)&=\sum_m\langle m|\mathrm{P}^{(k)}(\rho)|m\rangle\\&=\sum_{i,m}\langle m|K_i^{(k)}\rho K_i^{(k)\dagger}|m\rangle,\end{aligned}$$

so

$$K'^{(k)}_{i,m}=\langle m|K_i^{(k)}. \tag{A13}$$

Analogous reasoning leads to

$$K''^{(k)}_{j,n}=K_j^{(k-1)}\langle n|. \tag{A14}$$

By relation (A7) channels (A11) and (A12) are equal thus $K'(k)$ and $K''(k)$ are two Kraus representations of the same channel. It follows that there exists an isometry $V_{im,jn}$ connecting those two representations [58]. They are related by (A3). Finally:

$$\begin{aligned}W^{(k)}&\overset{(A8)}{=}\sum_i|i\rangle K_i^{(k)}\overset{(A13)}{=}\sum_{i,m}|m\rangle|i\rangle K'^{(k)}_{i,m}\\&\overset{(A3)}{=}\sum_{i,j,m,n}V_{im,jn}|m\rangle|i\rangle K''^{(k)}_{j,n}\end{aligned}$$

$$\overset{(A14)}{=} \sum_{i,j,m,n} V_{im,jn}|m\rangle\langle n| \otimes |i\rangle K_j^{(k-1)}$$
$$\overset{(A8)}{=} \sum_{i,j,m,n} V_{im,jn}|m\rangle\langle n| \otimes |i\rangle\langle j| W^{(k-1)}$$
$$= \left(\sum_{i,j,m,n} V_{im,jn}|m\rangle\langle n| \otimes |i\rangle\langle j| \right) W^{(k-1)},$$

thus the matrix in brackets is an isometry that satisfies (A10) and this is indeed the matrix (A4) which we give in the last step of our procedure.

Appendix B. Tensor network representation of quantum combs

Let us introduce the basic concepts of tensor networks formalism. Any n-index tensor $T_{i_1 i_2 \ldots i_n}$ can be represented as a rectangle with n legs:

$$T_{i_1 i_2 \ldots i_n} = \quad (B1)$$

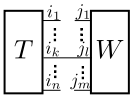
All the links in the depicted example come from the right side of a rectangle, generally the side from which a link comes out does not have any mathematical meaning.

Let $T_{i_1 i_2 \ldots i_n}$ and $W_{j_1 j_2 \ldots j_m}$ be two tensors, and let us assume that the range of indices i_k and j_l is the same: $i_k, j_l \in \{0, 1, \ldots, d-1\}$. Then, a tensor S can be constructed by contraction of indices i_k and j_l:

$$S = T_{i_1 \ldots i_{k-1} i i_{k+1} \ldots i_n} W_{j_1 \ldots j_{l-1} i j_{l+1} \ldots j_n}, \quad (B2)$$

the summation convention was used above—the sum over i ranging from 0 to $d-1$ was performed. Notice, that S has $n+m-2$ indices: $i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n, j_1, \ldots, j_{l-1}, j_{l+1}, \ldots, j_m$. Tensor S can be graphically represented as

$$S = \quad (B3)$$



Free legs correspond to $n+m-2$ indices of S, connected legs correspond to contracted indices. We can combine more tensors into a network, in which each link denotes a contraction of one indices pair. For a more detailed introduction to a tensor networks formalism see [84].

Let us now demonstrate the construction of multi-index tensors representing quantum channels, which allow to express a link product using tensor networks formalism. Let $C \in \mathrm{Lin}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_N)$ be a CJ matrix of a channel whose input and output subspaces are $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_N$ (for our further considerations it does not matter which spaces are outputs and which are inputs). Operator C can be written as

$$C = C^{i_1 i_2 \ldots i_N}_{i_1' i_2' \ldots i_N'} |i_1 i_2 \ldots i_N\rangle\langle i_1' i_2' \ldots i_N'|, \quad (B4)$$

where indices i_j, i_j' run from 0 to $d_j - 1$, $d_j = \dim(\mathcal{H}_j)$; vectors $|0\rangle, |1\rangle, \ldots, |d_j - 1\rangle$ form o.-n. basis of $\mathcal{H}_j$. We can construct an equivalent representation of C by concatenating indices i_j and i_j' into one index k_j, ranging from 0 to $d_j^2 - 1$. Then, we obtain an N-index tensor $\tilde{C}$, whose elements are

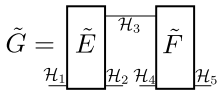
$$\tilde{C}^{k_1 k_2 \ldots k_N} = C^{i_1 i_2 \ldots i_N}_{i_1' i_2' \ldots i_N'} \quad \text{for} \quad k_j = d_j i_j + i_j'. \quad (B5)$$

Notice, that for $N = 1$, the described procedure corresponds to a matrix vectorization.

Let $E : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3)$, $F : \mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_4) \to \mathcal{L}(\mathcal{H}_5)$ be quantum channels, and $E \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3)$, $F \in \mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_4 \otimes \mathcal{H}_5)$ the corresponding CJ operators. We can construct another channel, $G : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_4) \to \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_5)$ by using part of E ($\mathcal{H}_3$) a part of input of F. Then, by construction, $G \in \mathrm{Comb}[(\mathcal{H}_1, \mathcal{H}_2), (\mathcal{H}_4, \mathcal{H}_5)]$, and the corresponding CJ matrix G can be written using link product:

$$G = E \star F = \mathrm{Tr}_{\mathcal{H}_3}\left[(E \otimes \mathbb{1}_{\mathcal{H}_4 \otimes \mathcal{H}_5}) \left(\mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{H}_2} \otimes F^{T_{\mathcal{H}_3}} \right) \right]. \quad (B6)$$

Equivalently, we can write down the relation between $\tilde{G}, \tilde{E}, \tilde{F}$ using tensor network formalism:

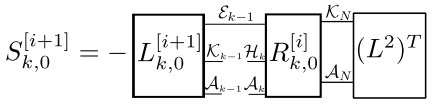


(B7)

This procedure can be directly generalized to channels with more input and output subspaces. We can also concatenate more channels, to represent quantum combs consisting of many teeth. In the main text, to simplify the notation, we remove $\sim$ symbol, and write the symbols of CJ matrices (e.g. G, E, F) in the tensor network notation. Formally, one should understand this notation as indices contraction in the corresponding tensors $\tilde{G}, \tilde{E}, \tilde{F}$.

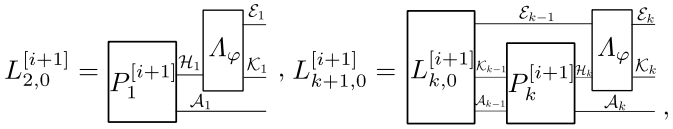
Appendix C. Tensor network based optimization: technical details

Let us present an efficient way to compute S_k matrices which are used to optimize over strategy teeth P_k, see (39). Firstly, to be more precise, let us define $P_k^{[i]}$ as a value of kth tooth of an estimation strategy obtained after i full iterations of an optimization algorithm. To update this value to $P_k^{[i+1]}$, we need to construct a matrix $S_k^{[i+1]}$—this is a notation for a matrix S_k with updated values of P_l for $l<k$ ($P_{l<k} = P_l^{[i+1]}$) and the values of $P_{l>k}$ from the previous iteration ($P_{l>k} = P_l^{[i]}$). This is related with the order of optimization—we optimize from left to right. Let us also define $S_{k,+} = \sum_{i=1}^{N} S_{k,i}$, this means that $S_k = S_{k,0} + S_{k,+}$, see (37). The matrix $S_{k,0}^{[i+1]}$ can be represented as

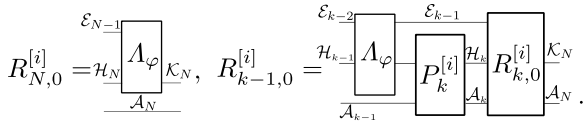


(C1)

where $L_{k,0}^{[i+1]}$, $R_{k,0}^{[i]}$ correspond to contraction of elements from left and right of P_k respectively, and are defined by the following iterative relations:

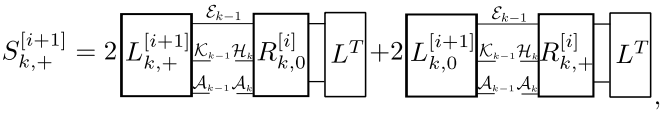


(C2)



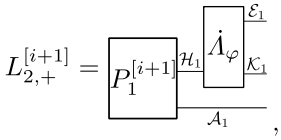
(C3)

Notice, that we do not define $L_{1,0}$ since there are not any object at the left side of P_1. Instead of calculating matrices $S_{k,i}$ for $i>0$ independently, we can calculate their sum $S_{k,+}$ directly, using the following relation

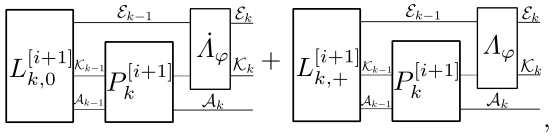


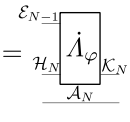
(C4)

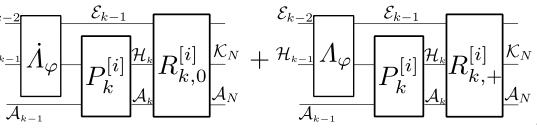
where $L_{k,+}^{[i+1]}$, $R_{k,+}^{[i]}$ are defined by the following iterations:



(C5)

$$L_{k+1,+}^{[i+1]} = $$  (C6)

$$R_{N,+}^{[i]} = $$  (C7)

$$R_{k-1,+}^{[i]} = $$  (C8)

It is straightforward to see that after applying these iterative relations, the figure of merit defined in (33) can be indeed written down as (36). Moreover, all tensor networks contractions required to find $P_1^{[i+1]}, P_2^{[i+1]}, \ldots, P_N^{[i+1]}, L^{[i+1]}$ given $P_1^{[i]}, P_2^{[i]}, \ldots, P_N^{[i]}, L^{[i]}$ can be done in time $O(N)$. To achieve this, we follow the following algorithm.

1. Tensors $R_{k,0}^{[i]}, R_{k,+}^{[i]}$ are computed using (C3), (C7), (C8) for $k \in \{1,2,\ldots,N\}$. When the iterative relations are applied directly, then the time of computation of all of them is $O(N)$.
2. The optimization over P_1 is performed, and $P_1^{[i]}$ is replaced with $P_1^{[i+1]}$. Then, $L_{2,0}^{[i+1]}$ and $L_{2,+}^{[i+1]}$ are computed using (C2) and (C5). This allows us to compute $S_{2,0}^{[i+1]}$ and $S_{2,+}^{[i+1]}$ using (C1) and (C4). Then, we perform optimization over P_2 and find $P_2^{[i+1]}$, which is used to calculate $L_{3,0}^{[i+1]}$, $L_{3,+}^{[i+1]}$ with the help of (C2) and (C6). This procedure is continued until we find $P_N^{[i+1]}$. Notice, that the time of computation of $L_{k+1,0}^{[i+1]}$, $L_{k+1,+}^{[i+1]}$ using $L_{k,0}^{[i+1]}$, $L_{k,+}^{[i+1]}$ does not depend on N. Therefore, the time of all contractions and optimizations described in this step is $O(N)$.
3. Finally, we find $L^{[i+1]}$ using a standard procedure, described, for example, in [32, 39].

Appendix D. Optimal metrological protocols for select examples

D.1. Perpendicular dephasing

Here we describe the almost optimal protocol for perpendicular dephasing, $N = 3$ and $d_{\mathcal{A}} = 2$. Let us start with the optimal protocol for $N = 2$. We take as an input state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{\mathcal{HA}} + |11\rangle_{\mathcal{HA}})$. After the action of the channel this becomes a mixture of

$$\frac{1}{\sqrt{2}}\left(|00\rangle + e^{i\varphi}|11\rangle\right), \quad \frac{1}{\sqrt{2}}\left(|10\rangle + e^{i\varphi}|01\rangle\right), \tag{D1}$$

with probabilities p and $1-p$ respectively. Then the error can be detected by projecting the state on one of the two subspaces $\mathcal{C} = \text{span}\{|00\rangle, |11\rangle\}$ and $\mathcal{E} = \text{span}\{|01\rangle, |10\rangle\}$ and then corrected. The action of the second channel leads to the mixture of

$$\frac{1}{\sqrt{2}}\left(|00\rangle + e^{i2\varphi}|11\rangle\right), \quad \frac{1}{\sqrt{2}}\left(|10\rangle + |01\rangle\right), \tag{D2}$$

with probabilities p and $1-p$ respectively. This state gives QFI at level $4p$ which is an optimal value for $p \geqslant 0.5$. For $p \leqslant 0.5$ the occurrence of σ_x error is more probable than not thus it is more beneficial to treat nonoccurence of σ_x as an error and correct the state when it is in $\text{span}\{|00\rangle, |11\rangle\}$. Therefore finally we get $F_Q^{(2)} = 2(1 + |1-2p|)$.

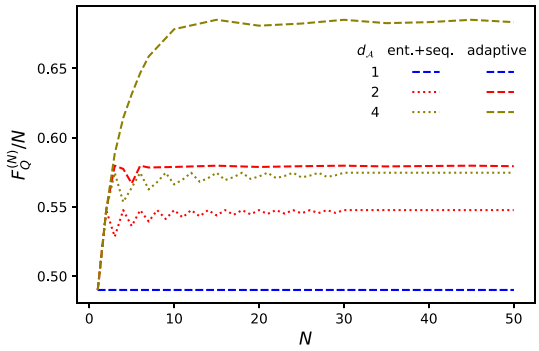


Figure 7. Optimal values of QFI per channel ($F_{\mathrm{Q}}^{(N)}/N$) as a function of the number of channel uses (N). General adaptive strategies (dashed line) and strategies involving entanglement and simple sequentiality (dotted line) are compared for different total number of entangled qubits: 1 (blue, $d_{\mathcal{A}}=1$, for this case dashed and dotted line overlap), 2 (red, $d_{\mathcal{A}}=2$) and 3 (yellow, $d_{\mathcal{A}}=4$). For both cases we allow for splitting N channel uses into k experiments with n_i channel uses ($\sum_{i=1}^{k} n_i = N$).

Our protocol for $N=3$ does everything as in the case of $N=2$ and then after the action of the second channel it transforms all states in $\mathcal{E}$ (or $\mathcal{C}$ for $p<0.5$) into:

$$\frac{1}{2}\left(|0\rangle+|1\rangle\right)\otimes\left(|0\rangle+\mathrm{e}^{-\mathrm{i}\pi/4}|1\rangle\right). \tag{D3}$$

Numerical computations showed that this results in a QFI that is at worst 8% percent smaller than the optimal protocol for $d_{\mathcal{A}}=2$.

Note, that in case when signal comes before a noise the matrix U_φ acts on a state immediately after it was corrected. Thus it can be ensured that this state will always be pure which allows for a complete eradication of a noise and $F_{\mathrm{Q}}^{(N)}=N^2$.

D.2. Parallel dephasing—adaptive vs entangled based strategy analysis

In this appendix, we present the data showing the comparison between entanglement based and adaptive strategies utilizing the same total number of entangled qubits, see figure 7. For general adaptive strategies (dashed lines), we probe the channels sequentially, and allow for arbitrary control operations in between. For entanglement based strategies, we do not allow for any intermediate control, but we can optimally choose the input entangled state. Moreover, in the latter case, channels can be probed sequentially, which means that each qubit of an input can be evolved through a number of probed channels. In both cases, we allow for the same total number of channel uses and for the same number of entangled qubits. Notice that $d_{\mathcal{A}}=2^m$ (m ancillary qubits) in adaptive strategy correspond to $m+1$ entangled qubits in total. In some cases, it is advantageous to split N probed channels into subsets of $n_1, n_2, \ldots, n_k$ channels, and repeat the protocol k times, in ith use, n_i channels are utilized. The optimal division of channels into different coherent runs of the protocol is allowed in both considered cases.

D.3. Perpendicular amplitude damping

Let us describe in more details the optimal estimation protocol introduced in section 5.3, and calculate the associated FI . For a moment, let us assume that after ith action of a channel Λ_φ the probe qubit is in a state

$$\rho_\varphi^{(i)} = |\psi\rangle_{c_i\varphi}\langle\psi| + \mathcal{O}\left(\varphi^3\right), \tag{D4}$$

where

$$|\psi\rangle_{c_i\varphi} = \mathrm{e}^{-\frac{\mathrm{i}}{2}c_i\varphi\sigma_z}|-\rangle = \cos\left(\frac{c_i\varphi}{2}\right)|-\rangle - \mathrm{i}\sin\left(\frac{c_i\varphi}{2}\right)|+\rangle, \tag{D5}$$

the ancillary qubit is in a state $|0\rangle$. In the calculation, we neglect higher order in φ, since the estimation around $\varphi=0$ is considered. Using identity

$$V_i|\psi\rangle_{c_i\varphi}\otimes|0\rangle = \cos\left(\frac{c_i\varphi}{2}\right)|-\rangle\otimes|0\rangle + \tag{D6}$$

$$-\mathrm{i}t_i\sin\left(\frac{c_i\varphi}{2}\right)|+\rangle\otimes|0\rangle - \mathrm{i}\sqrt{1-t_i^2}\sin\left(\frac{c_i\varphi}{2}\right)|-\rangle\otimes|1\rangle \tag{D7}$$

and (D4), we can write the output of the control operation as

$$V_i\left(\rho_\varphi^{(i)}\otimes|0\rangle\langle 0|\right)V_i^\dagger=|\chi_i\rangle\langle\chi_i|+\mathcal{O}\left(\varphi^3\right), \tag{D8}$$

where

$$|\chi_i\rangle=|\psi\rangle_{c_i t_i\varphi}\otimes|0\rangle-\mathrm{i}\sqrt{1-t_i^2}\frac{c_i\varphi}{2}|-\rangle\otimes|1\rangle. \tag{D9}$$

The ancillary qubit of this output is measured in a computational basis

- With probability $p_{1,i}=\frac{c_i^2\varphi^2}{4}(1-t_i^2)+\mathcal{O}(\varphi^3)$ we measure ancillary qubit in state $|1\rangle$. Then the 1st qubit is in a state $|-\rangle$ and carries no information about φ. We can start the whole protocol again.
- With probability $p_0=1-\mathcal{O}(\varphi^2)$ we measure ancillary qubit in $|0\rangle$, and the 1st qubit is then in a state $|\psi\rangle_{c_i t_i\varphi}\langle\psi|+\mathcal{O}(\varphi^3)$.

In the 2nd case, we use the output state as input to next channel Λ_φ—the output of Λ_φ can be then calculated using (45), and after expanding around $\varphi=0$, we obtain

$$\Lambda_\varphi\left(|\psi\rangle_{c_i t_i\varphi}\langle\psi|\right)=|\psi\rangle_{c_{i+1}\varphi}\langle\psi|+\mathcal{O}\left(\varphi^3\right), \tag{D10}$$

where

$$c_{i+1}=c_i t_i\sqrt{p}+1. \tag{D11}$$

This justifies our initial assumption about the form of the input state of V_i—we start with a state $|-\rangle=|\psi_0\rangle$, and during the whole protocol the state is in the form $|\psi\rangle_{c_i\varphi}\langle\psi|+\mathcal{O}(\varphi^3)$. The coefficients c_i are given by the recursive relation (D11) with an initial condition $c_0=0$.

The total probability of a protocol termination in step i is

$$p_i=(1-p_{1,1})\ldots(1-p_{1,i-1})p_{1,i}=\frac{c_i^2\varphi^2}{4}\left(1-t_i^2\right)+\mathcal{O}\left(\varphi^4\right), \tag{D12}$$

note that it is equal to $p_{1,i}$ when terms of order φ^4 and higher are neglected. With probability $1-\mathcal{O}(\varphi^2)$ the protocol will not terminate after N channel uses, and at the output we will obtain a probe state $|\psi\rangle_{c_N\varphi}\langle\psi|+\mathcal{O}(\varphi^3)$ and ancillary qubit in state $|0\rangle$. The QFI associated with this output state is

$$F_{\text{out}}=c_N^2, \tag{D13}$$

the measurement which allows to achieve the classical FI equal to QFI is, for example, the measurement in $|\pm\rangle$ basis.

The total FI achieved in the described protocol is

$$F^{(N)}=\sum_{i=1}^{n-1}\frac{\dot{p}_i^2}{p_i}+F_{\text{out}}=\sum_{i=1}^{n-1}c_i^2\left(1-t_i^2\right)+c_N^2. \tag{D14}$$

To get an optimal performance, we need to optimize the result over $\{t_i\}$. This can be done step by step—for a given c_i we pick the value t_i that maximizes the sum of the FI associated with the measurement of ancillary system after V_i and the QFI of the probe output state. Therefore, we find t_i maximizing the function

$$\frac{\dot{p}_i^2}{p_i}+c_{i+1}^2=c_i^2\left(1-t_i^2\right)+\left(c_i t_i\sqrt{p}+1\right)^2. \tag{D15}$$

This is a quadratic function, and it can be easily shown that the optimal choice of t_i is

$$t_i=\begin{cases}\frac{\sqrt{p}}{c_i(1-p)}, & \text{when} \quad \frac{\sqrt{p}}{(1-p)}\leqslant c_i\\ 1 & \text{othwerwise}\end{cases}. \tag{D16}$$

After inserting this to (D11), we obtain

$$c_{i+1}=\begin{cases}\frac{1}{1-p}, & \text{when} \quad \frac{\sqrt{p}}{(1-p)}\leqslant c_i\\ c_i\sqrt{p}+1 & \text{othwerwise}\end{cases}. \tag{D17}$$

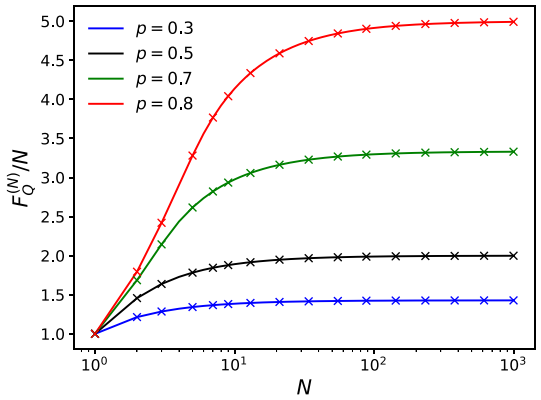


Figure 8. The QFI per channel ($F_Q^{(N)}/N$) as a function of number of channels N for perpendicular amplitude damping noise for different noise parameters p. The fundamental bound (solid line) is compared with the FI achieved by the described protocol (crosses). The bound is saturable using this protocol, so crosses coincide with solid lines.

Notice, we always have $c_i \leqslant c_i\sqrt{p}+1 \leqslant \frac{1}{1-p}$ for $c_i \leqslant \frac{\sqrt{p}}{(1-p)}$, which means that $c_i \leqslant c_{i+1}$ and $c_i \leqslant \frac{1}{1-p}$ for any i. In the first stage of the protocol (first k channel uses, see figure 4) the optimal value of t_i is 1, which means that no error correction is performed. In that stage, c_i keeps growing. Eventually, for $i=k$, c_k becomes larger than $\frac{\sqrt{p}}{(1-p)}$, and then $t_k < 1$ must be picked, according to (D16). Moreover, according to (D17), it means that $c_{k+1} = \frac{1}{1-p}$. Since now, protocol is in its stable phase—it can be easily seen from (D17) and (D16), that $c_i = \frac{1}{1-p}$ and $t_i = \sqrt{p}$ for $i \geqslant k+1$. The error correction with this value of t_i keeps the probe state in a state $|\psi\rangle_{\varphi/(1-p)}$ as long as ancillary qubit is measured in $|0\rangle$. Then, the FI increase associated with each new channel is $\frac{1}{1-p}$, which is an optimal asymptotic value of QFI per channels for this noise, as a result of bounds derived in [16].

To show that this protocol is indeed optimal, we calculated the FI analytically using (D17), (D14), and compared the result with the fundamental bound. The bound turned out to be saturated in all the cases, see figure 8.

D.4. Parallel amplitude damping

Here we provide details of the optimal strategies for parallel damping, with restriction to 1 or 2 qubits of ancilla, as introduced in section 5.4. We provide explicit numerical values for strategies for $p=0.5$, as a comparison with the strategy presented in [45].

D.4.1. 1-qubit ancilla

The Kraus operators $\{L_{i,j}\}$ of the ith tooth of the strategy are as follows. The first tooth prepares an entangled state:

$$L_{0,0} = |\psi_0\rangle, \tag{D18}$$

where $|\psi_0\rangle = a_1|00\rangle + a_2|11\rangle$ and for $p=0.5$ $a_1 = 0.678, a_2 = 0.735$. The second tooth's Kraus operators are as follows:

$$\begin{aligned} L_{1,0} &= U(|00\rangle + |11\rangle)(\langle 00| + \langle 11|), \\ L_{1,1} &= |\psi_1\rangle\langle 01|, \\ L_{1,2} &= |10\rangle\langle 10|, \end{aligned} \tag{D19}$$

where $|\psi_1\rangle = b_1|00\rangle + b_2|11\rangle$ and for $p=0.5 : b_1 = 0.590, b_2 = 0.800 + 0.105i$, $U = 0.959|00\rangle\langle 00| + (-0.282 - 0.002i)|00\rangle\langle 11| + (0.278 + 0.047i)|11\rangle\langle 00| + (0.945 + 0.166i)|11\rangle\langle 11|$.

D.4.2. 2-qubit ancilla

For this ancilla size the adaptive bound from [16] is saturated. The Kraus operators $\{L_{i,j}\}$ of the ith tooth of the optimal strategy are as follows. The first tooth prepares an entangled state:

$$L_{0,0} = |\psi_0\rangle, \tag{D20}$$

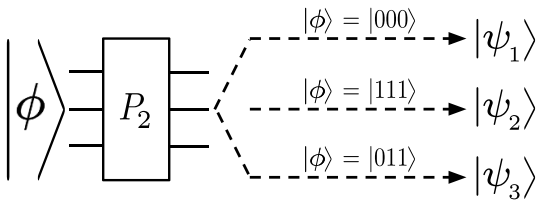


Figure 9. Second tooth, P_2, of the optimal strategy for parallel damping noise model with two uses of channel Λ_φ, when ancilla is restricted to 2 qubits. After a state $|\psi_0\rangle$ is prepared by the first tooth and after the action of signal and noise, the 3-qubit input state to P_2 is $|\phi\rangle$. P_2, acts as follows: if $|\phi\rangle$ is $|000\rangle$, $|\psi_1\rangle$ is prepared; if $|\phi\rangle$ is $|111\rangle$, $|\psi_2\rangle$ is prepared; if $|\phi\rangle$ is $|011\rangle$, $|\psi_3\rangle$ is prepared.

where $|\psi_0\rangle = a_1|000\rangle + a_2|111\rangle$ and for $p = 0.5:\ \ a_1 = 0.675, a_2 = 0.738$. The second tooth's action is represented schematically in figure 9. The second tooth's Kraus operators are as follows:

$$\begin{aligned} L_{1,0} &= |\psi_1\rangle\langle 000| + |\psi_2\rangle\langle 111|, \\ L_{1,1} &= |\psi_3\rangle\langle 011|, \\ L_{1,k} &= |\psi_k\rangle\langle\psi_k| \text{ for } \psi_k \in \{|001\rangle, |010\rangle, |100\rangle, |101\rangle, |110\rangle\}, \end{aligned} \tag{D21}$$

where:

$$\begin{aligned} |\psi_1\rangle &= b_1|000\rangle + b_2|101\rangle, \\ |\psi_2\rangle &= c_1|000\rangle + c_2|101\rangle, \\ |\psi_3\rangle &= e_1|000\rangle + e_2|010\rangle + e_3|011\rangle + e_4|101\rangle \\ &\quad + e_5|110\rangle + e_6|111\rangle. \end{aligned} \tag{D22}$$

For $p = 0.5 : b_1 = -0.956, b_2 = -0.160 - 0.248i, c_1 = 0.295, c_2 = -0.519 - 0.803i, e_1 = 0.062, e_2 = -0.041 + 0.519i, e_3 = 0.289 + 0.237i, e_4 = 0.042 + 0.065i, e_5 = 0.370 + 0.247i, e_6 = -0.580 + 0.225i$.

ORCID iDs

Stanisław Kurdziałek https://orcid.org/0000-0001-8877-1421
Piotr Dulian https://orcid.org/0009-0006-7312-2844
Joanna Majsak https://orcid.org/0000-0001-6612-328X
Sagnik Chakraborty https://orcid.org/0000-0001-6345-9757
Rafał Demkowicz-Dobrzański https://orcid.org/0000-0001-5550-4431

References

[1] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222
[2] Degen C L, Reinhard F and Cappellaro P 2017 *Rev. Mod. Phys.* **89** 035002
[3] Pezzè L, Smerzi A, Oberthaler M K, Schmied R and Treutlein P 2018 *Rev. Mod. Phys.* **90** 035005
[4] Pirandola S, Bardhan B R, Gehring T, Weedbrook C and Lloyd S 2018 *Nat. Photon.* **12** 724
[5] Fujiwara A and Imai H 2008 *J. Phys. A: Math. Theor.* **41** 255304
[6] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nat. Phys.* **7** 406
[7] Hall M J W and Wiseman H M 2012 *Phys. Rev.* X **2** 041006
[8] Demkowicz-Dobrzański R, Kołodyński J and Guţă M 2012 *Nat. Commun.* **3** 1063
[9] Demkowicz-Dobrzański R and Maccone L 2014 *Phys. Rev. Lett.* **113** 250801
[10] Jarzyna M and Demkowicz-Dobrzański R 2015 *New J. Phys.* **17** 013010
[11] Demkowicz-Dobrzański R, Czajkowski J and Sekatski P 2017 *Phys. Rev.* X **7** 041009
[12] Zhou S, Zhang M, Preskill J and Jiang L 2018 *Nat. Commun.* **9** 78
[13] Zhou S and Jiang L 2021 *PRX Quantum* **2** 010343
[14] Górecki W, Demkowicz-Dobrzański R, Wiseman H M and Berry D W 2020 *Phys. Rev. Lett.* **124** 030501
[15] Wan K and Lasenby R 2022 *Phys. Rev. Res.* **4** 033092
[16] Kurdziaek S, Górecki W, Albarelli F and Demkowicz-Dobrzański R 2023 *Phys. Rev. Lett.* **131** 090801
[17] Caves C M 1981 *Phys. Rev.* D **23** 1693
[18] Ulam-Orgikh D and Kitagawa M 2001 *Phys. Rev.* A **64** 052106
[19] Danilishin S L and Khalili F Y 2012 *Living Rev. Relativ.* **15** 5
[20] Schnabel R 2017 *Phys. Rep.* **684** 1
[21] Schulte M, Lisdat C, Schmidt P O, Sterr U and Hammerer K 2020 *Nat. Commun.* **11** 5955
[22] Kaubruegger R, Vasilyev D V, Schulte M, Hammerer K and Zoller P 2021 *Phys. Rev.* X **11** 041045
[23] Yang J, Pang S, Chen Z, Jordan A N and del Campo A 2022 *Phys. Rev. Lett.* **128** 160505
[24] Zhang M, Yu H-M, Yuan H, Wang X, Demkowicz-Dobrzański R and Liu J 2022 *Phys. Rev. Res.* **4** 043057
[25] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 *Phys. Rev.* A **54** R4649

[26] Giovannetti V, Lloyd S and Maccone L 2006 *Phys. Rev. Lett.* **96** 010401
[27] Berry D W and Wiseman H M 2000 *Phys. Rev. Lett.* **85** 5098
[28] Huelga S F, Macchiavello C, Pellizzari T, Ekert A K, Plenio M B and Cirac J I 1997 *Phys. Rev. Lett.* **79** 3865
[29] Dorner U, Demkowicz-Dobrzanski R, Smith B J, Lundeen J S, Wasilewski W, Banaszek K and Walmsley I A 2009 *Phys. Rev. Lett.* **102** 040403
[30] Knysh S I, Chen E H and Durkin G A 2014 arXiv:1402.0495
[31] Demkowicz-Dobrzański R, Jarzyna M and Kołodyński J 2015 *Progress in Optics* vol 60 ed E Wolf (Elsevier) pp 345–435
[32] Macieszczak K 2013 arXiv:1312.1356
[33] Tóth G and Vértesi T 2018 *Phys. Rev. Lett.* **120** 020506
[34] Demkowicz-Dobrzański R 2011 *Phys. Rev.* A **83** 061802
[35] Macieszczak K, Fraas M and Demkowicz-Dobrzański R 2014 *New J. Phys.* **16** 113002
[36] Demkowicz-Dobrzański R, Banaszek K and Schnabel R 2013 *Phys. Rev.* A **88** 041802
[37] Ma J, Wang X, Sun C and Nori F 2011 *Phys. Rep.* **509** 89
[38] Schollwöck U 2011 *Ann. Phys., NY* **326** 96
[39] Chabuda K, Dziarmaga J, Osborne T J and Demkowicz-Dobrzanski R 2020 *Nat. Commun.* **11** 250
[40] Chabuda K and Demkowicz-Dobrzański R 2022 *Comput. Phys. Commun.* **274** 108282
[41] Sekatski P, Skotiniotis M, Kołodyński J and Dür W 2017 *Quantum* **1** 27
[42] Chiribella G, D'Ariano G M and Perinotti P 2009 *Phys. Rev.* A **80** 022339
[43] Chiribella G 2012 *New J. Phys.* **14** 125008
[44] Altherr A and Yang Y 2021 *Phys. Rev. Lett.* **127** 060501
[45] Liu Q, Hu Z, Yuan H and Yang Y 2023 *Phys. Rev. Lett.* **130** 070803
[46] Durkin G A and Dowling J P 2007 *Phys. Rev. Lett.* **99** 070801
[47] Hayashi M 2011 *Commun. Math. Phys.* **304** 689
[48] Rubio J and Dunningham J 2020 *Phys. Rev.* A **101** 032114
[49] Demkowicz-Dobrzański R, Górecki W and Guţă M 2020 *J. Phys. A: Math. Theor.* **53** 363001
[50] Meyer J J, Khatri S, Stilck França D, Eisert J, and Faist P 2023 arXiv:2307.06370
[51] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (Academic)
[52] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
[53] Kolodynski J and Demkowicz-Dobrzanski R 2013 *New J. Phys.* **15** 073043
[54] Diamond S and Boyd S 2016 *J. Mach. Learn. Res.* **17** 1 (arXiv:1603.00943 [math.OC])
[55] Boyd S and Vandenberghe L 2004 *Convex Optimization* (Cambridge University Press)
[56] Bavaresco J, Lipka-Bartosik P, Sekatski P, and Mehboudi M 2023 arXiv:2311.01513
[57] Tóth G, Vértesi T, Horodecki P and Horodecki R 2020 *Phys. Rev. Lett.* **125** 020402
[58] Bengtsson I and Zyczkowski K 2006 *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge Univeristy Press)
[59] Bisio A, D'Ariano G M, Perinotti P and Chiribella G 2011 *Phys. Rev.* A **83** 022325
[60] Younis E, Iancu C C, Lavrijsen W, Davis M and Smith E (USDOE) 2021 Berkeley quantum synthesis toolkit (bqskit) v1 (https://doi.org/10.11578/dc.20210603.2)
[61] Iten R, Reardon-Smith O, Malvetti E, Mondada L, Pauvert G, Redmond E, Kohli R S and Colbeck R 2021 Introduction to UniversalQCompiler (arXiv:1904.01072 [quant-ph])
[62] Fannes M, Nachtergaele B and Werner R F 1992 *Commun. Math. Phys.* **144** 443
[63] MOSEK ApS 2024 The MOSEK optimization toolbox for Python manual. Version 10.2. (available at: http://docs.mosek.com/latest/pythonapi/index.html)
[64] Arrad G, Vinkler Y, Aharonov D and Retzker A 2014 *Phys. Rev. Lett.* **112** 150801
[65] Kessler E M, Lovchinsky I, Sushkov A O and Lukin M D 2014 *Phys. Rev. Lett.* **112** 150802
[66] Dür W, Skotiniotis M, Fröwis F and Kraus B 2014 *Phys. Rev. Lett.* **112** 080801
[67] Jarzyna M and Demkowicz-Dobrzański R 2013 *Phys. Rev. Lett.* **110** 240405
[68] Dorner U 2012 *New J. Phys.* **14** 043011
[69] Jeske J, Cole J H and Huelga S F 2014 *New J. Phys.* **16** 073039
[70] Layden D and Cappellaro P 2018 *npj Quantum Inf.* **4** 30
[71] Altenburg S, Oszmaniec M, Wölk S and Gühne O 2017 *Phys. Rev.* A **96** 042319
[72] Zhou S 2024 arXiv:2402.18765
[73] Katariya V and Wilde M M 2021 *Quantum Inf. Process* **20** 78
[74] Bavaresco J, Murao M and Quintino M T 2021 *Phys. Rev. Lett.* **127** 200504
[75] Katariya V and Wilde M M 2021 *Phys. Rev.* A **104** 052406
[76] Bergh B, Datta N, Salzmann R and Wilde M M 2024 *IEEE Trans. Inf. Theory* **70** 2617
[77] Bavaresco J, Murao M and Quintino M T 2022 *J. Math. Phys.* **63** 042203
[78] Albarelli F, Friel J F and Datta A 2019 *Phys. Rev. Lett.* **123** 200503
[79] Hayashi M and Ouyang Y 2024 arXiv:2401.05886
[80] Pollock F A, Rodríguez-Rosario C, Frauenheim T, Paternostro M and Modi K 2018 *Phys. Rev.* A **97** 012127
[81] White G, Pollock F, Hollenberg L, Modi K and Hill C 2022 *PRX Quantum* **3** 020344
[82] Butler E P, Fux G E, Ortega-Taberner C, Lovett B W, Keeling J and Eastham P R 2024 *Phys. Rev. Lett.* **132** 060401
[83] Liu Q and Yang Y 2024 Efficient tensor networks for control-enhanced quantum metrology arXiv:2403.09519
[84] Orús R 2014 *Ann. Phys., NY* **349** 117

Erratum

Equation (D9) in appendix D is

$$|\chi_i\rangle = |\psi\rangle_{c_i t_i \varphi} \otimes |0\rangle - i\sqrt{1-t_i^2}\frac{c_i\varphi}{2}\,|-\rangle \otimes |1\rangle\,,$$

while the correct version should be

$$|\chi_i\rangle = \sqrt{1-\frac{c_i^2\varphi^2}{4}(1-t_i^2)}\,|\psi\rangle_{c_i t_i \varphi} \otimes |0\rangle - i\sqrt{1-t_i^2}\frac{c_i\varphi}{2}\,|-\rangle \otimes |1\rangle\,.$$

This minor mistake does not affect any furhter calculations, in particular, the formulas for probabilities $p_{1,i}$ and p_0 are correct.

Chapter 3

Using adaptiveness and causal superpositions against noise in quantum metrology

Commentary

This work establishes asymptotically tight bounds for adaptive strategies with uncorrelated channels (C2, Fig. 1.1). Although noise and signal correlations are not studied, the methodology developed here proves essential for tackling the correlated scenarios. This is demonstrated in Chapter 4, where the bound for schemes C3 is derived.

In Chapter 2, it was shown how to use the MOP technique to compute the channel QFI, see eq. (5) in Section 2.1. This formula can be, in principle, easily extended to the case of parallel estimation of N channels (B2), simply by replacing Λ_φ with $\Lambda_\varphi^{\otimes N}$

Unfortunately, the resulting minimization problem quickly becomes unfeasible because its complexity grows exponentially with N. Nevertheless, for large N (or even asymptotically), the MOP is still useful to derive the QFI *upper bounds*. To demonstrate this, we briefly recapitulate the derivation of an upper bound for parallel schemes (B2); details can be found in Refs. [Fuj08; Dem12].

The inequality

$$F_Q(\Lambda_\varphi) \leq 4\|\alpha\|, \quad \alpha = \sum_{k=1}^{r} \dot{K}_{\varphi,k}^\dagger \dot{K}_{\varphi,k}, \tag{3.1}$$

holds for any Kraus representation $\{K_{\varphi,k} : k \in \{1, \ldots, r\}\}$ of Λ_φ. This is a direct consequence of Eq. (5) from the work presented in Chapter 2, whose notation we adopt here.

We apply this inequality to $\Lambda_\varphi^{\otimes N}$. Among the many possible Kraus representations of

$\Lambda_\varphi^{\otimes N}$, we choose the one formed by tensor products of single-channel Kraus operators:

$$K_{\varphi,\boldsymbol{k}}^{(N)} = \bigotimes_{i=1}^{N} K_{\varphi,k_i}, \quad \boldsymbol{k} = (k_1, k_2, \ldots, k_N) \in \{1, 2, \ldots, r\}^{\times N}, \tag{3.2}$$

which leads to

$$F_Q(\Lambda_\varphi^{\otimes N}) \leq 4\|\alpha^{(N)}\|, \quad \alpha^{(N)} = \sum_{\boldsymbol{k}\in\{1,2,\ldots,r\}^{\times N}} \dot{K}_{\varphi,\boldsymbol{k}}^{(N)\dagger} \dot{K}_{\varphi,\boldsymbol{k}}^{(N)}. \tag{3.3}$$

This construction still leaves us the freedom of choosing any Kraus representation of a single channel.

Expanding $\alpha^{(N)}$ under the operator norm in (3.3), and using a triangle inequality (see [Fuj08] for details), one gets

$$\|\alpha^{(N)}\| \leq N\|\alpha\| + N(N-1)\|\beta\|^2, \tag{3.4}$$

where

$$\alpha = \sum_{k=1}^{r} \dot{K}_{\varphi,k}^\dagger \dot{K}_{\varphi,k}, \quad \beta = \sum_{k=1}^{r} \dot{K}_{\varphi,k}^\dagger K_{\varphi,k} \tag{3.5}$$

are constructed using a *single* channel Kraus representation, which makes them numerically tractable. The above reasoning yields the following bounds

$$F_Q(\Lambda_\varphi^{\otimes N}) \leq 4\left(N\|\alpha\| + N(N-1)\|\beta\|^2\right), \tag{3.6}$$

which can be further tightened by minimizing over Kraus representations of a single channel Λ_φ,

$$F_Q(\Lambda_\varphi^{\otimes N}) \leq 4 \min_{\{K_k\}} \left(N\|\alpha\| + N(N-1)\|\beta\|^2\right). \tag{3.7}$$

This minimization can be translated into an SDP [Dem12; Koł13] by parametrizing all relevant Kraus representations using a hermitian $r \times r$ matrix h, as we did in eq. (6) in Chapter 2. Importantly, the complexity of the resulting SDP is constant with N. From this bound, it is clear that when $\beta = 0$ for *at least one* Kraus representation of Λ_φ, the QFI exhibits only SS. Conversely, HS scaling is possible only when $\beta \neq 0$ for *all* Kraus representations.

Even the optimized bound (3.7) is not guaranteed to be tight, as the minimization is restricted to the subclass of Kraus representations of $\Lambda_\varphi^{\otimes N}$ generated using (3.2), and inequality (3.4) is not always saturated. Nevertheless, this method yields informative bounds that are asymptotically saturable as $N \to \infty$ [Zho21].

Similar reasoning was used to derive bounds for adaptive schemes (C2) [Dem14; Zho18]. For SS cases, parallel and adaptive bounds proved asymptotically equivalent, establishing the asymptotic equivalence of B2 and C2 strategies. However, HS cases exhibited an asymptotic discrepancy between parallel and adaptive bounds, leaving unclear whether this reflected a genuine advantage or merely loose bounds.

This work resolves this question: previous adaptive bounds were not tight. We derive tighter adaptive bounds that are asymptotically equivalent to parallel bounds and more informative for finite N. This is achieved through an iterative approach that bounds the maximal QFI increase from adding a single channel to the sequential chain in C2. Such approach had been used earlier, but either led to loose bounds [Per21], or was restricted to continuous-time scenarios only [Wan22].

Additionally, we generalize our result to even broader class of metrological schemes, involving *causal superpositions* of different orders of channels probing [Zha20; Cha21; Liu23]. Interestingly, we prove that such schemes are also asymptotically equivalent to parallel schemes.

We also show how to apply the derived bounds to continuous-time models, which is further explored in Ref. [Das25b].

The adaptive scheme considered in this work may appear to be less general than the C2 scheme from Fig. 1.1 because general control operations C_i are replaced with unitary operations V_i. However, this does not lead to any loss of generality, as long as we assume the unlimited size of an ancillary system—then we can replace each control C_i with its purification, and such procedure will never decrease the resulting QFI.

My contribution

Together with WG, I proved the main result of the paper—the precision bound for adaptive metrological protocols—which I subsequently generalized independently to protocols involving causal superpositions. I developed and implemented all the numerical procedures presented in the publication, and generated and analyzed numerical data. I wrote appendices C and D, read and coedited the main text and remaining appendices, I am the corresponding author.

Using Adaptiveness and Causal Superpositions Against Noise in Quantum Metrology

Stanisław Kurdziałek,[1,*] Wojciech Górecki,[1,*] Francesco Albarelli,[2,3] and Rafał Demkowicz-Dobrzański[1]
[1]Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland
[2]Dipartimento di Fisica "Aldo Pontremoli," Università degli Studi di Milano, via Celoria 16, 20133 Milan, Italy
[3]Istituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, 20133 Milan, Italy

(Received 5 January 2023; accepted 11 July 2023; published 30 August 2023)

We derive new bounds on achievable precision in the most general adaptive quantum metrological scenarios. The bounds are proven to be asymptotically saturable and equivalent to the known parallel scheme bounds in the limit of a large number of channel uses. This completely solves a long-standing conjecture in the field of quantum metrology on the asymptotic equivalence between parallel and adaptive strategies. The new bounds also allow us to easily assess the potential benefits of invoking nonstandard causal superposition strategies, for which we prove, similarly to the adaptive case, the lack of asymptotic advantage over the parallel ones.

DOI: 10.1103/PhysRevLett.131.090801

Introduction.—In the field of quantum information and quantum technologies, one can distinguish three levels of quantumness that are behind the boost in performance of various communication [1,2], computational [3], or metrological tasks [4–6]. The most rudimentary one is quantum coherence (C), which refers to the potential of having a single quantum system in the state of quantum superposition. This is already enough to implement secure quantum key distribution protocols [7] or even reach the Heisenberg limit in noiseless quantum metrology, provided a given quantum probe can pass through a sensing channel multiple times [8,9]. The next level is entanglement (E), where quantum coherence present in multipartite systems manifests itself in the form of nonclassical correlations. This quantumness level is crucial to guarantee quantum speedup in computational tasks [10] as well as to assure the ultimate security in the so-called device-independent quantum key distribution [11]. In quantum metrology, it had long been appreciated as the way to boost the precision in optical and atomic interferometric tasks [12–16], either in the form of N00N states [17,18] or much more practical optical and atomic squeezed states [19,20]. Finally, exploiting the quantum potential to its limits, one can consider adaptive (AD) or active quantum feedback strategies, where the probes are entangled with noiseless ancillary systems, and quantum control operations may actively modify the probe system that will be sent to the subsequent channel based on the information obtained so far [21–25], see Fig. 1. Such protocols represent the most general channel sensing schemes, containing E as a special case and encompassing in particular all quantum error-correcting strategies widely used in the whole field of quantum information processing to counter noise [26–29].

Interestingly, in the absence of noise, AD strategies provide no advantage over optimal E strategies [30]. In the presence of noise, however, some advantages have been observed in the small-number-of-uses regime where a direct search of optimal metrological protocols could be carried out [21,31–35]. In 2014 a conjecture was formulated [21] predicting no asymptotic advantage of AD over E. A notable progress in answering this fundamental question was made in 2021 [36,37], when it was demonstrated that in the models where quantum coherence cannot be protected against noise on an arbitrary scale, and hence the Heisenberg scaling (HS) is not achievable, AD strategies offer no asymptotic advantage over E. Still, the full answer to the question was lacking, mainly due to the fact that the bounds used there were not tight enough.

In this Letter, utilizing our new bounds, we indeed answer the conjecture in an affirmative way, proving in full

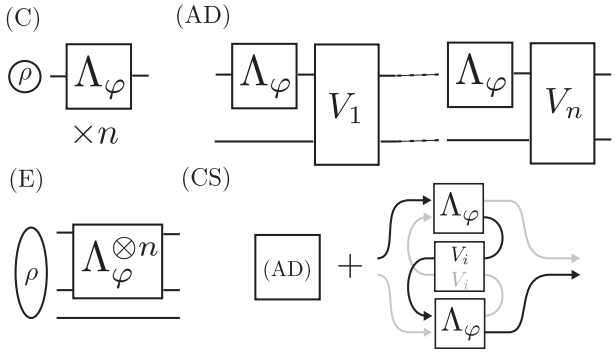


FIG. 1. Metrological schemes utilizing "four levels of quantumness": (C) channels probed independently (basic use of quantum coherence); (E) channels probed in parallel using a general entangled state, with ancillary systems potentially involved; (AD) general adaptive (active quantum feedback) strategies; (CS) causal superposition strategies, where additionally channels may be probed in a superposition of different causal orders.

generality that AD strategies provide no asymptotic advantage over E. As negative as it may sound, the result by no means implies that AD strategies are useless. In fact, our bounds allow us to clearly pinpoint the potential advantage one may expect in the finite number-of-uses regime, and easily observe how the advantage fades away when approaching the asymptotic limit of a large number of channel uses. On a more practical side, adaptive strategies may sometimes be in fact easier to implement than parallel, as they may not necessarily require entangling a large number of particles, while obtaining the same effect via small scale entanglement and active feedback. Even though the "three levels of quantumness" listed above appear to cover all quantum aspects of metrological protocols, an intriguing idea was put forward of considering causal superposition (CS) strategies where different channels are being probed in a superposition of different causal orders [35,38–44]. Advantages of such a strategy over the most general AD strategy have been observed, but no efficiently computable bounds have been proposed. In this Letter, we provide bounds valid also for this more general class of protocols and show their asymptotic equivalence to AD and E, which also means that CS strategies cannot surpass the HS [45].

Introductory example.—Let us start with the most elementary yet very illuminating example of a noisy metrological model, where it is possible to remove noise while assuring the preservation of HS of precision in the asymptotic regime. Consider a single qubit channel $\Lambda_\varphi(\cdot) = \sum_k K_{\varphi,k} \cdot K_{\varphi,k}^\dagger$, where $K_{\varphi,k} = U_\varphi K_k$,

$$U_\varphi = e^{-\frac{i}{2}\sigma_z\varphi}, \qquad K_1 = \sqrt{p}\mathbb{1}, \qquad K_2 = \sqrt{1-p}\sigma_x. \quad (1)$$

The channel represents dephasing of a qubit along the x axis of the Bloch ball (the operator K_2 may be understood as a σ_x error occurring with probability $1-p$) and the subsequent rotation U_φ of the state around the z axis by angle φ, where φ is the parameter to be estimated—a similar model has been used in an experimental demonstration of quantum error-correction enhanced metrology in NV-center sensing setups [46], as well as in [47] where the possibility of beating the standard scaling (SS) in the presence of transversal noise was shown. In the case of a single channel use, $n = 1$, the effect of noise may be completely mitigated by choosing the input state as $|\psi^{(1)}\rangle = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. This state is not affected by the σ_x error and the output state $|\psi_\varphi\rangle = (|0\rangle + e^{i\varphi}|1\rangle)/\sqrt{2}$ represents a noiseless phase encoding. We will quantify the performance of a given protocol using the quantum Fisher information (QFI) [48,49] of the output state, which in this case is $F^{(1)} = 1$ (we recall the definition of the QFI in [50], Section A).

Assume now that we can use the channel twice, $n = 2$. If we send the optimal single qubit probes independently to each of the channels, we get the QFI value $F_C^{(2)} = 2$. We can, however, also consider a parallel strategy involving an entangled input state $|\psi^{(2)}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ (the N00N state [52]). In this case if either zero or two σ_x errors occur, the final state will again correspond to the noiseless phase encoding $|\psi_\varphi^{(2)}\rangle = U_\varphi^2|\psi^{(2)}\rangle = (|00\rangle + e^{2i\varphi}|11\rangle)/\sqrt{2}$ for which the QFI equals 4. Whereas, if only a single σ_x occurs the state will contain no information about the phase at all. As a result the final QFI reads $F_E^{(2)} = 4[p^2 + (1-p)^2] \geq F_C^{(2)}$. Interestingly, this result may be further improved via a simple adaptive strategy. The protocol involves entangling the initial single probe qubit with a single ancillary qubit, so that the input state is again $|\psi^{(2)}\rangle$. After a single action of the channel, $\Lambda_\varphi \otimes \mathcal{I}$, an error correction operation is performed, where we check if a σ_x error occurred and correct the error accordingly. Then the channel acts on the probe state again, and with probability p yields the ideal state $|\psi_\varphi^{(2)}\rangle$; while if another σ_x error occurs, the final unitary rotation U_φ removes all the phase information from the state. Consequently, the protocol yields a QFI equal to $4p$. This protocol is actually the optimal one provided $p \geq 0.5$. If $p < 0.5$, then one simply needs to modify the recovery operation in a way that instead of correcting a single σ_x error on the probe system the σ_x operation is applied to the ancillary qubit. In the end the optimal QFI reads $F_{\mathrm{AD}}^{(2)} = 2(1 + |1-2p|) \geq F_E^{(2)}$ (see Ref. [50] Section A for details).

With this example in mind, one may wonder how to prove that the actual protocols are indeed optimal and what is (if any) the potential benefit of using even more general CS strategies ($F_{\mathrm{CS}}^{(2)} > F_{\mathrm{AD}}^{(2)}$). For larger n the task becomes even more challenging, and no brute-force optimization approach can tell what happens in the asymptotic limit $n \to \infty$. The methods developed in this Letter allow us to answer all these questions.

State-of-the-art bounds.—The most powerful state-of-the-art bounds for the performance of E as well as AD strategies, are based on the concept of minimization of certain operator norm expressions over different Kraus representations of the channel $\Lambda_\varphi = \sum_k K_{\varphi,k} \cdot K_{\varphi,k}^\dagger$ [21,22,28,31,37,53–56]—in what follows we drop subscript φ in Kraus operators for conciseness. For E strategies, the upper bound on the achievable QFI reads

$$F_E^{(n)} \leq \min_{\{K_k\}} 4[n\|\alpha\| + n(n-1)\|\beta\|^2], \quad (2)$$

where $\|\cdots\|$ denotes the operator norm, $\alpha = \sum_k \dot{K}_k^\dagger \dot{K}_k$, $\beta = \sum_k \dot{K}_k^\dagger K_k$, and $\dot{K}_k = \partial_\varphi K_k$. If a Kraus representation exists for which $\beta = 0$, the QFI scales asymptotically at most linearly with n—SS models—and the optimal quantum enhancement amounts to a constant factor improvement [21,54–56]. If no such representation exists, then the

HS can be preserved asymptotically [28,37]. Interestingly, the above bound has been proven to be asymptotically tight for both SS ($\beta = 0$) and HS ($\beta \neq 0$) models [37].

Moving to AD strategies, the best state-of-the-art universally valid bound reads [21,31,37]

$$F_{\mathrm{AD}}^{(n)} \leq \min_{\{K_i\}} 4\left[n\|\alpha\| + n(n-1)\|\beta\|\left(\|\beta\| + 2\sqrt{\|\alpha\|}\right)\right]. \quad (3)$$

It is asymptotically equivalent to the parallel bound, Eq. (2), in the case of SS models ($\beta = 0$), and, since the parallel bound is asymptotically saturable, this implies no asymptotic advantage of AD strategies over E. Still, the bound leaves space for improvement for finite n and does not exclude an asymptotic advantage for HS models—the term quadratic in n has a larger coefficient than the one in Eq. (2).

Iterative bound.—Below, we derive a tighter adaptive bound than the one given above, and prove it is asymptotically equivalent to the parallel one—consequently, this implies no asymptotic advantage of AD over E for all models (both SS and HS).

Let $\Lambda_\varphi^{(n)}(\cdot) = \sum_{\boldsymbol{k}^{(n)}} K_{\boldsymbol{k}^{(n)}} \cdot K_{\boldsymbol{k}^{(n)}}^\dagger$ represent a combined action of n channels Λ_φ in a general adaptive strategy where they are intertwined with control operations V_i acting on probe and ancillary systems, as in Fig. 1 (AD). $K_{\boldsymbol{k}^{(n)}}$ denote the corresponding Kraus operators, which can be computed via the following iteration relation: $K_{\boldsymbol{k}^{(1)}} = V_1(K_{k_1} \otimes \mathbb{1})$,

$$K_{\boldsymbol{k}^{(i+1)}} = V_{i+1}(K_{k_{i+1}} \otimes \mathbb{1})K_{\boldsymbol{k}^{(i)}}, \quad (4)$$

where $\boldsymbol{k}^{(i)} = (k_i, \ldots, k_1)$, and $\mathbb{1}$ is acting on the ancillary system (we will drop it in what follows for conciseness of notation).

The starting point for the derivation of the state-of-the art bounds as reported in Eqs. (2), (3), is an observation that, given a channel $\Lambda_\varphi^{(n)}$, maximization of the QFI of the output state over all inputs and sets of control operations can be upper bounded by [53]

$$F_{\mathrm{AD}}^{(n)} = \max_{\rho_0, \{V_i\}} F\left[\Lambda_\varphi^{(n)}(\rho_0)\right] \leq \max_{\{V_i\}} \min_{\{K_{\boldsymbol{k}^{(n)}}\}} 4\|\alpha^{(n)}\|, \quad (5)$$

where $\alpha^{(n)} = \sum_{\boldsymbol{k}^{(n)}} \dot{K}_{\boldsymbol{k}^{(n)}}^\dagger \dot{K}_{\boldsymbol{k}^{(n)}}$, the minimization is performed over all equivalent Kraus representations of $\Lambda_\varphi^{(n)}$. Note that for a large enough ancillary system the inequality becomes equality. As such, this inequality is not of much practical use due to the infeasibility of performing the minimization over all Kraus representations for larger values of n, as well as the need to additionally perform the optimization over the control operations V_i. The usefulness of this inequality stems from the fact, that it is possible to further upper bound the rhs of Eq. (5) with norms of operators defined in terms of Kraus operators of the *elementary channel* Λ_φ. This is how bounds (2) and (3) were obtained [21,31,37,54,55].

In what follows we provide a novel step-by-step approach, where at each step we bound the maximal *increase* in the final QFI thanks to the additional usage of a single quantum channel [57]. Using Eq. (4) we have

$$\begin{aligned}\alpha^{(i+1)} &= \sum_{k_{i+1}, \boldsymbol{k}^{(i)}} \left(K_{\boldsymbol{k}^{(i)}}^\dagger \dot{K}_{k_{i+1}}^\dagger + \dot{K}_{\boldsymbol{k}^{(i)}}^\dagger K_{k_{i+1}}^\dagger\right) \times \mathrm{H.c.} \\ &= \sum_{\boldsymbol{k}^{(i)}} K_{\boldsymbol{k}^{(i)}}^\dagger \alpha K_{\boldsymbol{k}^{(i)}} + K_{\boldsymbol{k}^{(i)}}^\dagger \beta \dot{K}_{\boldsymbol{k}^{(i)}} + \dot{K}_{\boldsymbol{k}^{(i)}}^\dagger \beta^\dagger K_{\boldsymbol{k}^{(i)}} + \alpha^{(i)}.\end{aligned} \quad (6)$$

We will now use the following operator norm inequality (see Ref. [50] Section B for the proof):

$$\left\|\sum_k L_k^\dagger A Q_k\right\| \leq \sqrt{\left\|\sum_k L_k^\dagger L_k\right\|}\|A\|\sqrt{\left\|\sum_k Q_k^\dagger Q_k\right\|}, \quad (7)$$

which, together with the triangle inequality and the trace preservation condition, $\sum_{\boldsymbol{k}^{(i)}} K_{\boldsymbol{k}^{(i)}}^\dagger K_{\boldsymbol{k}^{(i)}} = \mathbb{1}$, yields

$$\|\alpha^{(i+1)}\| \leq \|\alpha^{(i)}\| + \|\alpha\| + 2\|\beta\|\sqrt{\|\alpha^{(i)}\|}. \quad (8)$$

Let us define the following iteration:

$$a^{(i+1)} = a^{(i)} + \|\alpha\| + 2\|\beta\|\sqrt{a^{(i)}}, \qquad a^{(0)} = 0, \quad (9)$$

which, in light of Eqs. (5) and (8), yields $F_{\mathrm{AD}}^{(n)} \leq 4a^{(n)}$. The resulting bound $4a^{(n)}$ may be optimized over the choice of Kraus representation of the elementary channel in each iteration *separately* (how to efficiently implement this iteration numerically is described in [50], Section D) or, in a weaker variant, over a *single* Kraus representation identically used in each step (for which the resulting bound will also be valid for CS strategies—see Ref. [50] Section C for the proof). Since $a^{(n)}$ is strategy independent, the maximization over $\{V_i\}$, or, more generally, over all CS strategies, is no longer necessary. This finally yields

$$F_{\mathrm{AD}}^{(n)} \leq \min_{\{K_k\}^{\times n}} 4a^{(n)}, \qquad F_{\mathrm{CS}}^{(n)} \leq \min_{\{K_k\}} 4a^{(n)}. \quad (10)$$

Interestingly, the possibility to use a different Kraus representation for each channel use allows us to tighten the bound also for parallel strategies, see Ref. [50] Section D3.

Closed formula bounds.—In order to appreciate how much tighter the obtained bounds are compared to the state-of-the-art bounds, we will provide some closed formulas for the bounds that result from a relaxed variants of the iteration procedure. First, observe that from Eq. (7) we get

$\|\beta\| \leq \sqrt{\|\alpha\|}$. From Eq. (9) it then follows that $a^{(n)} \leq n^2\|\alpha\|$ (the bound obtained in [34]), which when put back into the iteration formula results in

$$F^{(n)}_{\mathrm{AD,CS}} \leq \min_{\{K_k\}} 4\left(n\|\alpha\| + n(n-1)\|\beta\|\sqrt{\|\alpha\|}\right). \tag{11}$$

Note, that the bound is noticeably tighter than Eq. (3) and is also valid for CS strategies, as the same Kraus representation is used in each step. We also see that the difference between this bound and Eq. (2) amounts to replacing one $\|\beta\|$ with $\sqrt{\|\alpha\|}$. It might be tempting to conjecture that this difference reflects the asymptotic gain of AD over E strategies. This is not the case, however, as we demonstrate below.

For any fixed $\|\alpha\|$, $\|\beta\|$ consider the following function $f(n) = n\|\alpha\| + n(n-1)\|\beta\|^2 + n\log n(\|\alpha\| - \|\beta\|^2)$. For $n \geq 0$ it can be shown (see Ref. [50] Section E) that $f(n+1) \geq f(n) + \|\alpha\| + 2\|\beta\|\sqrt{f(n)}$. Hence, in light of Eq. (9) we get $f(n) \geq a^{(n)}$ and as a result

$$F^{(n)}_{\mathrm{AD,CS}} \leq \min_{\{K_k\}} 4\left[n\|\alpha\| + n(n-1)\|\beta\|^2\left(1 + \frac{c\log n}{n-1}\right)\right], \tag{12}$$

where $c = (\|\alpha\| - \|\beta\|^2)/\|\beta\|^2$. Since we know that the parallel bound, Eq. (2), is asymptotically saturable this implies that

$$\lim_{n\to\infty}\left(F^{(n)}_{\mathrm{AD,CS}}/F^{(n)}_{E}\right) = 1 \tag{13}$$

and, hence, there is no asymptotic advantage of AD nor CS over E.

Interestingly, lack of asymptotic advantage thanks to adaptiveness has also been demonstrated for continuous-time models [25], a result which can be regarded as a limiting case of the theory we develop here (see Ref. [50] Section F for details).

Examples.—In order to illustrate the practical applications of the bounds, we compute them for four representative models and compare the results with the actual performance of the optimal protocols that can be determined numerically for a small number of channel uses ($n \leq 4$) via semidefinite programming (SDP) as described in [21] (parallel strategies), [33] (adaptive protocols), and [35] (causal superposition protocols). The results are presented in Fig. 2. As a figure of merit we plot the achievable QFI with n uses of a channel normalized by n times $F^{(1)}$ (the maximal QFI for single-channel sensing with a possible use of ancillary systems).

Figure 2(a) presents results corresponding to the introductory example of the perpendicular dephasing model, Eq. (1)—in all the models that follow we also assume the convention that $K_{\varphi,k} = U_\varphi K_k$ (signal comes after noise). Among the four models presented, this is the only one that admits asymptotic HS—hence the linear increase of the figure of merit. Interestingly, the bounds are saturated for $n = 2$ and the optimal QFI values are equal to the ones obtained for the protocols discussed in the introductory example, proving they are indeed optimal. For larger n, the bounds are very tight, and, as expected, the bounds for AD and CS converge asymptotically to the E bound (unlike the state-of-the-art bound).

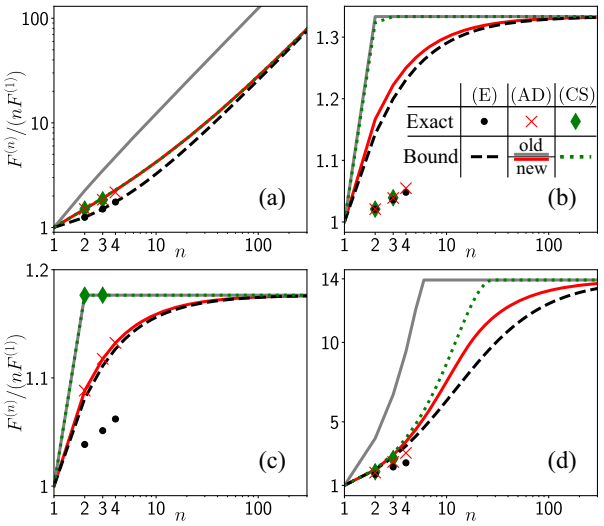


FIG. 2. Achievable QFI as a function of the number of channels probed for parallel (E, black), adaptive (AD, red), and causal superposition strategies (CS, green) normalized by n times the single-channel QFI. Points represent the result of the exact optimization, while curves represent the respective bounds. The best previously known adaptive bound (gray) is depicted for comparison. The four plots correspond to different metrological models with a qualitatively different behavior: (a) dephasing perpendicular to the signal, Eq. (1) ($p = 0.75$); (b) dephasing parallel to the signal, Eq. (14) ($p = 0.75$); (c) damping perpendicular to the signal, Eq. (15) ($p = 0.15$); (d) damping parallel to the signal, Eq. (16) ($p = 0.75$).

Results depicted in Fig. 2(b) refer to the parallel dephasing model (both the unitary encoding and the dephasing are with respect to the z axis), where the Kraus operators read

$$K_1 = \sqrt{p}\mathbb{1}, \qquad K_2 = \sqrt{1-p}\sigma_z. \tag{14}$$

In this case, gains due to adaptiveness or causal superpositions are very modest, and the bounds are not particularly tight for low n—still, thanks to the general theorem, we know they are tight asymptotically.

Figure 2(c) illustrates results for the perpendicular amplitude damping model (unitary encoding with respect to the z axis, amplitude damping with respect to the x axis):

$$K_1 = |-\rangle\langle -| + \sqrt{p}|+\rangle\langle +|, \qquad K_2 = \sqrt{1-p}|-\rangle\langle +|, \tag{15}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ are the eigenvectors of σ_x. This model is of particular interest as the finite-n bounds are saturated here both for AD and CS for all n. This suggests that it is highly unlikely that any tighter metrological bounds can be derived solely from the properties of the single-channel Kraus operators.

Finally, Fig. 2(d) depicts results for the parallel amplitude damping model with

$$K_1 = |0\rangle\langle 0| + \sqrt{p}|1\rangle\langle 1|, \qquad K_2 = \sqrt{1-p}|0\rangle\langle 1|. \quad (16)$$

This model illustrates particulary well how much tighter the novel bounds are, when compared with the previous state-of-the-art ones.

Conclusions and open problems.—With the results presented in this Letter, we dare to say that the theory of single-parameter quantum metrology in the presence of uncorrelated noise is now complete. Universal asymptotically saturable bounds are known as well as efficiently computable bounds in the regime of finite (but potentially large) number of channel uses. This, together with exact algorithms to find optimal protocols for small n, provides a complete landscape of achievable quantum enhancement in realistic quantum metrology. This said, we need to admit that in the case of multiparameter models [58,59], Bayesian models [60,61], and most importantly models involving temporally or spatially correlated noise [33,62–65], the quest for a full understanding of quantum metrological potential is still not complete. Nevertheless, these achievements compare favorably to the ones obtained in the related field of (binary) quantum channel discrimination [66,67]. Interestingly, adaptive strategies are not useful asymptotically for asymmetric hypothesis testing [68–70], while an advantage is possible in the symmetric setting [71,72]. However, easily computable asymptotic bounds, as well as practical strategies to attain them for arbitrary channels are still missing, unlike in quantum metrology. Moreover, the asymptotic analysis of causal superposition strategies for quantum channel discrimination [44,73] is still an open question.

This work was supported by the National Science Center (Poland) Grant No. 2020/37/B/ST2/02134. F. A. acknowledges financial support from MUR under the "PON Ricerca e Innovazione 2014-2020" project EEQU.

[*] These authors contributed equally to this work.

[1] N. Gisin and R. Thew, Quantum communication, Nat. Photonics **1**, 165 (2007).
[2] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Secure quantum key distribution with realistic devices, Rev. Mod. Phys. **92**, 025002 (2020).
[3] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).
[5] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).
[6] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nat. Photonics **12**, 724 (2018).
[7] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, New York, 1984), pp. 175–179.
[8] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free heisenberg-limited phase estimation, Nature (London) **450**, 393 (2007).
[9] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Quantum-enhanced measurements without entanglement, Rev. Mod. Phys. **90**, 035006 (2018).
[10] R. Jozsa and N. Linden, On the role of entanglement in quantum-computational speed-up, Proc. R. Soc. A **459**, 2011 (2003).
[11] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography Against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).
[12] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Improvement of Frequency Standards with Quantum Entanglement, Phys. Rev. Lett. **79**, 3865 (1997).
[13] L. Pezzé and A. Smerzi, Entanglement, Nonlinear Dynamics, and the Heisenberg Limit, Phys. Rev. Lett. **102**, 100401 (2009).
[14] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum limits in optical interferometry, in *Progress in Optics*, Volume 60, edited by E. Wolf (Elsevier, Amsterdam, 2015), pp. 345–435.
[15] J. P. Dowling and K. P. Seshadreesan, Quantum optical technologies for metrology, sensing, and imaging, J. Lightwave Technol. **33**, 2359 (2015).
[16] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).
[17] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Optimal frequency measurements with maximally correlated states, Phys. Rev. A **54**, R4649 (1996).
[18] J. P. Dowling, Quantum optical metrology—the lowdown on high-n00n states, Contemp. Phys. **49**, 125 (2008).
[19] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D **23**, 1693 (1981).
[20] R. Schnabel, Squeezed states of light and their applications in laser interferometers, Phys. Rep. **684**, 1 (2017).
[21] R. Demkowicz-Dobrzański and L. Maccone, Using Entanglement Against Noise in Quantum Metrology, Phys. Rev. Lett. **113**, 250801 (2014).
[22] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive Quantum Metrology under General Markovian Noise, Phys. Rev. X **7**, 041009 (2017).
[23] S. Pang and A. N. Jordan, Optimal adaptive control for quantum metrology with time-dependent Hamiltonians, Nat. Commun. **8**, 14695 (2017).

[24] S. Pirandola and C. Lupo, Ultimate Precision of Adaptive Noise Estimation, Phys. Rev. Lett. **118**, 100502 (2017).
[25] K. Wan and R. Lasenby, Bounds on adaptive quantum metrology under Markovian noise, Phys. Rev. Res. **4**, 033092 (2022).
[26] D. Kribs, R. Laflamme, and D. Poulin, Unified and Generalized Approach to Quantum Error Correction, Phys. Rev. Lett. **94**, 180501 (2005).
[27] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).
[28] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).
[29] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Ancilla-Free Quantum Error Correction Codes for Quantum Metrology, Phys. Rev. Lett. **122**, 040502 (2019).
[30] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, Phys. Rev. Lett. **96**, 010401 (2006).
[31] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, Quantum **1**, 27 (2017).
[32] Y. Yang, Memory Effects in Quantum Metrology, Phys. Rev. Lett. **123**, 110501 (2019).
[33] A. Altherr and Y. Yang, Quantum Metrology for Non-Markovian Processes, Phys. Rev. Lett. **127**, 060501 (2021).
[34] J. L. Pereira, L. Banchi, and S. Pirandola, Bounding the Benefit of Adaptivity in Quantum Metrology Using the Relative Fidelity, Phys. Rev. Lett. **127**, 150501 (2021).
[35] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal Strategies of Quantum Metrology with a Strict Hierarchy, Phys. Rev. Lett. **130**, 070803 (2023).
[36] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, Phys. Rev. Res. **2**, 013235 (2020).
[37] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).
[38] M. Araújo, F. Costa, and i. c. v. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, Phys. Rev. Lett. **113**, 250402 (2014).
[39] C. Mukhopadhyay, M. K. Gupta, and A. K. Pati, Superposition of causal order as a metrological resource for quantum thermometry, arXiv:1812.07508.
[40] M. Frey, Indefinite causal order aids quantum depolarizing channel identification, Quantum Inf. Process. **18**, 96 (2019).
[41] X. Zhao, Y. Yang, and G. Chiribella, Quantum Metrology with Indefinite Causal Order, Phys. Rev. Lett. **124**, 190503 (2020).
[42] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, Phys. Rev. A **103**, 032615 (2021).
[43] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, Quantum circuits with classical versus quantum control of causal order, PRX Quantum **2**, 030335 (2021).
[44] J. Bavaresco, M. Murao, and M. T. Quintino, Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination, Phys. Rev. Lett. **127**, 200504 (2021).
[45] A CS strategy was shown to achieve super-Heisenberg scaling in a quantum metrology problem with infinite-dimensional systems [41]. This does not contradict our results, since in this work we derive bounds for finite-dimensional systems.
[46] T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M. B. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A. O. Sushkov, M. D. Lukin, A. Retzker, B. Naydenov, L. P. McGuinness, and F. Jelezko, Quantum Metrology Enhanced by Repetitive Quantum Error Correction, Phys. Rev. Lett. **116**, 230502 (2016).
[47] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, Noisy Metrology Beyond the Standard Quantum Limit, Phys. Rev. Lett. **111**, 120401 (2013).
[48] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[49] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, Phys. Rev. Lett. **72**, 3439 (1994).
[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.090801 for more information, which includes Ref. [51].
[51] P. Sekatski and M. Perarnau-Llobet, Optimal nonequilibrium thermometry in Markovian environments, Quantum **6**, 869 (2022).
[52] We call it a N00N state because $|\psi^{(2)}\rangle = (|2,0\rangle + |0,2\rangle)/\sqrt{2}$ in occupation number representation.
[53] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, J. Phys. A **41**, 255304 (2008).
[54] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, Nat. Phys. **7**, 406 (2011).
[55] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, Nat. Commun. **3**, 1063 (2012).
[56] J. Kolodynski and R. Demkowicz-Dobrzanski, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).
[57] Similar philosophy may be found in [25,34], yet the final results obtained there lacked either generality or tightness.
[58] M. Szczykulska, T. Baumgratz, and A. Datta, Multi-parameter quantum metrology, Adv. Phys. **1**, 621 (2016).
[59] F. Albarelli and R. Demkowicz-Dobrzański, Probe Incompatibility in Multiparameter Noisy Quantum Metrology, Phys. Rev. X **12**, 011039 (2022).
[60] M. J. W. Hall and H. M. Wiseman, Heisenberg-style bounds for arbitrary estimates of shift parameters including prior information, New J. Phys. **14**, 033040 (2012).
[61] J. Rubio and J. Dunningham, Bayesian multiparameter quantum metrology with limited data, Phys. Rev. A **101**, 032114 (2020).
[62] Y. Matsuzaki, S. C. Benjamin, and J. Fitzsimons, Magnetic field sensing beyond the standard quantum limit under the effect of decoherence, Phys. Rev. A **84**, 012103 (2011).
[63] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum Metrology in Non-Markovian Environments, Phys. Rev. Lett. **109**, 233601 (2012).
[64] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański, Ultimate Precision Limits for Noisy Frequency Estimation, Phys. Rev. Lett. **116**, 120801 (2016).

[65] F. Beaudoin, L. M. Norris, and L. Viola, Ramsey interferometry in correlated quantum noise environments, Phys. Rev. A **98**, 020102(R) (2018).
[66] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, Fundamental limits to quantum channel discrimination, npj Quantum Inf. **5**, 50 (2019).
[67] V. Katariya and M. M. Wilde, Geometric distinguishability measures limit quantum channel estimation and discrimination, Quantum Inf. Process. **20**, 78 (2021).
[68] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, Lett. Math. Phys. **110**, 2277 (2020).
[69] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, Phys. Rev. Res. **1**, 033169 (2019).
[70] K. Fang, O. Fawzi, R. Renner, and D. Sutter, Chain Rule for the Quantum Relative Entropy, Phys. Rev. Lett. **124**, 100501 (2020).
[71] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Adaptive versus nonadaptive strategies for quantum channel discrimination, Phys. Rev. A **81**, 032339 (2010).
[72] F. Salek, M. Hayashi, and A. Winter, Usefulness of adaptive strategies in asymptotic quantum channel discrimination, Phys. Rev. A **105**, 022419 (2022).
[73] J. Bavaresco, M. Murao, and M. T. Quintino, Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies, J. Math. Phys. (N.Y.) **63**, 042203 (2022).

Using adaptiveness and causal superpositions against noise in quantum metrology: Supplementary Material

Stanisław Kurdziałek,[1,*] Wojciech Górecki,[1,*] Francesco Albarelli,[2,3] and Rafał Demkowicz-Dobrzański[1]

[1] *Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland*
[2] *Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, via Celoria 16, 20133 Milan, Italy*
[3] *Istituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, 20133 Milan, Italy*

Appendix A: Optimal estimation strategies in case of the perpendicular dephasing model ($n = 2$).

Before presenting the analysis of the example, let us recall the definition of the QFI for completeness of the presentation. Given a parameter dependent state ρ_φ, the corresponding QFI reads:

$$F(\rho_\varphi) = \mathrm{Tr}\left(\rho_\varphi L_\varphi^2\right), \quad \frac{d\rho_\varphi}{d\varphi} = \frac{1}{2}\left(\rho_\varphi L_\varphi + L_\varphi \rho_\varphi\right), \quad \text{(A1)}$$

where L_φ is the symmetric logarithmic derivative (SLD) and can be computed from the rightmost formula above. For pure states, the formula simplifies to

$$F(|\psi_\varphi\rangle) = 4\left(\langle\dot{\psi}_\varphi|\dot{\psi}_\varphi\rangle - |\langle\dot{\psi}_\varphi|\psi_\varphi\rangle|^2\right), \quad \text{(A2)}$$

where $|\dot{\psi}_\varphi\rangle = \frac{d|\psi_\varphi\rangle}{d\varphi}$. Operationally, the inverse of the QFI provides a lower bound on the minimal achievable variance of the estimated parameter, via the quantum Cramér-Rao bound [1, 2], $\Delta^2\tilde{\varphi} \geq 1/F$.

In particular for ideal phase encoding, where the phase is coherently imprinted k times on a balanced superposition state: $|\psi_\varphi\rangle = (|0\rangle + e^{ik\varphi}|1\rangle)/\sqrt{2}$, $F = k^2$. This implies, the result $F^{(1)} = 1$, for single use, $n = 1$, of the channel in the example discussed in the main text.

Given two uses of a channel, $n = 2$, and utilizing no entanglement or adaptiveness, the most straightforward strategy is to use the two channels independently, and send optimal single qubit probes to each of them. In this case we get the QFI $F_{\mathrm{C}}^{(2)} = 2$ (twice the single-channel QFI).

This value may be improved by considering a parallel strategy involving an entangled input state $|\psi^{(2)}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ (the N00N state), which results in the output state of the form:

$$\rho_{\varphi,\mathrm{E}}^{(2)} = \Lambda_\varphi^{\otimes 2}\left(|\psi^{(2)}\rangle\langle\psi^{(2)}|\right) = [p^2 + (1-p)^2]\,|\psi_\varphi^{(2)}\rangle\langle\psi_\varphi^{(2)}| + 2p(1-p)\,|\phi^{(2)}\rangle\langle\phi^{(2)}|, \quad \text{(A3)}$$

where $|\psi_\varphi^{(2)}\rangle = (|00\rangle + e^{2i\varphi}|11\rangle)/\sqrt{2}$, is the state corresponding to an effectively noiseless phase encoding (when two or none σ_x errors occur), while the orthogonal state $|\phi^{(2)}\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, contains no φ information (when one σ_x error occurred). As a result the QFI reads $F_{\mathrm{E}}^{(2)} = (p^2 + (1-p)^2) \times 4 \geq 2$—larger than when the two channels are used independently.

Interestingly, this result may be further improved via a simple adaptive strategy. The protocol involves entangling the initial single probe qubit with a single ancillary qubit in the same state as in the parallel strategy: $|\psi^{(2)}\rangle$. After a single action of the channel, $\lambda_\varphi \otimes \mathcal{I}$, an error correction operation is performed $\mathcal{R}$, where we check if a σ_x error occurred by projecting the state on one of the two subspaces $\mathcal{C} = \mathrm{span}(|00\rangle, |11\rangle)$ and $\mathcal{E} = \mathrm{span}(|10\rangle, |01\rangle)$, and correct the error accordingly. Afterwards, the second action of the channel is applied. The resulting state at the end of the protocol reads:

$$\rho_{\varphi,\mathrm{AD}}^{(2)} = (\Lambda_\varphi \otimes \mathcal{I}) \circ \mathcal{R} \circ (\Lambda_\varphi \otimes \mathcal{I})\left(|\psi^{(2)}\rangle\langle\psi^{(2)}|\right) = p\,|\psi_\varphi^{(2)}\rangle\langle\psi_\varphi^{(2)}| + (1-p)\,|\phi^{(2)}\rangle\langle\phi^{(2)}|. \quad \text{(A4)}$$

This protocols yield $F_{\mathrm{AD}}^{(2)} = 4p$ and is optimal provided $p \geq 0.5$. If $p < 0.5$, then one simply needs to modify the recovery operation in a way that instead of correcting a single σ_x error on the probe system the σ_x operation is applied to the ancillary qubit. This yields effectively $F_{\mathrm{AD}}^{(2)} = 4(1-p)$. So in the end the optimal QFI reads $F_{\mathrm{AD}}^{(2)} = 2(1 + |1-2p|)$, which is always larger than the parallel $F_{\mathrm{E}}^{(2)}$ except for the $p = 0.5$ case when they are equal.

Having said that, it needs to be noticed that if we restrict ourselves to utilizing coherence only, we should in principle also allow a scenario where a single probe goes through the two channels sequentially. Consider $|+\rangle$ as an input (which is the optimal choice). As a result of going through the the two channels the output state reads:

$$\rho_{\varphi,\mathrm{C}}^{(2)} = p\,|\psi_{2\varphi}\rangle\langle\psi_{2\varphi}| + (1-p)\,|+\rangle\langle+|. \quad \text{(A5)}$$

This state looks very similar to the output state resulting form the adaptive strategy, Eq. (A4). Still, in Eq. (A4) we may perform a measurement that unambiguously discriminates between the ideally phase-encoded state, $|\psi_\varphi^{(2)}\rangle$, and the non-informative state $|\phi^{(2)}\rangle$. In case of Eq. (A5) this is no longer possible as the "signal" state $|\psi_{2\varphi}\rangle$ and non-informative $|+\rangle$ are not orthogonal in general and hence one cannot simply filter out the

* These two authors contributed equally to the project.

non-informative term. Indeed, direct computation of the QFI, using Eq. (A1), for the state $\rho^{(2)}_{\varphi,C}$ yields

$$F^{(2)}_{\mathrm{C}} = 2p[1 + p + (1-p)\cos 2\varphi]. \tag{A6}$$

While in general this value is smaller than $F^{(2)}_{\mathrm{AD}}$, it approaches the performance of the adaptive scheme for $\varphi \approx 0$, and $p \geq 0.5$, $F^{(2)}_{\mathrm{C}} \approx 4p$. This shows that sometimes even a basic sequential scheme (utilizing only coherence, and no adaptiveness) may be superior to the optimal entanglement-based scheme.

It should be pointed out, however, that for larger n the simple sequential strategy will quickly become ineffective due to a build-up of unchecked decoherence effects which cause the resulting QFI to be damped. In this regime, more advanced strategies are required (E, AD or CS) in order to exploit the full potential of quantum enhanced sensing.

Appendix B: Operator norm inequality

Here we prove Eq. (7),

$$\left\|\sum_{k=1}^{K} L_k^\dagger A Q_k\right\| \leq \sqrt{\left\|\sum_{k=1}^{K} L_k^\dagger L_k\right\|}\|A\|\sqrt{\left\|\sum_{k=1}^{K} Q_k^\dagger Q_k\right\|}, \tag{B1}$$

which is a key step in the derivation of the bounds. Let us introduce the following $K \times K$ block matrices:

$$\tilde{L} = \begin{bmatrix} L_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ L_K & 0 & \dots & 0 \end{bmatrix}, \ \tilde{Q} = \begin{bmatrix} Q_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ Q_K & 0 & \dots & 0 \end{bmatrix}, \tag{B2}$$

and $\tilde{A} = \mathbb{1}_{K\times K} \otimes A$. Note that

$$\tilde{L}^\dagger \tilde{A} \tilde{Q} = \begin{bmatrix} \sum_{k=1}^{K} L_k^\dagger A Q_k & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}. \tag{B3}$$

This implies that:

$$\left\|\sum_{k=1}^{K} L_k^\dagger A Q_k\right\| \;=\; \left\|\tilde{L}^\dagger \tilde{A} \tilde{Q}\right\| \;\leq\; \|\tilde{L}^\dagger\|\|\tilde{A}\|\|\tilde{Q}\|, \tag{B4}$$

where the inequality follows from the sub-multiplicavity property of the operator norm. Invoking now another operator norm property valid for arbitrary operators acting on a Hilbert space, $\|A\| = \sqrt{\|A^\dagger A\|}$ and $\|A^\dagger\| = \|A\|$, we get

$$\|\tilde{L}^\dagger\|\|\tilde{A}\|\|\tilde{Q}\| = \sqrt{\|\tilde{L}^\dagger \tilde{L}\|}\|\tilde{A}\|\sqrt{\|\tilde{Q}^\dagger \tilde{Q}\|}. \tag{B5}$$

Finally, noticing that $\|\tilde{L}^\dagger \tilde{L}\| = \|\sum_k L_k^\dagger L_k\|$, $\|\tilde{A}\| = \|A\|$ and $\|\tilde{Q}^\dagger \tilde{Q}\| = \|\sum_k Q_k^\dagger Q_k\|$ we arrive at Eq. (B1).

Appendix C: Derivation of the bound for causal superposition strategies

For CS strategies, any quantum superposition of $n!$ possible causal orders of n elementary channels can be created. This can be achieved by entangling all different causal orders with an external, $n!$-level quantum control system [3, 4]. As in AD strategies, we are allowed to put control unitaries between channels, moreover, for different causal orders, control unitaries may be different. Therefore, AD is a subclass of CS, in which only one causal order of channels is probed. The channel describing the action of n elementary channels Λ in the superposition of different causal orders is $\Lambda^{(n)}_{\mathrm{CS}}(\cdot) = \sum_{\boldsymbol{k^{(n)}}} K^{\mathrm{CS}}_{\boldsymbol{k^{(n)}}} \cdot K^{\mathrm{CS}\dagger}_{\boldsymbol{k^{(n)}}}$, where

$$K^{\mathrm{CS}}_{\boldsymbol{k^{(n)}}} = \sum_{\pi\in\sigma(n)} K^{\pi}_{\boldsymbol{k^{(n)}_\pi}} \otimes |\pi\rangle\langle\pi|, \tag{C1}$$

$\sigma(n)$ is the set of all n-element permutations, $\boldsymbol{k^{(n)}_\pi} = (k_{\pi(n)}, ..., k_{\pi(1)})$. The Kraus operators $K^{\pi}_{\boldsymbol{k^{(n)}}}$ are defined iteratively, as in the adaptive case: $K^{\pi}_{\boldsymbol{k^{(1)}}} = V^{\pi}_1(K_{k_1}\otimes\mathbb{1})$,

$$K^{\pi}_{\boldsymbol{k^{(i+1)}}} = V^{\pi}_{i+1}(K_{k_{i+1}} \otimes \mathbb{1})K^{\pi}_{\boldsymbol{k^{(i)}}}, \tag{C2}$$

V^{π}_i is the control unitary applied after i-th channel when the order of channels is described by permutation π, $|\pi\rangle$ is a corresponding state of a control system.

The CS family of strategies has been studied in the context of quantum metrology in Ref. [4], where numerical values of the QFI were obtained for $n = 2$ and $n = 3$ (the Authors use abbreviation "sup" instead of "CS"). Notice, that the formal definition of CS strategies from Appendix V.D of Ref. [4] does not contain the explicit form of the channel $\Lambda^{(n)}_{\mathrm{CS}}$. The explicit Kraus representation of a similar channel is present, for example, in [5] (for $n = 2$). The Choi-Jamiołkowski matrix of a channel involving superpositions of different causal orders for $n > 2$ is explicitly written down in [6].

Before we proceed further, let us show that we defined $\Lambda^{(n)}_{\mathrm{CS}}$ in a way that does not depend on the choice Kraus represenations of elementary channels Λ in (C2). Starting from an arbitrary Kraus representation $\{K_i\}$, we can construct any equivalent representation of the same channel $\{K'_j\}$ as $K'_j = \sum_i u_{ji}K_i$, where coefficients u_{ij} form a unitary matrix [7]. For a fixed π, the channel $\Lambda^{(n)}_{\pi}$, given by Kraus operators $K^{\pi}_{\boldsymbol{k^{(n)}}}$, is the concatenation of elementary channels Λ and unitary operations V^{π}_1, V^{π}_2,...,V^{π}_n— $\Lambda^{(n)}_{\pi}$ is well defined without referring to a particular Kraus representations of the elementary channel Λ. For this reason, changing the representation of each Λ may only change the Kraus operators $\{K^{\pi}_{\boldsymbol{k^{(n)}}}\}$ to an equivalent representation $\{K^{'\pi}_{\boldsymbol{l^{(n)}}}\}$, where $K^{'\pi}_{\boldsymbol{l^{(n)}}} = \sum_{\boldsymbol{k^{(n)}}} u^{\pi}_{\boldsymbol{l^{(n)}},\boldsymbol{k^{(n)}}} K^{\pi}_{\boldsymbol{k^{(n)}}}$, the coefficients $u^{\pi}_{\boldsymbol{l^{(n)}},\boldsymbol{k^{(n)}}}$ form a unitary matrix. Consequently, the Kraus opera-

tors $\{K^{\mathrm{CS}}_{\boldsymbol{k}^{(n)}}\}$ transform to $\{K'^{\mathrm{CS}}_{\boldsymbol{l}^{(n)}}\}$, where

$$K'^{\mathrm{CS}}_{\boldsymbol{l}^{(n)}} = \sum_{\pi\in\sigma(n)} \left(\sum_{\boldsymbol{k}^{(n)}} u^{\pi}_{\boldsymbol{l}^{(n)},\boldsymbol{k}^{(n)}} K^{\pi}_{\boldsymbol{k}^{(n)}} \right) \otimes |\pi\rangle\langle\pi| . \quad \text{(C3)}$$

This can be written as

$$K'^{\mathrm{CS}}_{\boldsymbol{l}^{(n)}} = \sum_{\boldsymbol{k}^{(n)}} u_{\boldsymbol{l}^{(n)},\boldsymbol{k}^{(n)}} K^{\mathrm{CS}}_{\boldsymbol{k}^{(n)}}, \quad \text{(C4)}$$

where

$$u_{\boldsymbol{l}^{(n)},\boldsymbol{k}^{(n)}} = \sum_{\pi} u^{\pi}_{\boldsymbol{l}^{(n)},\boldsymbol{k}^{(n)}} \otimes |\pi\rangle\langle\pi| \quad \text{(C5)}$$

is a unitary matrix. Therefore, the Kraus operators $K'^{\mathrm{CS}}_{\boldsymbol{l}^{(n)}}$ form an equivalent representation of a channel $\Lambda^{(n)}_{\mathrm{CS}}$, so physical properties of this channel do not depend on the choice of represantions of Λ.

In what follows, we derive an upper bound for the QFI achievable with any CS strategy involving n elementary channels, which can be efficiently computed for large n. We know that

$$F^{(n)}_{\mathrm{CS}} \leq \min_{\left\{K^{\mathrm{CS}}_{\boldsymbol{k}^{(n)}}\right\}} \left\|\alpha^{(n)}_{\mathrm{CS}}\right\|, \quad \alpha^{(n)}_{\mathrm{CS}} = \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\mathrm{CS}\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}^{\mathrm{CS}}_{\boldsymbol{k}^{(n)}}, \quad \text{(C6)}$$

where the minimization is taken over all Kraus representations of the channel $\Lambda^{(n)}_{\mathrm{CS}}$. Such minimization is very hard to perform, but we can obtain a less tight but more feasible upper bound by minimizing over representations of a single channel only, assuming that the representation of each elementary channel Λ is the same:

$$F^{(n)}_{\mathrm{CS}} \leq \min_{\{K_k\}} \left\|\alpha^{(n)}_{\mathrm{CS}}\right\|. \quad \text{(C7)}$$

For this family of representations, we have

$$\alpha^{(n)}_{\mathrm{CS}} = \sum_{\boldsymbol{k}^{(n)}} \sum_{\pi\in\sigma(n)} \dot{K}^{\pi\dagger}_{\boldsymbol{k}^{(n)}_{\pi}} \dot{K}^{\pi}_{\boldsymbol{k}^{(n)}_{\pi}} \otimes |\pi\rangle\langle\pi| . \quad \text{(C8)}$$

The matrix $\alpha^{(n)}_{\mathrm{CS}}$ has a block diagonal structure, where different blocks correspond to different π because the states $|\pi\rangle$ are orthogonal to each other. Therefore,

$$\left\|\alpha^{(n)}_{\mathrm{CS}}\right\| = \max_{\pi\in\sigma(n)} \left\| \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\pi\dagger}_{\boldsymbol{k}^{(n)}_{\pi}} \dot{K}^{\pi}_{\boldsymbol{k}^{(n)}_{\pi}} \right\|. \quad \text{(C9)}$$

The summation inside a norm runs over all possible vectors $\boldsymbol{k}^{(n)}$, so it does not change when lower indices $\boldsymbol{k}^{(n)}_{\pi}$ are replaced with $\boldsymbol{k}^{(n)}$, so we have

$$\left\|\alpha^{(n)}_{\mathrm{CS}}\right\| = \max_{\pi\in\sigma(n)} \left\| \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\pi\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}^{\pi}_{\boldsymbol{k}^{(n)}} \right\| \quad \text{(C10)}$$

In this step, the assumption that the Kraus representations of all n elementary channels are the same, was crucial—otherwise, a different π would correspond to a different order of representations, so (C10) would not follow from (C9). That is why we cannot minimize over different representations for different channels, as we did for AD strategies, and the CS bound is less tight.

For a fixed π, the Kraus operators $K^{\pi}_{\boldsymbol{k}^{(n)}}$ describe an adaptive strategy on n elementary channels with control unitaries V^{π}_i. In (C10), maximization over π is equivalent to maximization over different sets of control unitaries V^{π}_i with a fixed causal order of channels, so we have

$$\left\|\alpha^{(n)}_{\mathrm{CS}}\right\| \leq \max_{\{V_i\}} \left\| \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}_{\boldsymbol{k}^{(n)}} \right\|, \quad \text{(C11)}$$

where the maximization is over all sets of control unitaries, and $K_{\boldsymbol{k}^{(n)}}$ for a given $\{V_i\}$ are defined in (4). The upper bound for the r.h.s. for a fixed elementary channel representation $\{K_k\}$ is

$$\max_{\{V_i\}} \left\| \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}_{\boldsymbol{k}^{(n)}} \right\| \leq a^{(n)}, \quad \text{(C12)}$$

which follows directly from the derivation of the AD bound. Finally, we can minimize both sides of (C11) and (C12) over $\{K_k\}$ to obtain

$$\min_{\{K_k\}} \left\|\alpha^{(n)}_{\mathrm{CS}}\right\| \leq \min_{\{K_k\}} \max_{\{V_i\}} \left\| \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}_{\boldsymbol{k}^{(n)}} \right\| \leq \min_{\{K_k\}} a^{(n)}. \quad \text{(C13)}$$

This inequality, together with (C7), leads directly to the upper bound for $F^{(n)}_{\mathrm{CS}}$ from (10).

The considered class of strategies CS is quite general, since it allows superpositions of all different causal orders of channels combined with all possible intermediate controls—moreover, for different causal orders, unitary controls may be different. However, quantum mechanics in principle allows for even more general non-causal strategies, the most general class of strategies is schematically shown in Fig. 1.e in Ref. [4], where the Authors refer to it using abbreviation "ICO". The advantage of ICO over CS strategies is usually very small, but possible to demonstrate numerically already for $n = 2$ [4]. Therefore, it is interesting to ask, whether our newly derived bound can be applied to the most general non-causal strategies ICO. Unfortunately, the answer is negative, and a counterexample exists already for $n = 3$. Let us consider a single qubit channel studied in the *introductory example*, where noise is described by the Kraus operators from (1) with $p = 0.75$. For $n = 3$, exact values of the maximal QFI associated with CS and ICO strategies (obtained using procedures from [4]) are $F^{(3)}_{\mathrm{CS}} = 5.52$, $F^{(3)}_{\mathrm{ICO}} = 5.84$, whereas the upper bound for QFI for CS strategies is $\min_{\{K_k\}} 4a^{(3)} = 5.73$, so we have $F^{(3)}_{\mathrm{CS}} < \min_{\{K_k\}} 4a^{(3)} < F^{(3)}_{\mathrm{ICO}}$. Therefore, the problem of finding a feasible upper bound on the QFI for the most

general non-causal strategies is still open. However, it is not clear whether all ICO strategies are possible to implement—probably this fundamental question should be addressed first.

Appendix D: Computing the iterative bounds

Firstly, let us describe more precisely the construction of the AD bound. When we construct $a^{(n)}$, defined in (9), we can choose the Kraus representation of the elementary channel $\{K_k\}$ independently in each step, so the norms $\|\alpha\|$, $\|\beta\|$ are different in each step. Then, we want to minimize $a^{(n)}$ over all possible sets of n represenations $\{K_k\}^{\times n}$, the i-th element of such a set corresponds to the representation chosen in i-th step. However, $a^{(i+1)}$ depends only on $a^{(i)}$ and the representation chosen in the step $i+1$, moreover, $a^{(i+1)}$ is an increasing function of $a^{(i)}$. Therefore, the minimum of $a^{(i+1)}$ over $\{K_k\}^{\times(i+1)}$ can be obtained by inserting the minimum of $a^{(i)}$ over $\{K_k\}^{\times i}$ into (9), and then minimizing the resulting expression over the representation of $(i+1)$-th channel.

As a result, the full optimization over n Kraus representations used in n steps may be easily obtained by performing n iterative optimizations over a single Kraus representation in each successive step, namely:

$$\begin{aligned} &\min_{\{K_k\}^{\times n}} a^{(n)} = \tilde{a}^{(n)}, \quad \text{where} \quad \tilde{a}^{(0)} = 0, \\ &\tilde{a}^{(i+1)} = \min_{\{K_k\}} \left[\tilde{a}^{(i)} + \left\|\alpha_{\{K_k\}}\right\| + 2\left\|\beta_{\{K_k\}}\right\| \sqrt{\tilde{a}^{(i)}}\right]. \end{aligned} \tag{D1}$$

We used notation $\alpha_{\{K_k\}}$ and $\beta_{\{K_k\}}$ to indicate the dependence of an elementary channel representation $\{K_k\}$, over which we minimize in step $(i+1)$. We are now ready to demonstrate the numerical procedures for computing AD and CS bounds.

1. Minimizing $\|\alpha\|$ assuming $\|\beta\| \leq b$

Let us show how to formulate the following minimization problem

$$g(b) = \quad \text{minimize}_{\{K_k\}} \quad \left\|\alpha_{\{K_k\}}\right\| \tag{D2}$$

$$\text{subject to} \quad \left\|\beta_{\{K_k\}}\right\| \leq b \tag{D3}$$

as an SDP. Firstly, as shown in [7, 8], we can significantly limit the class of Kraus representations over which we minimize. Starting from arbitrary $\{K_k\}$, it is enough to consider representations $\{\tilde{K}_k\}$ satisfying

$$\tilde{K}_k = K_k, \quad \dot{\tilde{K}}_k = \dot{\tilde{K}}_k - i\sum_j h_{kj} K_j \tag{D4}$$

to obtain all possible values of $(\|\alpha\|, \|\beta\|)$ for a given channel Λ, the coefficients h_{kj} must form a hermitian matrix $h \in \mathrm{Herm}(\mathbb{C}^{r\times r})$. We can replace the minimization over $\{K_k\}$ with a minimization over h in (D2). Following the path from [8, 9], let us construct the following matrices

$$A = \left(\begin{array}{c|ccc} \lambda \mathbb{1}_d & \dot{\tilde{K}}_1^\dagger & \dots & \dot{\tilde{K}}_r^\dagger \\ \hline \dot{\tilde{K}}_1 & & & \\ \vdots & & \mathbb{1}_{d\cdot r} & \\ \dot{\tilde{K}}_r & & & \end{array}\right), \tag{D5}$$

$$B = \left(\begin{array}{c|c} b\mathbb{1}_d & i\sum_{k=1}^r \dot{\tilde{K}}_k^\dagger K_k \\ \hline \left(i\sum_{k=1}^r \dot{\tilde{K}}_k^\dagger K_k\right)^\dagger & b\mathbb{1}_d. \end{array}\right) \tag{D6}$$

It can be shown, using Schur's complement condition, that

$$A \succeq 0 \iff \lambda \geq \|\alpha_{\{\tilde{K}_k\}}\|, \quad B \succeq 0 \iff b \geq \|\beta_{\{\tilde{K}_k\}}\|. \tag{D7}$$

Therefore, the constraint $B \succeq 0$ is equivalent to upper bounding $\|\beta\|$ by b, and minimization of λ is equivalent to minimization of $\|\alpha\|$. All elements of A and B are linear in λ and h, so we can use constraints for positivity of A and B in the SDP. All this observations allow us to write down the minimization problem (D2) as an SDP described in Algorithm 1.

Algorithm 1 Minimize $\|\alpha\|$ given $\|\beta\| \leq b$

Input: $K_k, \dot{K}_k \in \mathbb{C}^{d\times d}$ for $k \in \{1, ..., r\}$ ▷ `Kraus operators of elementary channel Λ and their derivatives`
Input: $b \geq 0$ ▷ `Upper bound for ‖β‖`
Output: $g(b) = \min_{\{K_k\}} \|\alpha\|$ s.t. $\|\beta\| \leq b$

1: variables: $\lambda \geq 0$, $h \in \mathrm{Herm}(\mathbb{C}^{r\times r})$ ▷ `Variables to optimize in SDP`
2: **for** k in $(1, ..., r)$ **do**
3: $\quad \dot{\tilde{K}}_k := \dot{K}_k - i\sum_{j=1}^r h_{kj} K_j$
4: **end for**
5: construct matrices A, B defined in (D5), (D6)
6: minimize$_{\lambda,h}$ λ
subject to $A \succeq 0$, $B \succeq 0$ ▷ `SDP`
7: **Output** λ

2. Computing AD and CS bounds

Let us prove the following lemma:

Lemma 1 Let $f(x, y)$ be an increasing function of x and y with non-negative values. Then,

$$\min_{\{K_k\}} f\left(\|\alpha_{\{K_k\}}\|, \|\beta_{\{K_k\}}\|\right) = \min_{b\in[l,r]} f\left(g(b), b\right), \tag{D8}$$

where $l = \min_{\{K_k\}} \|\beta_{\{K_k\}}\|$, $r = \min_{\{K_k\}} \sqrt{\|\alpha_{\{K_k\}}\|}$, $g(b)$ is defined in (D2).

Proof First note that

$$\min_{\{K_k\}} f\left(\|\alpha_{\{K_k\}}\|, \|\beta_{\{K_k\}}\|\right) = \min_{b\geq l} \min_{\{K_k\}:\|\beta_{\{K_k\}}\|\leq b} f\left(\|\alpha_{\{K_k\}}\|, b\right) = \min_{b\geq l} f\left(g(b), b\right), \tag{D9}$$

where we used the fact that $f(x,y)$ is increasing with y (1st equality) and x (2nd equality). What remains to be proven is that the minimal value is always obtained for $b \leq r$. From (7), we have $\|\beta_{\{K_k\}}\| \leq \sqrt{\|\alpha_{K_k}\|}$, which means that $l \leq r$. Let $\{K'_k\}$ be the Kraus representation minimizing $\|\alpha\|$, which means that $\sqrt{a'} = \sqrt{\|\alpha_{\{K'_k\}}\|} = r$, $b' = \|\beta_{\{K'_k\}}\| \leq \sqrt{\|\alpha_{\{K'_k\}}\|} = r$. Obviously, $a' \leq a$, so when we choose $\{K_k\}$ such that $b > r$, we have $a' \leq a$ and $b' = r < b$, so $f(a',b') < f(a,b)$, which means that the minimum cannot be achieved for $b > r$.■

The values of l and r can be found using an SDP, as descibed in [8, 9]—the result of a minimization of λ with the constraint $A \succeq 0$ is r^2, and the minimum of b with the constraint $B \succeq 0$ equals to l.

Notice, that $\tilde{a}^{(i+1)}$ defined in (D1) is an increasing function of $\|\alpha\|$ and $\|\beta\|$, so we can use Lemma 1 to perform the minimization over $\{K_k\}$:

$$\tilde{a}^{(i+1)} = \min_{b\in[l,r]} \left[\tilde{a}^{(i)} + g(b) + 2b\sqrt{\tilde{a}^{(i)}}\right] \tag{D10}$$

In practice, we do not have an analytical form of $g(b)$, and we need to solve an SDP problem for each argument b independently to obtain $g(b)$. Therefore, minimization over $b \in [l,r]$ is approximated with minimization over $b \in \text{Linspace}(l,r,p)$, where $\text{Linspace}(l,r,p)$ is a p-element arithmetic sequence whose first element is l and last element is r. Notice, that this approximation can only increase the final result—therefore, even for small sampling precision p, we obtain valid, but not necessary tight, upper bounds. We used the value $p = 500$ to generate the data presented in this paper, further precision increase did not lead to any significant difference.

To sum up, the procedure to compute the upper bound for $F_{\text{AD}}^{(n)}$ is as follows. Firstly, we compute l and r using the SDP. Secondly, $g(b)$ is calculated for $b \in \text{Linspace}(l,r,p)$. Finally, we initialize $\tilde{a}^{(0)} = 0$, and compute $\tilde{a}^{(i)}$ for larger i iteratively using (D1). Notice, that the computational cost of each iteration is very low since we computed all the required values of $g(b)$ at the beginning, and we keep them in the memory. The computation of $g(b)$ for p different values of b is the most time-consuming part, but even for a high precision ($p = 500$), the computation time did not exceed two minutes. Once we do this computation, we can easily get an upper bound for $F_{\text{AD}}^{(n)}$ even for large values of n.

The procedure to compute an upper bound for $F_{\text{CS}}^{(n)}$ is very similar. Notice, that $a^{(i)}$ defined in (9) is an increasing function of $\|\alpha\|$ and $\|\beta\|$, so we can use Lemma 1 to find its minimum over a single $\{K_k\}$:

$$\min_{\{K_k\}} a^{(n)} = \min_{b\in[l,r]} \hat{a}^{(n)}(b) \quad \text{, where} \quad \hat{a}^{(0)}(b) = 0, \quad \hat{a}^{(i+1)}(b) = \hat{a}^{(i)}(b) + g(b) + 2b\sqrt{\hat{a}^{(i)}(b)} \tag{D11}$$

Subsequent values of $\hat{a}^{(i)}(b)$ are found iteratively for all $b \in \text{Linspace}(l,r,p)$, and then minimization over b is performed with the help of previously calculated values of $g(b)$. The whole procedure of computing bounds for $F_{\text{AD}}^{(n)}$ and $F_{\text{CS}}^{(n)}$ is summarized in Algorithm 2.

Algorithm 2 Iterative bounds for $F_{\text{AD}}^{(n)}$ and $F_{\text{CS}}^{(n)}$

Input: $K_k, \dot{K}_k \in \mathbb{C}^{d\times d}$ for $k \in \{1,...,r\}$
Input: $n \in \mathbb{Z}_+$ ▷ `Max number of elementary channels`
Input: $p \in \mathbb{Z}_+$ ▷ `Precision of sampling`
Output: AD, CS: lists of size n, AD/CS[k] is an upper bound for $F_{\text{AD/CS}}^{(k)}$

$l := \min_{\{K_k\}} \|\beta\|$
$r := \min_{\{K_k\}} \sqrt{\|\alpha\|}$
variables: a[p], b[p], t1[p], t2[p], AD[n], CS[n] ▷ `Lists of real numbers of size` $[\cdot]$`, initialized with 0s`
b := Linspace (l,r,p) ▷ b[1] $= l$, b[p] $= r$, b[1], b[2], ..., b[p] `is arithmetic sequence`
for j in $(1,...,p)$ **do**
    a[j] := g(b[j]) ▷ `Use Alg. 1`
end for
▷ `Computing AD bounds`
$m := 0$
for k in $(1,...,n)$ **do**
    for j in $(1,...,p)$ **do**
        t1[j] := m + a[j] + 2 * b[j] * $\sqrt{m}$
    end for
    m := Minimum[t1]
    AD[k] := 4 * m
end for
▷ `Computing CS bounds`
for k in $(1,...,n)$ **do**
    for j in $(1,...,p)$ **do**
        t2[j] := t2[j] + a[j] + 2 * b[j] * $\sqrt{\text{t2}[j]}$
    end for
    CS[k] := 4 * Minimum[t2]
end for
Output AD, CS

The old bound for AD strategies can be also computed numerically using Lemma 1—the minimization from (3) simplifies to:

$$\min_{\{K_i\}} 4\left[n\|\alpha\| + n(n-1)\|\beta\|\left(\|\beta\| + 2\sqrt{\|\alpha\|}\right)\right] = \min_{b\in[l,r]} 4\left[ng(b) + n(n-1)b\left(b + 2\sqrt{g(b)}\right)\right]. \tag{D12}$$

Analogously, we can compute the new analytical bound from (12), or the bound for E strategies (2). Therefore, the function $g(b)$ fully characterizes an elementary channel Λ_φ from a metrological point of view—if we know $g(b)$ (at least for some values of b), we can easily compute all metrological bounds.

3. Tightening of the bound for parallel strategies

Looking at the original derivation of the bound for parallel strategies in [7], we see that it may be tightened by allowing different Kraus representations for each elementary Hilbert space in Eq. (16) of [7] . However, the full optimization here is a bit more complicated to perform than in the adaptive case, while the advantage over the standard parallel bound (2) is not significant for the typical examples (what we checked numerically). Still, it is worth noting that such an optimized bound is guaranteed to be tighter than the newly derived adaptive one (10), which we show below.

The channel describing the parallel action of n elementary channels Λ is $\Lambda_{\mathrm{E}}^{(n)}(\cdot) = \sum_{\boldsymbol{k}^{(n)}} K^{\mathrm{E}}_{\boldsymbol{k}^{(n)}} \cdot K^{\mathrm{E}\dagger}_{\boldsymbol{k}^{(n)}}$, where $K^{\mathrm{E}}_{\boldsymbol{k}^{(1)}} = K_{k_1} \otimes \mathbb{1}$,

$$K^{\mathrm{E}}_{\boldsymbol{k}^{(i+1)}} = K_{k_{i+1}} \otimes K^{\mathrm{E}}_{\boldsymbol{k}^{(i)}}, \tag{D13}$$

$\mathbb{1}$ is acting on an ancillary system. Using this definition, we obtain the following iteration relations for $\alpha_{\mathrm{E}}^{(n)} = \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\mathrm{E}\dagger}_{\boldsymbol{k}^{(n)}} \dot{K}^{\mathrm{E}}_{\boldsymbol{k}^{(n)}}$, $\beta_{\mathrm{E}}^{(n)} = \sum_{\boldsymbol{k}^{(n)}} \dot{K}^{\mathrm{E}\dagger}_{\boldsymbol{k}^{(n)}} K^{\mathrm{E}}_{\boldsymbol{k}^{(n)}}$:

$$\alpha_{\mathrm{E}}^{(i+1)} = \alpha \otimes \mathbb{1} + \mathbb{1} \otimes \alpha_{\mathrm{E}}^{(i)} - 2\beta \otimes \beta_{\mathrm{E}}^{(i)}, \tag{D14}$$

$$\beta_{\mathrm{E}}^{(i+1)} = \beta \otimes \mathbb{1} + \mathbb{1} \otimes \beta_{\mathrm{E}}^{(i)}, \tag{D15}$$

$\mathbb{1}$ acting on an ancillary system was omitted. From the triangle inequality:

$$\|\alpha_{\mathrm{E}}^{(i+1)}\| \leq \|\alpha_{\mathrm{E}}^{(i)}\| + \|\alpha\| + 2\|\beta\|\|\beta_{\mathrm{E}}^{(i)}\|, \tag{D16}$$

$$\|\beta_{\mathrm{E}}^{(i+1)}\| \leq \|\beta_{\mathrm{E}}^{(i)}\| + \|\beta\|. \tag{D17}$$

Let us define

$$a_{\mathrm{E}}^{(i+1)} = a_{\mathrm{E}}^{(i)} + \|\alpha\| + 2\|\beta\| b_{\mathrm{E}}^{(i)}, \tag{D18}$$

$$b_{\mathrm{E}}^{(i+1)} = b_{\mathrm{E}}^{(i)} + \|\beta\|, \quad a_{\mathrm{E}}^{(0)} = 0, \quad b_{\mathrm{E}}^{(0)} = 0. \tag{D19}$$

Using the inequality $\|\alpha_{\mathrm{E}}^{(n)}\| \leq a_{\mathrm{E}}^{(n)}$, we obtain the upper bound for the QFI associated with E strategies

$$F_{\mathrm{E}}^{(n)} \leq \min_{\{K_k\}} 4a_{\mathrm{E}}^{(n)}, \tag{D20}$$

which is equivalent to the state-of-the-art bound presented in (2). Again, to make this bound tighter, we can allow for different Kraus representations of an elementary channel in each step of the iteration:

$$F_{\mathrm{E}}^{(n)} \leq \min_{\{K_k\}^{\times n}} 4a_{\mathrm{E}}^{(n)}. \tag{D21}$$

From the inequality $\|\beta\|^2 \leq \|\alpha\|$ and from the induction principle follows $b_{\mathrm{E}}^{(i)} \leq \sqrt{a_{\mathrm{E}}^{(i)}}$. Therefore, if we compare (9) with (D18), we see that $a_{\mathrm{E}}^{(n)} \leq a^{(n)}$, so the iterative bound for E strategies is tighter than the one for AD strategies. However, the minimization from (D21) cannot be performed step by step as easily as in the adaptive case because the minimimal value of $a_{\mathrm{E}}^{(i+1)}$ is not always achieved for the minimal value of $a_{\mathrm{E}}^{(i)}$.

Appendix E: Asymptotic equivalence between adaptive and parallel strategies

For any fixed single-channel Kraus representation $\{K_k\}$ (not necessarily the optimal one), we have:

$$a^{(i+1)} = a^{(i)} + \|\alpha\| + 2\|\beta\|\sqrt{a^{(i)}}, \quad a^{(0)} = 0. \tag{E1}$$

We will show that $\forall_{n\geq 1} a^{(n)} \leq f(n)$, where:

$$f(n) = n\|\alpha\| + n(n-1)\|\beta\|^2 + n\log n(\|\alpha\| - \|\beta\|^2). \tag{E2}$$

First note that they are equal at $n = 1$:

$$f(1) = a^{(1)} = \|\alpha\|. \tag{E3}$$

Next, we will show that this function satisfies the iteration rule (E1) with inequality $\geq$ instead of $=$,

$$f(n+1) - f(n) \overset{?}{\geq} \|\alpha\| + 2\|\beta\|\sqrt{f(n)} \tag{E4}$$

which would lead to $f(n) \geq a^{(n)}$.

To prove (E4), let us write down the l.h.s. explicitly:

$$f(n+1) - f(n) = \|\alpha\| + 2\|\beta\|^2 n + (\|\alpha\| - \|\beta\|^2)((1+n)\log(1+n) - n\log(n)). \tag{E5}$$

Note that first derivative of $x\log(x)$ is a strictly increasing function, and therefore:

$$(1+n)\log(1+n) - n\log(n) = \int_n^{n+1} (x\log(x))' dx \geq (x\log(x))'\Big|_{x=n} = 1 + \log(n). \tag{E6}$$

Using (E6), we can deduce that (E4) follows from

$$\|\alpha\| + 2\|\beta\|^2 n + (\|\alpha\| - \|\beta\|^2)(1 + \log(n)) \overset{?}{\geq} \|\alpha\| + 2\|\beta\|\sqrt{\|\beta\|^2 n(n-1) + n\|\alpha\| + (\|\alpha\| - \|\beta\|^2)n\log(n)}, \tag{E7}$$

which, after subtracting $\|\alpha\|$, taking the square and subtracting the r.h.s. reduces to

$$(\|\alpha\| - \|\beta\|^2)^2(1 + \log(n))^2 \geq 0, \tag{E8}$$

which is always true.

Now, as $f(n) \geq a^{(n)}$ for any fixed Kraus representation, $4f(n)$ minimized over all Kraus representations constitutes a valid bound for the QFI (12).

Appendix F: Continuous time limit for Markovian dynamics

We consider a Markovian semigroup dynamics in continuous time, described by the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) master equation

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_{j=1}^{J} L_j \rho L_j^\dagger - \frac{1}{2}\rho L_j^\dagger L_j - \frac{1}{2} L_j^\dagger L_j \rho. \tag{F1}$$

A Kraus representation of the dynamical map at the lowest order in Δt is

$$K_0 = \mathbb{1} - \left(\mathrm{i}H + \frac{1}{2}\mathbf{L}^\dagger\mathbf{L}\right)\Delta t + O\left(\Delta t^2\right) \tag{F2}$$

$$\mathbf{K} = \mathbf{L}\sqrt{\Delta t} + O\left(\Delta t^{\frac{3}{2}}\right), \tag{F3}$$

where we introduced a vector notation $\mathbf{L} = [L_1, \dots L_J]^T$ for the J collapse operators, which means $\mathbf{L}^\dagger\mathbf{L} = \sum_{j=1}^{J} L_j^\dagger L_j$, and for the J Kraus operators $\mathbf{K} = [K_1, \dots, K_J]^T$ (the 0th operator is kept separate). Following Refs. [10, 11], we obtain the continuous time evolution as the limit $\Delta t \to 0$ of repeated sequential applications of the channel with the above Kraus representation.

Since $F^{(i)} = 4\|\alpha^{(i)}\|$ when the input state involves ancillary systems [7], we can rewrite Eq. (8) as an inequality for the difference quotient of the (ancilla-assisted) QFI

$$\frac{F^{(i+1)} - F^{(i)}}{\Delta t} \leq \frac{4}{\Delta t}\left(\|\alpha\| + \|\beta\|\sqrt{F^{(i)}}\right), \tag{F4}$$

where α and β depend on Δt through the Kraus operators in Eqs. (F2,F3).

The parameter φ to be estimated is arbitrary and can appear both in the Hamiltonian H and in the collapse operators L_j. The derivatives of the Kraus operators are

$$\dot{K}_0 = -\left(\mathrm{i}\dot{H} + \frac{1}{2}\dot{\mathbf{L}}^\dagger\mathbf{L} + \frac{1}{2}\mathbf{L}^\dagger\dot{\mathbf{L}}\right)\Delta t \tag{F5}$$

$$\dot{\mathbf{K}} = \dot{\mathbf{L}}\sqrt{\Delta t}. \tag{F6}$$

The different Δt-order of the 0th Kraus operator suggests to divide the matrix h in Eq. (D4) that specifies the Kraus representations in blocks as

$$h = \begin{bmatrix} h_{00} & \mathbf{h}^\dagger \\ \mathbf{h} & \mathfrak{h} \end{bmatrix}, \tag{F7}$$

with $h_{00} \in \mathbb{R}$ and $\mathfrak{h}^\dagger = \mathfrak{h}$. We expand this matrix in powers of Δt as follows

$$h = h^{(0)} + h^{(\frac{1}{2})}\sqrt{\Delta t} + h^{(1)}\Delta t, \tag{F8}$$

and it must be chosen such that the limit for $\Delta t \to 0$ remains finite and Eq. (F4) gives a meaningful bound on the time-derivative of the QFI. From Eq. (F4) it is evident that we need to choose h such that the terms of order less than Δt in α and β are zero, while the terms of higher order become irrelevant in the limit $\Delta t \to 0$.

After some algebra one realizes that we must impose $h_{00}^{(0)} = h_{00}^{(\frac{1}{2})} = 0$ and $\mathbf{h}^{(0)} = 0$ to have $\beta^{(0)} = \beta^{(\frac{1}{2})} = \alpha^{(0)} = \alpha^{(\frac{1}{2})} = 0$; the first-order terms that remain relevant in the limit are

$$-\mathrm{i}\beta^{(1)} = \dot{H} - \frac{\mathrm{i}}{2}\left(\dot{\mathbf{L}}^\dagger\mathbf{L} - \mathbf{L}^\dagger\dot{\mathbf{L}}\right) + h_{00}^{(1)}\mathbb{1} + \mathbf{L}^\dagger\mathbf{h}^{(\frac{1}{2})} + \mathbf{h}^{(\frac{1}{2})\dagger}\mathbf{L} + \mathbf{L}^\dagger\mathfrak{H}^{(0)}\mathbf{L} \tag{F9}$$

and

$$\alpha^{(1)} = \left[\mathbf{h}^{(\frac{1}{2})}\mathbb{1} + \mathfrak{h}^{(0)}\mathbf{L} + \mathrm{i}\dot{\mathbf{L}}\right]^\dagger\left[\mathbf{h}^{(\frac{1}{2})}\mathbb{1} + \mathfrak{h}^{(0)}\mathbf{L} + \mathrm{i}\dot{\mathbf{L}}\right]. \tag{F10}$$

The bound on the rate of increase of the QFI is

$$\frac{dF(t)}{dt} \leq 4 \min_{h_{00}^{(1)}, \vec{h}^{(\frac{1}{2})}, \mathfrak{h}^{(0)}} \left(\left\|\alpha^{(1)}\right\| + \left\|\beta^{(1)}\right\|\sqrt{F(t)}\right). \tag{F11}$$

Since Eq. (8) can also be derived when a different channel is applied at each step, the previous inequality also holds in the Markovian time-inhomogenous case with time-dependent H and L_j in Eq. (F1).

Recently, the authors of Ref. [12] derived a bound tighter than the one in Refs. [11, 13], without relying on a discretization of the time evolution. For a Hamiltonian parameter (i.e. $\dot{\mathbf{L}} = 0$), Eq. (18) of Ref. [12] gives a state-dependent bound for arbitrary initial states. The inequalities in Eq. (19) of Ref. [12] can further be applied to obtain a state-independent bound for finite-dimensional systems which coincides with the bound in Eq. (F11) by applying the substitutions $-\mathrm{i}\beta^{(1)} \leftrightarrow G$, $\alpha^{(1)} \leftrightarrow \sum_j A_j^\dagger A_j$, $h_{00}^{(1)} \leftrightarrow -x$, $h_j^{(\frac{1}{2})} \leftrightarrow -\beta_j$ and $\mathfrak{h}_{ij}^{(0)} \leftrightarrow -\gamma_{ij}$.

For a Hamiltonian parameter, the two asymptotic cases are

1. The Hamiltonian derivative $\dot{H}$ is in the Lindblad span and we can find h such that $\beta^{(1)} = 0$ leading to

$$F \leq 4t \min_{h_{00}^{(1)}, \vec{h}^{(\frac{1}{2})}, \mathfrak{h}^{(0)}, \beta^{(1)}=0} \left\|\alpha^{(1)}\right\| \tag{F12}$$

2. The Hamiltonian derivative $\dot{H}$ is not in the Lindblad span, then the asymptotic solution of the differential equation for large t gives

$$F \leq 4t^2 \min_{h_{00}^{(1)}, \vec{h}^{(\frac{1}{2})}, \mathfrak{h}^{(0)}} \left\|\beta^{(1)}\right\|^2. \tag{F13}$$

Both bounds are asymptotically attainable with the parallel strategies described in Refs. [13, 14]. However, Eq. (F11) can give tighter bounds for short times [12]. Finally, while the estimation of a parameter appearing in the collapse operators is less studied, the bounds of Refs. [11, 13] have recently been applied to study optimal quantum thermometry in Markovian environments [15], thus these tighter bounds may also be useful in that context.

[1] C. W. Helstrom, *Quantum detection and estimation theory* (Academic press, 1976).

[2] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. **72**, 3439 (1994).

[3] M. Araújo, F. Costa, and i. c. v. Brukner, Computational advantage from quantum-controlled ordering of gates, Phys. Rev. Lett. **113**, 250402 (2014).

[4] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal strategies of quantum metrology with a strict hierarchy, Phys. Rev. Lett. **130**, 070803 (2023).

[5] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, Phys. Rev. A **103**, 032615 (2021).

[6] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, Quantum circuits with classical versus quantum control of causal order, PRX Quantum **2**, 030335 (2021).

[7] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, J. Phys. A **41**, 255304 (2008).

[8] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, Nat. Commun. **3**, 1063 (2012).

[9] J. Kolodynski and R. Demkowicz-Dobrzanski, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).

[10] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, Quantum **1**, 27 (2017).

[11] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive Quantum Metrology under General Markovian Noise, Phys. Rev. X **7**, 041009 (2017).

[12] K. Wan and R. Lasenby, Bounds on adaptive quantum metrology under Markovian noise, Phys. Rev. Research **4**, 033092 (2022).

[13] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).

[14] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, Phys. Rev. Res. **2**, 013235 (2020).

[15] P. Sekatski and M. Perarnau-Llobet, Optimal nonequilibrium thermometry in Markovian environments, Quantum **6**, 869 (2022).

Chapter 4

Universal bounds for quantum metrology in the presence of correlated noise

Commentary

In this work, universal upper bounds are derived for the most general metrological schemes, incorporating both correlations and quantum adaptiveness (C3, Fig. 1.1). As in Chapter 3, we bound the maximal QFI increment caused by increasing the number of estimated channels, which leads to recursive formula for the QFI upper bound.

Here, we propose a slightly different construction of a single recursion step, more in the spirit of the MOP technique. We observe that purifying the joint state of the system, environment, and ancilla at any stage of the protocol can only increase the resulting QFI. Moreover, using methods from Refs. [Alt21; Liu23], we derive an SDP for the QFI achievable with access to such a purification and m additional copies of Λ_φ. This allows us to upper bound the QFI with $N+m$ channels based on the QFI with N channels, which constitutes a single iteration step.

Interestingly, for uncorrelated cases (C2), this procedure is at least as good as that presented in the previous chapter and sometimes yields even tighter bounds for finite N. More importantly, the approach can be generalized to correlated cases.

However, correlations require additional care. To construct a tractable iterative procedure, we must cut the chain of many correlated channels Λ_φ (see C3, Fig. 1.1) into smaller segments. Crucially, one cannot assume that environmental information is lost between these segments, as this would underestimate the QFI.

Consequently, we make a different simplifying assumption: after every segment of m connected channels, the environmental information becomes fully accessible to the control operation C_i. This assumption may *increase* the resulting QFI but can never decrease it—thus guaranteeing a valid upper bound.

To make the bounds as tight as possible, one should minimize these information

leakages by increasing m. However, this significantly increases the computational complexity.

Nevertheless, as we demonstrate through examples of correlated dephasing and collisional thermometry, even small values of m yield informative bounds that nontrivially capture the impact of correlations. We benchmark the bounds' performance using the tensor-network based methods developed in Chapter 2.

We also construct SDPs that determine the asymptotic behavior of our recursive bounds and introduce correlated analogs of the α and β operators used for uncorrelated cases. When all probed channels are identical, two scaling regimes emerge: SS or HS, as in uncorrelated models.

Our bound is derived for scheme C3 but applies to all nine scheme classes depicted in Fig. 1.1, which are subclasses of C3. In particular, it provides the first informative universal bounds for classes A3 and B3. Whether tighter bounds for B3 or A3 can be derived by restricting the set of estimation strategies remains an open question.

My contribution

I conceived—through extensive discussions with RDD and FA—and developed all the main ideas presented in this work, performed all calculations and proofs. I developed and implemented numerical procedures used to calculate the bounds, I generated and analyzed all the numerical data. I wrote the whole main text and all appendices, I am the corresponding author.

Universal Bounds for Quantum Metrology in the Presence of Correlated Noise

Stanisław Kurdziałek,[1] Francesco Albarelli,[2] and Rafał Demkowicz-Dobrzański[1]
[1]Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland
[2]Scuola Normale Superiore, I-56126 Pisa, Italy

(Received 12 November 2024; accepted 31 July 2025; published 22 September 2025)

We derive fundamental bounds for general quantum metrological models involving both temporal or spatial correlations (mathematically described by quantum combs), which may be effectively computed in the limit of a large number of probes or sensing channels involved. Although the bounds are not guaranteed to be tight in general, their tightness may be systematically increased by increasing the numerical complexity of the procedure. Interestingly, this approach yields bounds tighter than the state of the art also for uncorrelated channels. We apply the bounds to study the limits for the most general adaptive phase estimation models in the presence of temporally correlated dephasing. We consider dephasing both parallel (no Heisenberg scaling) and perpendicular (Heisenberg scaling possible) to the signal. In the former case our new bounds show that negative correlations are beneficial; for the latter we show evidence that the bounds are tight. We also apply the bounds to collisional thermometry, i.e., estimation of a parameter of the environment, showing evidence that entangled probes may provide only a limited advantage.

DOI: 10.1103/jy3v-wkcb

Introduction—Reducing noise effects is the foremost challenge in advancing quantum technologies [1]. Real-world quantum systems suffer from environmental interactions, leading to decoherence and particle loss, yet quantum error correction (QEC) codes often help to achieve quantum advantage [2].

In quantum metrology [3–13], the most spectacular quantum advantage is the Heisenberg scaling (HS), a precision that scales as $1/N$, where N is the number of resources (such as particles or channel uses), whereas for standard scaling (SS) the precision is proportional to $1/\sqrt{N}$. Strategies that attain the HS are known for noiseless [3,14] and noisy [15–17] phase estimation—in the latter case, proper QEC is indispensable.

Recent developments in quantum channel estimation theory enable full characterization of precision limits under *uncorrelated* noise, allowing one to determine the optimal scaling (HS or SS) and the corresponding constant via a simple semidefinite program (SDP) [17–27]. However, noise and signal correlations are significant in many systems [28–37]. Despite various case studies including temporal [38–47] and spatial [48–53] correlations, or both [54,55], metrological bounds for these cases are missing, especially in large-N limit. As such, one cannot assess the fundamental optimality of a specific protocol. In the work presented here we fill this gap and derive universal bounds for correlated noise models.

State-of-the-art bounds—Numerous metrological tasks, e.g., phase estimation or superresolution imaging [56], amount to the estimation of a single parameter θ—the goal is to find a (locally) unbiased estimator $\tilde{\theta}$ with minimal variance $\Delta^2\tilde{\theta}$. When the parameter is encoded in a quantum state ρ_θ, the *quantum Cramér-Rao bound* (QCRB) says that $\Delta^2\tilde{\theta} \geq 1/F(\rho_\theta)$, where $F(\rho_\theta)$ is the *quantum Fisher information* (QFI) [57,58] (see [59] Sec. A for the definition). The QFI is a local quantity that depends only on ρ_θ and its derivative $\dot{\rho}_\theta = \partial_\theta \rho_\theta$ evaluated at some specific value of θ.

Let $|\Psi_\theta\rangle$ be a purification of ρ_θ, then it holds that $F(|\Psi_\theta\rangle) \geq F(\rho_\theta)$ because any measurement on ρ_θ can be implemented on $|\Psi_\theta\rangle$. Moreover, the QFI is equal to a minimization over all purifications of ρ_θ:

$$F(\rho_\theta) = 4 \min_{|\Psi_\theta\rangle} \langle \dot{\Psi}_\theta | \dot{\Psi}_\theta \rangle, \tag{1}$$

which follows from the existence of a *QFI nonincreasing purification* (QFI NIP) satisfying [18,60] $F(|\Psi_\theta\rangle) = F(\rho_\theta)$.

In quantum metrology, the parameter θ is encoded in a quantum channel $\Lambda_\theta(\bullet) = \sum_k K_{k,\theta} \bullet K_{k,\theta}^\dagger$, where $K_{k,\theta}$ are Kraus operators [20]. The ultimate precision of θ estimation is quantified by the *channel* QFI $\mathcal{F}(\Lambda_\theta) = \max_\rho F(\rho_\theta)$, where $\rho_\theta = (\Lambda_\theta \otimes \mathcal{I}_\mathcal{A})(\rho)$ and $\rho \in \mathcal{L}(\mathcal{H} \otimes \mathcal{A})$ is a possibly entangled input state of probe ($\mathcal{H}$) and noiseless ancillary ($\mathcal{A}$) systems, $\mathcal{L}(\bullet)$ is the set of linear operators acting on $\bullet$, and $\mathcal{I}$ the identity channel. Starting from (1) one finds [18]

$$\mathcal{F}(\Lambda_\theta) = 4 \min_h \|\alpha\|, \qquad \alpha = \sum_k \dot{\tilde{K}}_{k,\theta}^\dagger \tilde{K}_{k,\theta}, \tag{2}$$

where $\|\bullet\|$ is the operator norm, $\tilde{K}_{k,\theta} = \dot{K}_{k,\theta} - i h_{kk'} K_{k',\theta}$ and h is a hermitian matrix; the summation is performed

over repeated indices. Importantly, (2) can be recast as a simple SDP [20].

The real potential of quantum metrology appears when $N > 1$ channel copies are probed collectively using *adaptive* or *active quantum feedback* (AD) strategies, where channels $\Lambda_\theta\colon \mathcal{L}(\mathcal{H}_{2i-1}) \to \mathcal{L}(\mathcal{H}_{2i})$ for $i \in \{1, 2, \ldots, N\}$ are probed sequentially, and arbitrary quantum control channels $\mathrm{C}_i\colon \mathcal{L}(\mathcal{H}_{2i} \otimes \mathcal{A}_i) \to \mathcal{L}(\mathcal{H}_{2i+1} \otimes \mathcal{A}_{i+1})$ can act on a probe and arbitrarily large ancilla $\mathcal{A}_i$ after each Λ_θ. The AD class covers all strategies besides those involving indefinite causal order [61]—in particular *parallel* strategies, i.e., all channels simultaneously probed by an entangled state, form a subset of AD [22].

A general AD strategy can be represented by a *quantum comb*, which models the sequence of channels sharing a common environment [62]. Formally, $\mathrm{E} \in \mathrm{Comb}[(\mathcal{K}_1, \mathcal{K}_2), \ldots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$ when E is a quantum channel with inputs $\mathcal{K}_1, \mathcal{K}_3, \ldots \mathcal{K}_{2N-1}$ and outputs $\mathcal{K}_2, \mathcal{K}_4, \ldots \mathcal{K}_{2N}$, satisfying some extra linear conditions (see [59] Sec. B) ensuring that output $\mathcal{K}_{2j}$ may only depend on input $\mathcal{K}_{2k-1}$ when $j \leq k$. The Hilbert spaces $\mathcal{K}_{2k-1}, \mathcal{K}_{2k}$ can be interpreted as the input and output of the kth comb tooth, respectively.

Input state $\rho \in \mathcal{L}(\mathcal{H}_1)$ and controls C_i connected with ancillae $\mathcal{A}_i$ are represented using a comb $\mathrm{C}^{(N)} \in \mathrm{Comb}[(\varnothing, \mathcal{H}_1), (\mathcal{H}_2, \mathcal{H}_3), \ldots, (\mathcal{H}_{2N-2}, \mathcal{H}_{2N-1} \otimes \mathcal{A}_N)]$, where $\varnothing$ is a trivial space (no input), see Fig. 1(a). The parameter encoding is described by a comb $\Lambda_\theta^{(N)} \in \mathrm{Comb}[(\mathcal{H}_1, \mathcal{H}_2), \ldots, (\mathcal{H}_{2N-1}, \mathcal{H}_{2N})]$, which can model any type of noise and signal correlations; for uncorrelated channels it reduces to $\Lambda_\theta^{(N)} = \Lambda_\theta^{\otimes N}$. The output state $\rho_\theta \in \mathcal{L}(\mathcal{H}_{2N} \otimes \mathcal{A}_N)$ is obtained by concatenating the corresponding inputs and outputs of $\mathrm{C}^{(N)}$ and $\Lambda_\theta^{(N)}$ through the *link product* operation [62,63], $\rho_\theta = \mathrm{C}^{(N)} \star \Lambda_\theta^{(N)}$.

The *comb QFI* [46] $\mathcal{F}_{\mathrm{AD}}^{(N)} = \max_{\mathrm{C}^{(N)}} F(\mathrm{C}^{(N)} \star \Lambda_\theta^{(N)})$ quantifies the ultimate estimation precision for a parameter encoded in a comb $\Lambda_\theta^{(N)}$. To evaluate $\mathcal{F}_{\mathrm{AD}}^{(N)}$, we should decompose the Choi-Jamiołkowski (CJ) operator [64] of $\Lambda_\theta^{(N)}$ as $\Lambda_\theta^{(N)} = \sum_k |K_{k,\theta}^{(N)}\rangle\langle K_{k,\theta}^{(N)}|$, where $|K_{k,\theta}^{(N)}\rangle$ are vectorized Kraus operators; we use roman font for channels, and italics for the corresponding CJ operators. Analogously to α in (2), a *performance operator* is defined as $\boldsymbol{\alpha}^{(N)} = \mathrm{Tr}_{\mathrm{out}}(\sum_k |\dot{\tilde{K}}_{k,\theta}^{(N)}\rangle\langle \dot{\tilde{K}}_{k,\theta}^{(N)}|)$ [65], where $\mathrm{Tr}_{\mathrm{out}}$ is the partial trace over the last output space of $\Lambda_\theta^{(N)}$ $(\mathcal{H}_{2N})$, $|\dot{\tilde{K}}_{k,\theta}^{(N)}\rangle = |\dot{K}_{k,\theta}^{(N)}\rangle - i h_{kk'} |K_{k',\theta}^{(N)}\rangle$, h is a hermitian matrix. Thus, the comb QFI can be rewritten as [46,61]

$$\mathcal{F}_{\mathrm{AD}}^{(N)} = 4 \min_h \max_{\tilde{C}^{(N)}} \mathrm{Tr}\left(\boldsymbol{\alpha}^{(N)} \tilde{C}^{(N)}\right), \tag{3}$$

where $\tilde{C}^{(N)} = \mathrm{Tr}_{\mathcal{A}_N} C^{(N)}$. This optimization can be formulated as an SDP [46,61]. Unfortunately, its complexity grows exponentially with N, making it intractable for $N \gtrsim 5$.

For *uncorrelated* noise, one can circumvent this problem by computing bounds for $\mathcal{F}_{\mathrm{AD}}$, using the iteration [27]

$$\mathcal{F}_{\mathrm{AD}}^{(l+1)} \leq \mathcal{F}_{\mathrm{AD}}^{(l)} + 4 \min_h \left[\|\alpha\| + \sqrt{\mathcal{F}_{\mathrm{AD}}^{(l)}} \|\beta\| \right], \tag{4}$$

where $\beta = \sum_k \dot{\tilde{K}}_k^\dagger K_k$ (see [59] Sec. C1 for a new, simplified derivation). If there exists h for which $\beta = 0$, $\mathcal{F}_{\mathrm{AD}}^{(N)}$ scales linearly with N for $N \gg 1$. Hence

$$\lim_{N\to\infty} \mathcal{F}_{\mathrm{AD}}^{(N)}/N \leq 4 \min_h \|\alpha\| \ \text{ s.t. } \beta = 0. \tag{5}$$

If instead $\beta \neq 0$ for all h, HS is in principle allowed, and the asymptotic form of (4) reads

$$\lim_{N\to\infty} \mathcal{F}_{\mathrm{AD}}^{(N)}/N^2 \leq 4 \min_h \|\beta\|^2. \tag{6}$$

The iterative bound (4) and asymptotic bounds (5), (6) can be formulated as SDPs [21,27]. The bound (4) is not tight in general, but its asymptotic limits (5) and (6) are always saturable using a parallel strategy [26].

Bound for correlated noise—In what follows, we present novel bounds analogous to (4)–(6) valid for all correlated noise models. Unlike the exact comb QFI formula (3), these bounds are efficiently calculable for arbitrarily large N.

The comb $\Lambda_\theta^{(N)}$ representing a correlated noise metrological model is a link product of its teeth $\overleftrightarrow{\Lambda}_\theta\colon \mathcal{L}(\mathcal{H}_{2i-1} \otimes \mathcal{R}_{i-1}) \to \mathcal{L}(\mathcal{H}_{2i} \otimes \mathcal{R}_i)$ for $i \in \{1, 2, \ldots, N\}$, where $\mathcal{R}_i$ are inaccessible environmental spaces modeling correlations; the fixed state σ_{in} inputs the first register $\mathcal{R}_0$, the last register $\mathcal{R}_N$ is traced out—see Fig. 1(a). The $\leftrightarrow$ symbol in $\overleftrightarrow{\Lambda}_\theta$ indicates unconcatenated environmental spaces. Subsequent teeth are assumed identical, though this assumption can be easily dropped.

We divide $\Lambda_\theta^{(N)}$ into blocks $\overleftrightarrow{\Lambda}_\theta^{(m)}$ of m teeth each [Fig. 1(b)], intertwined with control combs $\mathrm{C}^{(l,m)}$ connected via ancillae $\mathcal{A}_l$. Any global strategy $\mathrm{C}^{(N)}$ can be simulated using appropriate $C^{(l,m)}$ with sufficiently large $\mathcal{A}_l$.

Notably, combs $\mathrm{C}^{(l,m)}$ control not only $\mathcal{H}$ and $\mathcal{A}$ at each protocol stage, but also environmental subspaces $\mathcal{R}$ between different blocks $\overleftrightarrow{\Lambda}_\theta^{(m)}$. In a normal AD scheme, the environment is always directly sent to the next tooth of $\Lambda_\theta^{(N)}$, but here we allow for arbitrary operations on the environment every m teeth. These "environment leakages" often make the bound less tight, but are necessary when dividing N correlated channels into blocks of m while keeping the validity of the bound—neglecting the environmental information may lead to an underestimated result.

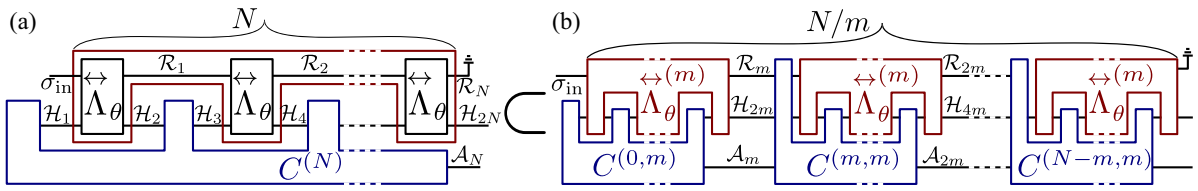


FIG. 1. The comb $\Lambda_\theta^{(N)}$ consists of teeth $\overleftrightarrow{\Lambda}_\theta$ connected with environmental spaces $\mathcal{R}$; physical estimation strategies $C^{(N)}$ act only on probe subspaces $\mathcal{H}$ (a). To derive an upper bound, we divide $\Lambda_\theta^{(N)}$ into blocks of m teeth $\overleftrightarrow{\Lambda}_\theta^{(m)}$ (b), and allow for more powerful strategies, in which a purified state of $\mathcal{R}$, $\mathcal{H}$, $\mathcal{A}$ after each block is an input of the next control $\mathrm{C}^{(l,m)}$.

Additionally, let us replace the joint mixed state $\rho_\theta^{(l)}$ of $\mathcal{H}, \mathcal{R}, \mathcal{A}$ after each block $\overleftrightarrow{\Lambda}_\theta^{(m)}$ with its QFI NIP $|\Psi_\theta^{(l)}\rangle$. This can only increase the final QFI, since any operation possible with a state is also possible with its purification.

Through purification, the space carrying the information about $\rho_\theta^{(l)}$ and $\dot{\rho}_\theta^{(l)}$ reduces to the qubit subspace $\mathcal{V}_l = \mathrm{span}(|\Psi_\theta^{(l)}\rangle, |\dot{\Psi}_\theta^{(l)}\rangle)$. After fixing a basis of $\mathcal{V}_l$ and constructing a comb $\Upsilon_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle|\Psi_\theta^{(l)}| \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}$ containing the purified and "qubitized" $\rho_\theta^{(l)}$ together with next m teeth, we get (see [59] Sec. D1 for a detailed derivation) $\mathcal{F}_{\mathrm{AD}}^{(i)} \leq F^{(i)}$, where $F^{(0)} = 0$,

$$F^{(l+m)} = \max_{\mathrm{C}^{(l,m)}} F\left(\mathrm{C}^{(l,m)} \star \Upsilon_\theta^{(l,m)}\right), \qquad \Upsilon_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overleftrightarrow{\Lambda}_\theta^{(m)},$$
$$\dot{\Upsilon}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overleftrightarrow{\dot{\Lambda}}_\theta^{(m)} + \frac{\sqrt{F^{(l)}}}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}, \tag{7}$$

maximizing over $\mathrm{C}^{(l,m)}$ amounts to computing the comb QFI of $\Upsilon_\theta^{(l,m)}$.

The bound generated by (7) can be applied to uncorrelated noise—then additional QFI NIPs are the only reason for the bound untightness, making it at least as tight as (4). The new bound can be then tighter than the old one even for $m = 1$, increasing m further tightens (7), but only for finite N—see [59] Sec. C2.

For correlated noise, for which (4) does not apply, the role of m is more significant, since information leaks from the environment every m steps. Consequently, increasing m tightens the bound also asymptotically.

Using (7), one can prove that (see [59] Sec. D2)

$$F^{(l+m)} \leq F^{(l)} + 4 \min_h \left[a^{(m)} + \sqrt{F^{(l)}} b^{(m)} \right],$$
$$a^{(m)} = \max_{\tilde{C}^{(m)}} \mathrm{Tr}\left(\boldsymbol{\alpha}^{(m)} \tilde{C}_{00}^{(m)}\right), \quad b^{(m)} = \max_{\tilde{C}^{(m)}} \mathrm{ReTr}\left(\boldsymbol{\beta}^{(m)} \tilde{C}_{10}^{(m)}\right), \tag{8}$$

where $\tilde{C}^{(m)}$ is a comb of the same type as any $\tilde{C}^{(l,m)} = \mathrm{Tr}_{\mathcal{A}_{l+m}} C^{(l,m)}$, $\boldsymbol{\alpha}^{(m)}$ is a performance operator of $\overleftrightarrow{\Lambda}_\theta^{(m)}$, $\boldsymbol{\beta}^{(m)}$ is a comb-adapted analog of β (defined as $\boldsymbol{\alpha}^{(m)}$, but without derivative acting on the $\langle\bullet|$ part), $\tilde{C}_{ij}^{(m)} = {}_{\mathcal{V}_l}\langle i|\tilde{C}^{(m)}|j\rangle_{\mathcal{V}_l}$. Equation (8) is similar to (4), but $\|\alpha\|$ and $\|\beta\|$ are replaced with their comb-adapted analogues $a^{(m)}$, $b^{(m)}$. Consequently, we get asymptotic bounds analogous to (5) and (6) valid for the correlated case; when $b^{(m)} = 0$ for some h, then

$$\lim_{N\to\infty} \frac{\mathcal{F}_{\mathrm{AD}}^{(N)}}{N} \leq \frac{4}{m} \min_h a^{(m)} \text{ s.t. } b^{(m)} = 0, \tag{9}$$

so the QFI scales at most linearly. If $b^{(m)} \neq 0$ for all h, then HS is allowed:

$$\lim_{N\to\infty} \frac{\mathcal{F}_{\mathrm{AD}}^{(N)}}{N^2} \leq \frac{4}{m^2} \min_h b^{(m)2}. \tag{10}$$

In [59] Sec. D5 we show how to translate (7), (9), and (10) into single SDPs using the technique from Refs. [46,61]. Even though bounds (7), (9), (10) are not generally tight, they can be tightened by increasing m (which increases the complexity of the resulting SDPs).

Example: Correlated dephasing—To illustrate the practical relevance of the new bounds, we consider phase estimation in the presence of correlated dephasing. Single-qubit dephasing shrinks the Bloch vector by a factor $\eta = \cos(\epsilon)$ toward the axis defined by a unit vector $\vec{n}$. It corresponds to a random rotation by angle $+\epsilon$ or $-\epsilon$ around $\vec{n}$, each with probability $1/2$.

To introduce basic correlations, we assume the consecutive rotational signs $r_i \in \{+, -\}$ follow a binary Markov chain given by conditional probabilities [66]

$$S_{i|i-1}(r_i|r_{i-1}) = (1 + r_i r_{i-1} C)/2, \tag{11}$$

where $C \in [-1, 1]$ controls correlation: $C = 0$ means no correlation, $C = \pm 1$ gives maximal (anti)correlation. The corresponding quantum model $\Lambda_\theta^{(N)}$ is constructed by alternating unitaries $V_\theta = R_{\hat{z}}(\theta) R_{\vec{n}}(+\epsilon) \otimes |+\rangle\langle+| + R_{\hat{z}}(\theta) R_{\vec{n}}(-\epsilon) \otimes |-\rangle\langle-|$ acting on $\mathcal{H}$ and $\mathcal{R}$ with mixing operations S performing the stochastic map (11) on the basis $|\pm\rangle$ of $\mathcal{R}$ [66], $R_{\vec{n}}(\varphi) = e^{-\frac{i}{2}\varphi \vec{n}\cdot\vec{\sigma}}$, $\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z]$, notice that noise (random rotation by angle $\pm\epsilon$) precedes signal (rotation by θ).

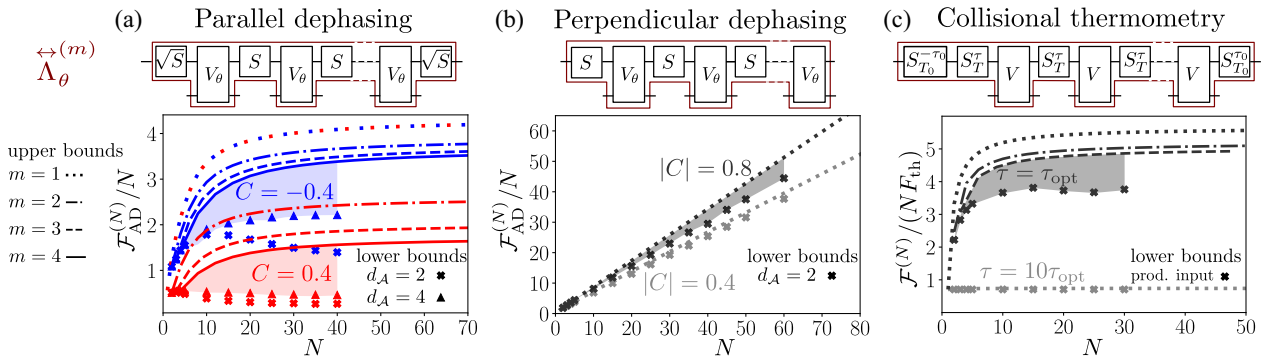


FIG. 2. Combs $\overleftrightarrow{\Lambda}_\theta^{(m)}$ and resulting upper bounds for QFI for three different models, shown respectively in panels (a)–(c). Lower bounds obtained using tensor networks techniques [66,67] are shown for comparison, areas between the corresponding tightest lower and upper bounds are shaded.

The comb $\Lambda_\theta^{(N)}$ can be split into pieces $\overleftrightarrow{\Lambda}_\theta^{(m)}$ in different ways, and this choice affects the tightness of the resulting bound. First, we consider *parallel dephasing* $\vec{n}=\hat{z}$, when noise commutes with signal. If we cut the chain after S and before V_θ, we get information on the sign of ϵ for the first dephasing channel in each block, while cutting the chain after V_θ and before S, we get this information for the *last* channel in each block. Consequently, the first or last channel in each block is effectively noiseless, so the bound manifests an unrealistic HS.

To resolve this issue, we write S as a concatenation of two identical stochastic maps, $\mathrm{S}=\sqrt{\mathrm{S}}\circ\sqrt{\mathrm{S}}$, see [59] Sec. E1 for the exact construction of $\sqrt{\mathrm{S}}$. The chain cuts are then made in the middle of every mth map S, the resulting $\overleftrightarrow{\Lambda}_\theta^{(m)}$ consists of V_θ with S between them and $\sqrt{\mathrm{S}}$ at both ends; see Fig. 2(a). We insert $\overleftrightarrow{\Lambda}_\theta^{(m)}$ into the recursive procedure (7) to derive bounds, sending the maximally mixed state as input for $\mathcal{R}$ for the first iteration (probabilities of $\pm$ are equal in the first channel).

We observe that the HS is not possible for correlated parallel dephasing for any $-1<C<1$ and $0\le\eta<1$. In Fig. 2(a) we compare upper bounds, calculated for positive ($C=0.4$) and negative ($C=-0.4$) correlations for $\eta=0.75$, with lower bounds—QFIs achievable using protocols with small ancilla dimension $d_\mathcal{A}$, calculated using the tensor-network algorithm from Refs. [66,68]. These bounds can be made arbitrarily tight by increasing m and $d_\mathcal{A}$, respectively; however, this quickly becomes numerically costly. We performed calculations up to $d_\mathcal{A}=4$ and $m=4$, for which the lower and upper bounds still do not coincide. Nevertheless, this is the first time that precision limits of metrological protocols for correlated noise with arbitrarily large $d_\mathcal{A}$ are explicitly evaluated. This allows us to deduce that negative correlations offer metrological advantage over positive correlations.

We also study *perpendicular dephasing* $\vec{n}=\hat{x}$, for which the HS is possible [27]. In this case, we observe that the tightest bound is obtained by splitting the comb after V_θ and before S; see Fig. 2(b). This is because noncommuting noise acts before the signal, and the information about rotation sign is useless after the signal. Interestingly, the resulting bounds are equally tight for $m\in\{1,2,3\}$, suggesting that the bound is already tight for $m=1$. This is further supported by the observation that the QFI achievable by an adaptive protocol with $d_\mathcal{A}=2$ is very close to the upper bound calculated for $m=1$; see Fig. 2(b). We also show that for perpendicular dephasing the QFI does not depend on the sign of C—all correlations act as an extra source of information, since the achievable precision increases with $|C|$.

Example: Collisional thermometry—From a different perspective, the environment register $\mathcal{R}$ may not represent noise, but an inaccessible quantum system, measured indirectly by sequential interactions with the mediators $\mathcal{H}_k$. This setup corresponds to quantum collision models [69] and, for factorized probes, to quantum Markov chains, widely studied for parameter estimation [70,71].

We consider a simple realization of *collisional quantum thermometry* [72–75], where a thermometer qubit ($\mathcal{R}$) probes a thermal bath, whose temperature T is to be estimated by measuring mediator qubits ($\mathcal{H}$) that sequentially interact with the thermometer. Each collision is a partial energy exchange between $\mathcal{H}$ and $\mathcal{R}$ governed by the unitary $V=e^{-igt(\sigma_+\otimes\sigma_-+\sigma_-\otimes\sigma_+)}$, g is a coupling constant, t is an interaction time, $\sigma_\pm=(\sigma_x\pm i\sigma_y)$; a full energy swap occurs for $gt=0.5\pi$. Between collisions, the thermometer evolves via a nonunitary channel $\mathrm{S}_T^\tau(\rho)=e^{\tau\mathcal{L}_T}(\rho)$, where $\mathcal{L}_T$ is a thermalizing Lindbladian, see [72] and [59] Sec. E2 for an exact definition. For $\tau\to\infty$, any input state gets fully thermalized, $\mathrm{S}_T^\infty(\rho)=\rho_{\mathrm{th},T}$, where $\rho_{\mathrm{th},T}=e^{-H_\mathcal{E}/k_BT}/Z$ is a Gibbs state, $H_\mathcal{E}=\hbar\Omega\sigma_z/2$. We assume that the thermometer is initialized in a thermal state $\rho_{\mathrm{th},T}$ [76].

Various strategies have been explored in this setting [72,73], yet their rigorous study for large N is challenging due to environment-induced correlations among measured mediator qubits. Our bounds fully account for these

correlations. Since parallel schemes are a subset of adaptive ones, the bounds also apply to nonadaptive scenarios with arbitrary initial entanglement among mediators and a final collective measurement.

We define *thermal Fisher information* as $F_{\mathrm{th}} = F(\rho_{\mathrm{th},T})$. Interestingly, the channel QFI of S_T^{τ} can exceed F_{th} for finite τ, reaching $\mathcal{F}(\mathrm{S}_T^{\tau_{\mathrm{opt}}}) \approx 5.66 F_{\mathrm{th},T}$ at the optimal thermalization time τ_{opt} when $k_B T/\hbar\Omega = 2$ [72]. We consider a collisional model with $gt = 0.35\pi < 0.5\pi$, which leads to correlations among output mediators. We then compute upper bounds for $\mathcal{F}_{\mathrm{AD}}$ for $m = 1, 2, 3$ in two regimes: $\tau = \tau_{\mathrm{opt}}$ and $\tau = 10\tau_{\mathrm{opt}}$.

To tighten the bounds, we insert an additional thermalization map $\mathrm{S}_{T_0}^{\tau_0}$ at the end of each $\overleftrightarrow{\Lambda}_\theta^{(m)}$, where T_0 denotes a fixed, reference temperature—i.e., $(d/dT)\mathrm{S}_{T_0}^{\tau_0} = 0$, so the map is insensitive to T; see Fig. 2(c). To preserve the overall channel structure, we prepend an inverse map $\mathrm{S}_{T_0}^{-\tau_0}$, ensuring the two cancel out when constructing $\Lambda_\theta^{(N)}$ from the building blocks $\overleftrightarrow{\Lambda}_\theta^{(m)}$. We choose $\tau_0 < \tau$ to keep $\mathrm{S}_T^{\tau} \circ \mathrm{S}_{T_0}^{-\tau_0}$ completely positive, and optimize τ_0 to obtain the tightest possible bounds.

For $\tau = \tau_{\mathrm{opt}}$, correlations play a significant role, as reflected by the tightening of asymptoptotic bounds with increasing m [see Fig. 2(c)]. In contrast, for $\tau = 10\tau_{\mathrm{opt}}$, the output of S_T^{τ} is nearly thermal for any input, correlations are negligible, and the bounds are almost insensitive to m.

To assess tightness, we compare our upper bounds with lower bounds obtained by optimizing the output QFI over input product states. This optimization, performed efficiently via tensor networks [67,68], shows near saturation of the upper bound for $\tau = 10\tau_{\mathrm{opt}}$, confirming that product states are optimal in this regime. For $\tau = \tau_{\mathrm{opt}}$, a gap remains, but the bounds still closely follow the QFI trend, indicating limited potential gain from entanglement or adaptivity.

Conclusions—These examples represent only a fraction of the possible applications of the new bounds. The formalism allows one to handle both classical and inherently quantum correlations, i.e., generic non-Markovian open quantum systems [29,30,32,34]. Although the sequential scheme suggests a focus on temporal correlations, the bound is also valid for analogous spatial correlations, where channels are sampled in parallel, since this is a subset of adaptive strategies. Furthermore, this approach provides a tighter bound for uncorrelated noise than the current state of the art [27].

This work is complementary to Refs. [66,77], where a tensor network approach was proposed to find AD estimation strategies with a limited ancilla dimension, providing a *lower* bound on $\mathcal{F}_{\mathrm{AD}}$. In combination, these works establish a comprehensive framework to study metrological protocols in the presence of correlated noise, for a large number of probed channels. The package containing the functions to compute the bounds as well as perform tensor-network optimization to obtain the lower bounds as in Fig. 2 is now publicly available [68].

A challenging extension will be to proceed toward the continuous time limit, where quantum combs (known as process tensors [34]) are routinely used in numerical studies of the dynamics and control of non-Markovian open systems [32,78–81]. Finally, it will be interesting to see if similar ideas apply to the related problem of discriminating quantum combs [82–89].

Acknowledgments—We thank Wojciech Górecki and Andrea Smirne for helpful discussions. This Skłodowskawork was supported by the National Science Center (Poland) Grant No. 2020/37/B/ST2/02134. F. A. acknowledges support from Marie-Curie Action EUHORIZON-MSCA-2021PF-01 (project QECANM, Grant No. 101068347).

[1] D. Suter and G. A. Álvarez, Colloquium: Protecting quantum information against environmental noise, Rev. Mod. Phys. **88**, 041001 (2016).
[2] *Quantum Error Correction*, edited by D. A. Lidar and T. A. Brun, 1st ed. (Cambridge University Press, Cambridge, England, 2013).
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306**, 1330 (2004).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).
[5] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum limits in optical interferometry, in *Progress in Optics*, Volume 60, edited by E. Wolf (Elsevier, Amsterdam, 2015), pp. 345–435.
[6] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).
[7] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).
[8] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nat. Photonics **12**, 724 (2018).
[9] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, AVS Quantum Sci. **2**, 024703 (2020).
[10] J. Liu, M. Zhang, H. Chen, L. Wang, and H. Yuan, Optimal scheme for quantum metrology, Adv. Quantum Technol. **5**, 2100080 (2022).
[11] L. Jiao, W. Wu, S.-Y. Bai, and J.-H. An, Quantum metrology in the noisy intermediate-scale quantum era, Adv. Quantum Technol. **8**, 2300218 (2025).
[12] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Fully-Optimized quantum metrology: Framework, tools, and applications, Adv. Quantum Technol. **7**, 2400094 (2024).
[13] V. Montenegro, C. Mukhopadhyay, R. Yousefjani, S. Sarkar, U. Mishra, M. G. Paris, and A. Bayat, Review: Quantum metrology and sensing with many-body systems, Phys. Rep. **1134**, 1 (2025).

[14] Z. Y. Ou, Fundamental quantum limit in precision phase measurement, Phys. Rev. A **55**, 2598 (1997).
[15] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Quantum error correction for metrology, Phys. Rev. Lett. **112**, 150802 (2014).
[16] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Improved quantum metrology using quantum error correction, Phys. Rev. Lett. **112**, 080801 (2014).
[17] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).
[18] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, J. Phys. A **41**, 255304 (2008).
[19] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, Nat. Phys. **7**, 406 (2011).
[20] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, Nat. Commun. **3**, 1063 (2012).
[21] J. Kołodyński and R. Demkowicz-Dobrzański, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).
[22] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, Phys. Rev. Lett. **113**, 250801 (2014).
[23] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, Quantum **1**, 27 (2017).
[24] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general Markovian noise, Phys. Rev. X **7**, 041009 (2017).
[25] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, Phys. Rev. Res. **2**, 013235 (2020).
[26] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).
[27] S. Kurdziałek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, Phys. Rev. Lett. **131**, 090801 (2023).
[28] K. Banaszek, A. Dragan, W. Wasilewski, and C. Radzewicz, Experimental demonstration of entanglement-enhanced classical communication over a quantum channel with correlated noise, Phys. Rev. Lett. **92**, 257901 (2004).
[29] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Quantum channels and memory effects, Rev. Mod. Phys. **86**, 1203 (2014).
[30] I. de Vega and D. Alonso, Dynamics of non-Markovian open quantum systems, Rev. Mod. Phys. **89**, 015001 (2017).
[31] P. Szańkowski, G. Ramon, J. Krzywda, D. Kwiatkowski, and Ł. Cywiński, Environmental noise spectroscopy with qubits subjected to dynamical decoupling, J. Phys. Condens. Matter **29**, 333001 (2017).
[32] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, Phys. Rev. A **97**, 012127 (2018).
[33] U. von Lüpke, F. Beaudoin, L. M. Norris, Y. Sung, R. Winik, J. Y. Qiu, M. Kjaergaard, D. Kim, J. Yoder, S. Gustavsson, L. Viola, and W. D. Oliver, Two-qubit spectroscopy of spatiotemporally correlated quantum noise in superconducting qubits, Phys. Rev. X Quantum **1**, 010305 (2020).
[34] S. Milz and K. Modi, Quantum stochastic processes and quantum non-Markovian phenomena, PRX Quantum **2**, 030201 (2021).
[35] R. Harper and S. T. Flammia, Learning correlated noise in a 39-qubit quantum processor, PRX Quantum **4**, 040311 (2023).
[36] F. A. Mele, G. D. Palma, M. Fanizza, V. Giovannetti, and L. Lami, Optical fibers with memory effects and their quantum communication capacities, IEEE Trans. Inf. Theory **70**, 8844 (2024).
[37] J. Preskill, Sufficient condition on noise correlations for scalable quantum computing, Quantum Inf. Comput. **13**, 181 (2013).
[38] Y. Matsuzaki, S. C. Benjamin, and J. F. Fitzsimons, Magnetic field sensing beyond the standard quantum limit under the effect of decoherence, Phys. Rev. A **84**, 012103 (2011).
[39] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum metrology in non-Markovian environments, Phys. Rev. Lett. **109**, 233601 (2012).
[40] K. Macieszczak, Zeno limit in frequency estimation with non-Markovian environments, Phys. Rev. A **92**, 010102(R) (2015).
[41] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański, Ultimate precision limits for noisy frequency estimation, Phys. Rev. Lett. **116**, 120801 (2016).
[42] J. F. Haase, A. Smirne, J. Kołodyński, R. Demkowicz-Dobrzański, and S. F. Huelga, Fundamental limits to frequency estimation: A comprehensive microscopic perspective, New J. Phys. **20**, 053009 (2018).
[43] P. Szańkowski, M. Trippenbach, and J. Chwedeńczuk, Parameter estimation in memory-assisted noisy quantum interferometry, Phys. Rev. A **90**, 063619 (2014).
[44] A. Smirne, A. Lemmer, M. B. Plenio, and S. F. Huelga, Improving the precision of frequency estimation via long-time coherences, Quantum Sci. Technol. **4**, 025004 (2019).
[45] D. Tamascelli, C. Benedetti, H.-P. Breuer, and M. G. A. Paris, Quantum probing beyond pure dephasing, New J. Phys. **22**, 083027 (2020).
[46] A. Altherr and Y. Yang, Quantum metrology for non-Markovian processes, Phys. Rev. Lett. **127**, 060501 (2021).
[47] X. Yang, X. Long, R. Liu, K. Tang, Y. Zhai, X. Nie, T. Xin, J. Li, and D. Lu, Control-enhanced non-Markovian quantum metrology, Commun. Phys. **7**, 282 (2024).
[48] J. Jeske, J. H. Cole, and S. F. Huelga, Quantum metrology subject to spatially correlated Markovian noise: Restoring the Heisenberg limit, New J. Phys. **16**, 073039 (2014).
[49] D. Layden and P. Cappellaro, Spatial noise filtering through error correction for quantum sensing, npj Quantum Inf. **4**, 30 (2018).
[50] J. Czajkowski, K. Pawłowski, and R. Demkowicz-Dobrzański, Many-body effects in quantum metrology, New J. Phys. **21**, 053031 (2019).
[51] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Ancilla-free quantum error correction codes for quantum metrology, Phys. Rev. Lett. **122**, 040502 (2019).

[52] D. Layden, M. Chen, and P. Cappellaro, Efficient quantum error correction of dephasing induced by a common fluctuator, Phys. Rev. Lett. **124**, 020504 (2020).

[53] G. Planella, M. F. B. Cenni, A. Acín, and M. Mehboudi, Bath-induced correlations enhance thermometry precision at low temperatures, Phys. Rev. Lett. **128**, 040502 (2022).

[54] F. Beaudoin, L. M. Norris, and L. Viola, Ramsey interferometry in correlated quantum noise environments, Phys. Rev. A **98**, 020102(R) (2018).

[55] F. Riberi, L. M. Norris, F. Beaudoin, and L. Viola, Frequency estimation under non-Markovian spatially correlated quantum noise: Restoring superclassical precision scaling, New J. Phys. **24**, 103011 (2022).

[56] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, Phys. Rev. X **6**, 031033 (2016).

[57] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[58] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. (Edizioni della Normale, Pisa, 2011).

[59] See Supplemental Material at http://link.aps.org/supplemental/10.1103/jy3v-wkcb for more information.

[60] J. Kołodyński, Precision bounds in noisy quantum metrology, Ph.D. thesis, University of Warsaw, 2014, arXiv:1409.0535.

[61] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal strategies of quantum metrology with a strict hierarchy, Phys. Rev. Lett. **130**, 070803 (2023).

[62] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, Phys. Rev. Lett. **101**, 060401 (2008).

[63] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[64] The CJ operator of a channel $\Lambda: \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}$ is defined as $\Lambda = \Lambda \otimes \mathcal{I}(|\Psi_+\rangle\langle\Psi_+|)$, where $|\Psi_+\rangle = \sum_i |i\rangle|i\rangle$, $|i\rangle$ is orthonormal basis of $\mathcal{H}_{\text{in}}$, $\mathcal{I}$ is and identity channel.

[65] The performance operator is often defined as transposition of this expression [46]. It does not affect the further results —for example, (3) remains valid irrespectively of the convention because $(C^{(N)})^T$ is a comb iff $C^{(N)}$ is a comb. The typical notation for performance operator is Ω, here we use $\boldsymbol{\alpha}$ because of analogy with a single-channel object.

[66] S. Kurdzialek, P. Dulian, J. Majsak, S. Chakraborty, and R. Demkowicz-Dobrzanski, Quantum metrology using quantum combs and tensor network formalism, New J. Phys. **27**, 013019 (2025).

[67] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański, Tensor-network approach for quantum metrology in many-body quantum systems, Nat. Commun. **11**, 250 (2020).

[68] P. Dulian, S. Kurdziałek, and R. Demkowicz-Dobrzański, Qmetro++, https://github.com/pdulian/qmetro (2025).

[69] F. Ciccarello, S. Lorenzo, V. Giovannetti, and G. M. Palma, Quantum collision models: Open system dynamics from repeated interactions, Phys. Rep. **954**, 1 (2022).

[70] M. Guţă, Fisher information and asymptotic normality in system identification for quantum Markov chains, Phys. Rev. A **83**, 062324 (2011).

[71] A. Godley and M. Guta, Adaptive measurement filter: Efficient strategy for optimal estimation of quantum Markov chains, Quantum **7**, 973 (2023).

[72] S. Seah, S. Nimmrichter, D. Grimmer, J. P. Santos, V. Scarani, and G. T. Landi, Collisional quantum thermometry, Phys. Rev. Lett. **123**, 180602 (2019).

[73] A. Shu, S. Seah, and V. Scarani, Surpassing the thermal Cramér-Rao bound with collisional thermometry, Phys. Rev. A **102**, 042417 (2020).

[74] E. O'Connor, B. Vacchini, and S. Campbell, Stochastic collisional quantum thermometry, Entropy **23**, 1634 (2021).

[75] T. M. Mendonça, D. O. Soares-Pinto, and M. Paternostro, Information flow-enhanced precision in collisional quantum thermometry, arXiv:2407.21618.

[76] This assumption differs from Ref. [72,73], in which the initial state was assumed to be stationary state of both S_T^τ and V. However, this assumption does not affect the asymptoptotic results.

[77] Q. Liu and Y. Yang, Efficient tensor networks for control-enhanced quantum metrology, Quantum **8**, 1571 (2024).

[78] M. R. Jørgensen and F. A. Pollock, Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-Markovian path integrals, Phys. Rev. Lett. **123**, 240602 (2019).

[79] G. E. Fux, E. P. Butler, P. R. Eastham, B. W. Lovett, and J. Keeling, Efficient Exploration of Hamiltonian parameter space for optimal control of non-Markovian open quantum systems, Phys. Rev. Lett. **126**, 200401 (2021).

[80] E. P. Butler, G. E. Fux, C. Ortega-Taberner, B. W. Lovett, J. Keeling, and P. R. Eastham, Optimizing performance of quantum operations with non-Markovian decoherence: The tortoise or the hare?, Phys. Rev. Lett. **132**, 060401 (2024).

[81] C. Ortega-Taberner, E. O'Neill, E. Butler, G. E. Fux, and P. R. Eastham, Unifying methods for optimal control in non-Markovian quantum systems via process tensors, J. Chem. Phys. **161**, 124119 (2024).

[82] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory effects in quantum channel discrimination, Phys. Rev. Lett. **101**, 180501 (2008).

[83] G. Gutoski, On a measure of distance for quantum strategies, J. Math. Phys. (N.Y.) **53**, 032202 (2012).

[84] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, Phys. Rev. Res. **1**, 033169 (2019).

[85] K. Nakahira and K. Kato, Generalized quantum process discrimination problems, Phys. Rev. A **103**, 062606 (2021).

[86] K. Nakahira and K. Kato, Simple upper and lower bounds on the ultimate success probability for discriminating arbitrary finite-dimensional quantum processes, Phys. Rev. Lett. **126**, 200502 (2021).

[87] C. Hirche, Quantum network discrimination, Quantum **7**, 1064 (2023).

[88] G. Zambon, Process tensor distinguishability measures, Phys. Rev. A **110**, 042210 (2024).

[89] J. Bavaresco, M. Murao, and M. T. Quintino, Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination, Phys. Rev. Lett. **127**, 200504 (2021).

Universal bounds for quantum metrology in the presence of correlated noise: Supplementary Material

Stanisław Kurdziałek,[1] Francesco Albarelli,[2] and Rafał Demkowicz-Dobrzański[1]
[1] *Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland*
[2] *Scuola Normale Superiore, I-56126 Pisa, Italy*

Appendix A: Basics of Quantum Fisher Information

The quantum Fisher information (QFI) of an arbitrary mixed state ρ_θ can be calculated using the formula

$$F(\rho_\theta) = \mathrm{Tr}(\rho_\theta L_\theta^2), \quad \dot{\rho}_\theta = \frac{1}{2}\left(\rho_\theta L_\theta + L_\theta \rho_\theta\right), \tag{A1}$$

where dot denotes the derivative over θ, and the symmetric logarithmic derivative (SLD) matrix L_θ can be calculated by solving the rightmost equation from (A1). For pure state models $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, the closed analytical formula for the SLD and the QFI can be easily derived:

$$L_\theta = 2\left(|\dot{\psi}_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\dot{\psi}_\theta|\right), \quad F(|\psi_\theta\rangle) = 4(\langle\dot{\psi}_\theta|\dot{\psi}_\theta\rangle - |\langle\psi_\theta|\dot{\psi}_\theta\rangle|^2). \tag{A2}$$

When $|\Psi_\theta\rangle$ is a purification of ρ_θ, then $F(|\Psi_\theta\rangle) \geq F(\rho_\theta)$. Additionally, for each ρ_θ we can construct its QFI non-increasing purification (QFI NIP) satisfying

$$F(|\Psi_\theta\rangle) = F(\rho_\theta), \quad \langle\Psi_\theta|\dot{\Psi}_\theta\rangle = 0 \tag{A3}$$

The second condition can be always satisfied by multiplying $|\Psi_\theta\rangle$ with a global, θ-dependent phase.

Appendix B: Basics of quantum combs

Let us introduce the formal definition of quantum combs [1]. We use roman font for channels and italics for corresponding Choi-Jamiołkowski (CJ) operators: the CJ operator of a channel $\Lambda : \mathcal{H}_{\mathrm{in}} \to \mathcal{H}_{\mathrm{out}}$ is defined as $\mathit{\Lambda} = \Lambda \otimes \mathcal{I}(|\Psi_+\rangle\langle\Psi_+|)$, where $|\Psi_+\rangle = \sum_i |i\rangle|i\rangle$, $|i\rangle$ is an orthonormal basis of $\mathcal{H}_{\mathrm{in}}$, $\mathcal{I}$ is the identity channel. Let E be a quantum channel with with inputs $\mathcal{K}_1, \mathcal{K}_3, ..., \mathcal{K}_{2N-1}$ and outputs $\mathcal{K}_2, \mathcal{K}_4, ...\mathcal{K}_{2N}$. Then, $\mathrm{E} \in \mathrm{Comb}[(\mathcal{K}_1, \mathcal{K}_2), (\mathcal{K}_3, \mathcal{K}_4), ..., (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$ if and only if (iff) there exists a sequence of operators $E^{(k)}$ for $k = 1, 2, ..., N$ such that $E = E^{(N)}$ and:

1. $\mathrm{Tr}_{2k} E^{(k)} = E^{(k-1)} \otimes \mathbb{1}_{2k-1}$ for $k = 2, 3, ..., N$
2. $\mathrm{Tr}_2 E^{(1)} = \mathbb{1}_1$
3. $E \succeq 0$

Intuitively, $\mathcal{K}_{2k-1}, \mathcal{K}_{2k}$ are input/output spaces of the kth tooth of a comb. The above conditions imply the causal order of the teeth—the output $\mathcal{K}_{2l}$ can depend on the input $\mathcal{K}_{2j-1}$ only if $l \geq j$. In other words, quantum combs can be used to simulate channels with memory, when each output may depend on all preceding inputs. Therefore, they are the ideal tool to represent sequences of correlated channels or adaptive quantum strategies. Notably, a sequence $\mathrm{E}^{(k)}$ consists of quantum combs with increasing number of teeth: $\mathrm{E}^{(k)} \in \mathrm{Comb}[(\mathcal{K}_1, \mathcal{K}_2), ...(\mathcal{K}_{2k-1}, \mathcal{K}_{2k})]$. Two combs can be connected with each other using a *link product* operation $\star$ [1]. When E, F are two quantum combs, then $G = E \star F$ is a comb created by connecting the input spaces of E with the corresponding output spaces of F and the output spaces of E with the corresponding input spaces of F (only common spaces of E and F get connected). Let E, F act on sets of spaces M, N respectively. The link product is then formally defined using CJ operators :

$$\mathrm{G} = \mathrm{E} \star \mathrm{F} \iff G = \mathrm{Tr}_{M\cap N}\left[\left(E^{\mathrm{T}_{M\cap N}} \otimes \mathbb{1}_{N\setminus M}\right)\left(\mathbb{1}_{M\setminus N} \otimes F\right)\right],$$

where $M \cap N$ is the set of common (linked) spaces, $\mathrm{T}_{M\cap N}$ is a partial transposition (with respect to common spaces only). When $M = N$, then $\mathrm{E} \star \mathrm{F}$ is a scalar, using the comb conditions and the fact that each output of E is an input of F and *vice versa*, it can be then shown that $\mathrm{E} \star \mathrm{F} = 1$. For more information about quantum combs and link product see Refs. [1–3].

Appendix C: Uncorrelated noise bound

The bound (4) has been derived in Ref. [4], and the asymptotic bounds (5) and (6) are its direct consequences, see equations (11), (12) from Ref. [4]. In what follows, we provide a simpler derivation of (4). From the new derivation it is clear, that the bound generated by (7) for uncorrelated noise is at least as tight as the one given by (4). The advantage of the old derivation is that it can be easily generalized to strategies involving causal superpositions.

1. Simplified derivation

We consider an adaptive (AD) scheme in which N independent channels Λ_θ are probed (see Fig. 1) and keep the notation from the main text. Let $\rho_\theta^{(l)} \in \mathcal{L}(\mathcal{H}_{2l} \otimes \mathcal{A}_l)$ be the probe and ancilla state after lth use of Λ_θ, and let

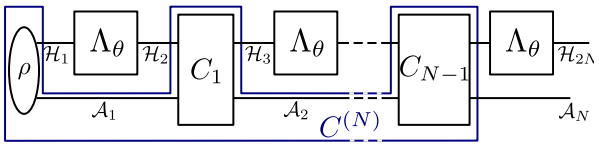


FIG. 1. Adaptive scheme for uncorrelated channels

$F(\rho_\theta^{(l)}) = F^{(l)}$. The state after the next control is $\mathrm{C}_l(\rho_\theta^{(l)})$, we construct its QFI NIP $|\Psi_\theta^{(l)}\rangle \in \mathcal{H}_{2l+1} \otimes \mathcal{A}_{l+1} \otimes \mathcal{E}_l$, where $\mathcal{E}_l$ is an artificial space added for purification purposes. Using (A3) and (A2) and the fact that a θ-independent channel cannot increase the QFI [5, p. 57], we obtain

$$4\,\langle\dot\Psi_\theta^{(l)}|\dot\Psi_\theta^{(l)}\rangle = F(\mathrm{C}_l(\rho_\theta^{(l)})) \le F^{(l)}. \tag{C1}$$

Let $\bar K_{k,\theta} = K_{k,\theta} \otimes \mathbb{1}_{\mathcal{A}_{l+1}\otimes\mathcal{E}_l}$, the state of probe and ancilla after the action of the $(l+1)$th channel can be written as

$$\rho_\theta^{(l+1)} = \mathrm{Tr}_{\mathcal{E}_l}\left(\sum_k \bar K_{k,\theta}\,|\Psi_\theta^{(l)}\rangle\,\langle\Psi_\theta^{(l)}|\,\bar K_{k,\theta}^\dagger\right).$$

Since the QFI of a subsystem is smaller or equal than the QFI of the whole system (i) and the QFI of a purification is larger or equal than the QFI of a purified state (ii), we obtain

$$F^{(l+1)} \overset{\text{(i)}}{\le} F\left(\sum_k \bar K_{k,\theta}\,|\Psi_\theta^{(l)}\rangle\,\langle\Psi_\theta^{(l)}|\,\bar K_{k,\theta}^\dagger\right) \overset{\text{(ii)}}{\le} F\left(\sum_k \bar K_{k,\theta}\,|\Psi_\theta^{(l)}\rangle \otimes |k\rangle_{\mathcal{E'}_l}\right), \tag{C2}$$

where $\mathcal{E'}_l$ is an additional space added for purification, and the vectors $|k\rangle_{\mathcal{E'}_l}$ form its o.-n. basis. After expanding the rightmost part of (C2) using (A2), we obtain

$$\begin{aligned} F^{(l+1)}/4 \le\ & \langle\dot\Psi_\theta^{(l)}|\sum_k \bar K_{\theta,k}^\dagger \bar K_{\theta,k}\,|\dot\Psi_\theta^{(l)}\rangle + \langle\dot\Psi_\theta^{(l)}|\sum_k \bar K_{\theta,k}^\dagger \dot{\bar K}_{\theta,k}\,|\Psi_\theta^{(l)}\rangle + \langle\Psi_\theta^{(l)}|\sum_k \dot{\bar K}_{\theta,k}^\dagger \bar K_{\theta,k}\,|\dot\Psi_\theta^{(l)}\rangle + \\ & + \langle\Psi_\theta^{(l)}|\sum_k \dot{\bar K}_{\theta,k}^\dagger \dot{\bar K}_{\theta,k}\,|\Psi_\theta^{(l)}\rangle \end{aligned} \tag{C3}$$

From (C1) and the identity $\sum_k \bar K_{\theta,k}^\dagger \bar K_{\theta,k} = \mathbb{1}$ we deduce that the first term is upper bounded by $F^{(l)}/4$; the last term is upper-bounded by $\|\alpha\|$ because $\langle\Psi_\theta^{(l)}|\Psi_\theta^{(l)}\rangle = 1$ and $\|A\otimes\mathbb{1}\| = \|A\|$. Both the second and the third term are upper-bounded by $\sqrt{F^{(l)}/4}\|\beta\|$, this follows from the inequality

$$\langle x|A|y\rangle \le \sqrt{\langle x|x\rangle}\|A\|\sqrt{\langle y|y\rangle},$$

which is a less general version of (7) from [4]. After taking this all together, we obtain

$$\mathcal{F}_{\mathrm{AD}}^{(l+1)} \le \mathcal{F}_{\mathrm{AD}}^{(l)} + 4\min_h\left[\|\alpha\| + \sqrt{\mathcal{F}_{\mathrm{AD}}^{(l)}}\|\beta\|\right], \tag{C4}$$

which is (4) from the main text.

To derive this bound, we assumed to have access to the QFI NIP of probe and ancilla after each control. Moreover, we maximized each term of (C3) independently—this is another factor making the bound not tight, since usually all the terms cannot be maximal at the same time. Interestingly, to derive (7), we only assumed to have access to the QFI NIP every m steps, no additional assumptions were required (see the exact derivation in the next Section). Therefore, the bound generated by (7) is guaranteed to be at least as tight as the one generated by (4) already for $m = 1$. Increasing m may tighten the bound even more.

2. Example

To illustrate the effectiveness of the new bound for uncorrelated noise, let us calculate it for $m \in \{1, 2, 3\}$ for phase estimation in the presence of amplitude damping noise, for which $K_{k,\theta} = K_k e^{-\frac{i\theta\sigma_z}{2}}$, with

$$K_1 = |0\rangle\langle 0| + \sqrt{p}\,|1\rangle\langle 1|\,, \quad K_2 = \sqrt{1-p}\,|0\rangle\langle 1|\,.$$

In Fig. 2, we demonstrate the results for $p = 0.5$. The newly introduced bound is tighter than the old one already for $m = 1$. For larger m, the bound becomes even tighter, and very close to the optimal QFI calculated exactly for $m \leq 4$ using the algorithm from Ref. [6].

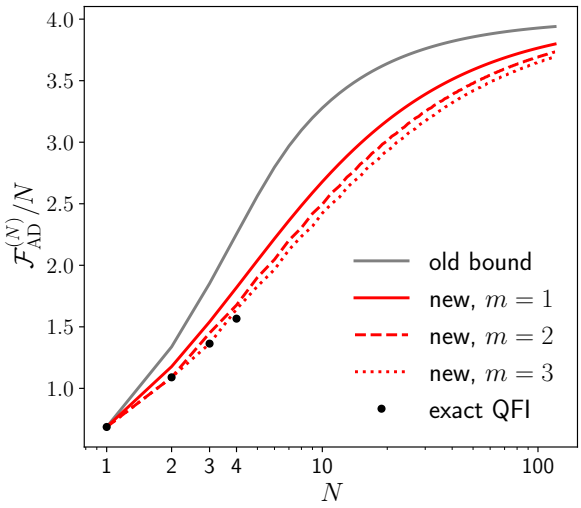


FIG. 2. Precision bounds for phase estimation in the presence of amplitude damping noise.

Appendix D: Correlated noise bounds

In this Section, we provide a detailed derivation of the recursive bounds (7) and (8) valid for correlated noise. Then, we derive the asymptotic bounds (9) and (10) using (8). Finally, we demonstrate how to formulate these new recursive and asymptotic bounds as SDPs.

1. Derivation of recursive bound (7)

Consider the joint state of the system, ancilla, and environment after l teeth of the comb $\Lambda^{(N)}_\theta$: $\rho^{(l)}_\theta \in \mathcal{H}_{2l} \otimes \mathcal{R}_l \otimes \mathcal{A}_l$, see Fig. 1 in the main text. To upper bound $F(\rho^{(l+m)}_\theta)$ in terms of $F(\rho^{(l)}_\theta)$, we replace $\rho^{(l)}_\theta$ with its QFI NIP $|\Psi^{(l)}_\theta\rangle$. This substitution can only increase the resulting QFI, since any operation possible on the original state is also possible on its purification. Next, we derive an upper bound on the QFI of the state obtained by evolving $|\Psi^{(l)}_\theta\rangle$ through the next m teeth, represented by $\overleftrightarrow{\Lambda}^{(m)}_\theta$. Arbitrary adaptive control is allowed during this evolution. To calculate this

upper bound, let us construct a comb $\Upsilon_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}| \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}$, such that $|\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}|$ is its first tooth (with trivial input), see Fig. 3.

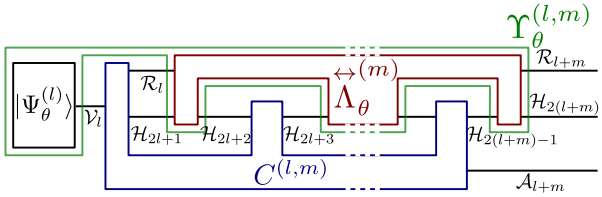


FIG. 3. The scheme of comb $\Upsilon_\theta^{(l,m)}$ and construction of the output state $\rho_\theta^{(l+m)}$

Then, the total output state of $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m} \otimes \mathcal{A}_{l+m}$ can be written as $\mathrm{C}^{(l,m)} \star \Upsilon_\theta^{(l,m)}$, where $\mathrm{C}^{(l,m)}$ is an arbitrary comb representing adaptive control, $\mathrm{C}^{(l,m)}$ acts on the spaces depicted in Fig. 3. Notice, that $C^{(l,m)}$ acts also on environmental subspace $\mathcal{R}_l$, which is impossible in a normal AD scheme. Therefore, the QFI of $\mathrm{C}^{(l,m)} \star \Upsilon_\theta^{(l,m)}$ might be larger than $F(\rho_\theta^{(l+m)})$, moreover $F(\rho_\theta^{(l+m)}) \leq \max_{\mathrm{C}^{(l,m)}} F(\mathrm{C}^{(l,m)} \star \Upsilon_\theta^{(l,m)})$ because all types of controls allowed in a normal AD scheme can be simulated by a proper choice of $\mathrm{C}^{(l,m)}$. These "environmental leakages" that happen every m teeth are one of the reasons for the potential looseness of the bound.

We are now nearly ready to compute the upper bound for $F(\rho_\theta^{(l+m)})$, expressed as the comb QFI of $\Upsilon_\theta^{(l,m)}$, which can be evaluated for any comb using the results of Refs. [6, 7]. To fully specify $\Upsilon_\theta^{(l,m)}$, we define the basis of $\mathcal{V}_l = \mathrm{span}(|\Psi_\theta^{(l)}\rangle, |\dot{\Psi}_\theta^{(l)}\rangle)$:

$$|0\rangle = |\Psi_\theta^{(l)}\rangle, \quad |1\rangle = 2\,|\dot{\Psi}_\theta^{(l)}\rangle / \sqrt{F(\rho_\theta^{(l)})}. \tag{D1}$$

This basis is orthonormal, which follows directly from quantum state normalization and from properties of the QFI NIP (A3). The density matrices of the state and its derivative can be expressed in this basis as

$$|\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \frac{\mathrm{d}}{\mathrm{d}\theta}\left(|\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}|\right) = \frac{\sqrt{F(\rho_\theta^{(l)})}}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Finally, we can explicitly write the bound for $F(\rho_\theta^{(l+m)})$ as $F(\rho_\theta^{(l+m)}) \leq F^{(l+m)}$, where $F^{(km)}$ is defined by the following iteration

$$F^{(0)} = 0, \quad F^{(l+m)} = \max_{C^{(l,m)}} F(\mathrm{C}^{(l,m)} \star \Upsilon_\theta^{(l,m)}), \quad \Upsilon_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}, \quad \dot{\Upsilon}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overleftrightarrow{\dot{\Lambda}}_\theta^{(m)} + \frac{\sqrt{F^{(l)}}}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}, \tag{D2}$$

where $F(\rho_\theta^{(l)})$ was changed to its upper bound $F^{(l)}$ in the expression for $\frac{\mathrm{d}}{\mathrm{d}\theta}\left(|\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}|\right)$.

The maximization over $\mathrm{C}^{(l,m)}$ can be performed using the SDP described in Refs. [6, 7]. The iteration D2 gives an upper bound for $F(\rho_\theta^{(l)})$ valid for all control combs C, so it is also an upper bound for $\mathcal{F}_{\mathrm{AD}}^{(l)}$. Interestingly, for uncorrelated cases there are no "environmental leakages", so the only reason for the lack of tightness of the bound is that the state $\rho_\theta^{(l)}$ is replaced with its QFI NIP every m steps. This makes the bound (D2) at least as tight as (C4) for uncorrelated cases.

2. Derivation of recursive bound (8)

Let $\overleftrightarrow{\Lambda}_\theta^{(m)} = \sum_k |K_{k,\theta}^{(m)}\rangle\langle K_{k,\theta}^{(m)}|$, since $\Upsilon_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle\langle\Psi_\theta^{(l)}| \otimes \overleftrightarrow{\Lambda}_\theta^{(m)}$, we can decompose

$$\Upsilon_\theta^{(l,m)} = \sum_k |L_{k,\theta}^{(l,m)}\rangle\langle L_{k,\theta}^{(l,m)}|, \tag{D3}$$

where $|L_{k,\theta}^{(l,m)}\rangle = |\Psi_\theta^{(l)}\rangle \otimes |K_{k,\theta}^{(m)}\rangle$. Using Leibniz rule and (D1) we obtain

$$|L_{k,\theta}^{(l,m)}\rangle = |0\rangle \otimes |K_{k,\theta}^{(m)}\rangle, \quad |\dot{L}_{k,\theta}^{(l,m)}\rangle = |0\rangle \otimes |\dot{K}_{k,\theta}^{(m)}\rangle + \sqrt{F^{(l)}}/2\,|1\rangle \otimes |K_{k,\theta}^{(m)}\rangle. \tag{D4}$$

Let $|\dot{\tilde{K}}^{(m)}_{k,\theta}\rangle = |\dot{K}^{(m)}_{k,\theta}\rangle - ih_{kk'}\,|K^{(m)}_{k',\theta}\rangle$, where h is a hermitian matrix generating equivalent Kraus representations of a channel; the summation is performed over repeated indices. Then, $|\dot{\tilde{L}}^{(l,m)}_{k,\theta}\rangle = |0\rangle \otimes |\dot{\tilde{K}}^{(m)}_{k,\theta}\rangle + \sqrt{F^{(l)}}/2\,|1\rangle \otimes |K^{(m)}_{k,\theta}\rangle$, where $|\dot{\tilde{L}}^{(l,m)}_{k,\theta}\rangle = |\dot{L}^{(l,m)}_{k,\theta}\rangle - ih_{kk'}\,|L^{(l,m)}_{k',\theta}\rangle$. The performance operator of $\overset{\leftrightarrow}{\Lambda}{}^{(m)}_{\theta}$ is $\boldsymbol{\alpha}^{(m)} = \mathrm{Tr}_{\mathrm{out}}\left(\sum_k |\dot{\tilde{K}}^{(m)}_{k,\theta}\rangle\langle\dot{\tilde{K}}^{(m)}_{k,\theta}|\right)$, the performance operator of $\Upsilon^{(l,m)}_{\theta}$ is $\boldsymbol{\alpha}^{(l,m)}_{\Upsilon} = \mathrm{Tr}_{\mathrm{out}}\left(\sum_k |\dot{\tilde{L}}^{(l,m)}_{k,\theta}\rangle\langle\dot{\tilde{L}}^{(l,m)}_{k,\theta}|\right)$, here $\mathrm{Tr}_{\mathrm{out}}$ is the partial trace over the subspace $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m}$, which is the last output of $\overset{\leftrightarrow}{\Lambda}{}^{(m)}_{\theta}$ and of $\Upsilon^{(l,m)}_{\theta}$. Using (D4), we get :

$$\boldsymbol{\alpha}^{(l,m)}_{\Upsilon} = |0\rangle\langle 0| \otimes \boldsymbol{\alpha}^{(m)} + \frac{1}{2}\sqrt{F^{(l)}}\,|0\rangle\langle 1| \otimes \boldsymbol{\beta}^{(m)} + \frac{1}{2}\sqrt{F^{(l)}}\,|1\rangle\langle 0| \otimes \boldsymbol{\beta}^{(m)\dagger} + \frac{1}{4}F^{(l)}\,|1\rangle\langle 1| \otimes \tilde{\Lambda}^{(m)}, \tag{D5}$$

where $\tilde{\Lambda}^{(m)} = \mathrm{Tr}_{\mathrm{out}}\overset{\leftrightarrow}{\Lambda}{}^{(m)}_{\theta}$, $\boldsymbol{\beta}^{(m)} = \mathrm{Tr}_{\mathrm{out}}\left(\sum_k |\dot{\tilde{K}}^{(m)}_{k,\theta}\rangle\langle K^{(m)}_{k,\theta}|\right)$. After writing the maximization in (D2) using the expression for the comb QFI (3) in terms of the performance operator, we get

$$F^{(l+m)} = \min_h \max_{\tilde{C}^{(m,l)}} 4\mathrm{Tr}\left(\boldsymbol{\alpha}^{(l,m)}_{\Upsilon}\tilde{C}^{(l,m)}\right),$$

where $\tilde{C}^{(l,m)} = \mathrm{Tr}_{\mathcal{A}_{l+m}}C^{(l,m)}$. After decomposing $\tilde{C}^{(l,m)} = \sum_{i,j=0}^{1}|i\rangle_{\mathcal{V}_l}\langle j| \otimes \tilde{C}^{(l,m)}_{ij}$, using (D5) and the normalization condition $\mathrm{Tr}(\tilde{C}^{(l,m)}_{11}\tilde{\Lambda}^{(m)}) = \tilde{\mathrm{C}}^{(l,m)}_{11} \star \Lambda^{(m)} = 1$, we obtain

$$4\mathrm{Tr}\left(\boldsymbol{\alpha}^{(l,m)}_{\Upsilon}\tilde{C}^{(l,m)}\right) = F^{(l)} + 4\mathrm{Tr}\left(\boldsymbol{\alpha}^{(m)}\tilde{C}^{(l,m)}_{00}\right) + 4\sqrt{F^{(l)}}\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}^{(l,m)}_{10}\right)\right),$$

and consequently

$$F^{(l+m)} = F^{(l)} + 4\min_h \max_{\tilde{C}^{(l,m)}}\left[\mathrm{Tr}\left(\boldsymbol{\alpha}^{(m)}\tilde{C}^{(l,m)}_{00}\right) + \sqrt{F^{(l)}}\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}^{(l,m)}_{10}\right)\right)\right] \tag{D6}$$

The normalization condition was derived using the fact that $\tilde{\mathrm{C}}^{(l,m)}_{11} \in \mathrm{Comb}[(\emptyset, \mathcal{H}_{2l+1} \otimes \mathcal{R}_l), (\mathcal{H}_{2l+2}, \mathcal{H}_{2l+3}), ..., (\mathcal{H}_{2(l+m)-2}, \mathcal{H}_{2(l+m)-1})]$ and $\tilde{\Lambda}^{(m)} \in \mathrm{Comb}[(\mathcal{H}_{2l+1}\otimes\mathcal{R}_l, \mathcal{H}_{2l+2}), (\mathcal{H}_{2l+3}, \mathcal{H}_{2l+4}), ..., (\mathcal{H}_{2(l+m)-1}, \emptyset)]$, so all outputs of $\tilde{\mathrm{C}}^{(l,m)}_{11}$ are inputs of $\tilde{\Lambda}^{(m)}$ and *vice versa*.

To further simplify the bound, we maximize each term of the RHS of (D6) independently and use the inequality between the maximum of a sum and the sum of maxima, which leads to (8) from the main text:

$$F^{(l+m)} \leq F^{(l)} + 4\min_h\left[a^{(m)} + \sqrt{F^{(l)}}b^{(m)}\right], a^{(m)} = \max_{\tilde{C}^{(m)}}\mathrm{Tr}\left(\boldsymbol{\alpha}^{(m)}\tilde{C}^{(m)}_{00}\right), \quad b^{(m)} = \max_{\tilde{C}^{(m)}}\mathrm{Re}\ \mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}^{(m)}_{10}\right). \tag{D7}$$

The bound (D7) is less tight than (D2), but, as we will prove in the next section, there is no difference between bounds generated by (D7) and (D2) asymptotically (for large N).

For simplicity, starting from (D7), we replace $\tilde{C}^{(l,m)}$ with $\tilde{C}^{(m)}$ under maximization, since $\tilde{C}^{(l,m)}$ has the same structure regardless of l . This is possible, when subsequent teeth of a comb $\overset{\leftrightarrow}{\Lambda}_{\theta}$ are the same, as we assumed from the very beginning in the main text. Notice that this assumption is not necessary if we just want to derive recursive bounds as in (D2), but it is important to derive an asymptotic bound. For simplicity and for the sake of deriving an asymptotic bound, let us also adopt convention $l = 0$ for naming of spaces $\mathcal{H}$, $\mathcal{R}$, $\mathcal{A}$. Then we have $\tilde{C}^{(m)} \in \mathrm{Comb}[(\mathcal{V}_0, \mathcal{R}_0 \otimes \mathcal{H}_1), (\mathcal{H}_2, \mathcal{H}_3), ..., (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1} \otimes \mathcal{A}_m)]$, see Fig. 3 for a comparison.

3. Derivation of asymptotic bounds (9) and (10)

To derive the asymptotic bounds we will use the following lemma.

Lemma 1 Let $x^{(n)}$ be a sequence of real numbers satisfying $x^{(n+1)} = x^{(n)} + A + 2B\sqrt{x^{(n)}}$ for any integer n, $x^{(0)} = 0$, $A \geq B^2$. Then

$$\lim_{n\to\infty} x^{(n)}/n = A \ \text{ for } \ B = 0; \quad \lim_{n\to\infty} x^{(n)}/n^2 = B^2 \ \text{ for } \ B \geq 0.$$

Proof When $B = 0$, then $x^{(n)} = An$, so the first part of lemma is obvious. When $B > 0$, then for any $n \geq 1$

$$x^{(n)} \leq f(n) = An + B^2n(n-1) + (A - B^2)n\log n, \tag{D8}$$

which is proven in the Supplementary Material of Ref. [4], Appendix E, in which $\|\alpha\|$ and $\|\beta\|$ play the role of A and B. Let us now prove that for any $n \geq 1$

$$x^{(n)} \geq B^2 n^2. \tag{D9}$$

This can be shown by induction, we have $x^{(1)} = A \geq B^2 1^2$, and when (D9) holds for some n, then for $n+1$ we have

$$x^{(n+1)} \geq B^2 n^2 + 2B^2 n + A \geq B^2 n^2 + 2B^2 n + B^2 = B^2 (n+1)^2.$$

From (D8) , (D9), and the fact that $\lim_{n\to\infty} f(n)/n^2 = B^2$ follows the second part of the lemma. ■

The asymptotic bounds for uncorrelated noise (5) and (6) are direct consequences of the iterative bound (4) and Lemma 1—notice that asymptotically (for large $\mathcal{F}_{\text{AD}}^{(i)}$) the minimum over h is always achieved when $\|\beta\|$ is minimal, since it is multiplied by $\sqrt{\mathcal{F}_{\text{AD}}^{(i)}}$ and consequently eventually dominates over the term with $\|\alpha\|$, unless $\|\beta\| = 0$.

The assumptions of Lemma 1 are true for the uncorrelated case because of the inequality $\|\alpha\| \geq \|\beta\|^2$. Let us show that the analogous inequality holds for the correlated case as well.

Lemma 2 $a^{(m)} \geq {b^{(m)}}^2$

Proof Let us use the notation $\tilde{\text{C}}^{\text{A}}$, $\tilde{\text{C}}^{\text{B}}$ for combs $\tilde{\text{C}}^{(m)}$ maximizing $a^{(m)}$ and $b^{(m)}$ respectively. Because $a^{(m)} = \text{Tr}\left(\boldsymbol{\alpha}^{(m)} \tilde{C}_{00}^{\text{A}}\right) \geq \text{Tr}\left(\boldsymbol{\alpha}^{(m)} \tilde{C}_{00}^{\text{B}}\right)$ and ${b^{(m)}}^2 \leq \left|\text{Tr}\left(\boldsymbol{\beta}^{(m)} \tilde{C}_{10}^{\text{B}}\right)\right|^2$, the inequality

$$\text{Tr}\left(\boldsymbol{\alpha}^{(m)} \tilde{C}_{00}^{\text{B}}\right) \geq \left|\text{Tr}\left(\boldsymbol{\beta}^{(m)} \tilde{C}_{10}^{\text{B}}\right)\right|^2$$

implies the thesis of our lemma. To prove this inequality, notice that

$$\begin{aligned}
\left|\text{Tr}\left(\boldsymbol{\beta}^{(m)} \tilde{C}_{10}^{\text{B}}\right)\right| &= \left|\text{Tr}\left(\text{Tr}_{\text{out}}\left(\sum_k |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle \langle K_{k,\theta}^{(m)}|\right) \tilde{C}_{10}^{\text{B}}\right)\right| = \left|\sum_k \langle K_{k,\theta}^{(m)}| \tilde{C}_{10}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle\right| = \\
&= \left|\sum_k \langle K_{k,\theta}^{(m)}| \left(\tilde{C}_{11}^{\text{B}}\right)^{\frac{1}{2}} \left(\tilde{C}_{11}^{\text{B}}\right)^{-\frac{1}{2}} \tilde{C}_{10}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle\right| \leq \\
&\overset{\text{(i)}}{\leq} \sqrt{\sum_k \langle K_{k,\theta}^{(m)}| \tilde{C}_{11}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |K_{k,\theta}^{(m)}\rangle} \cdot \sqrt{\sum_k \langle \dot{\tilde{K}}_{k,\theta}^{(m)}| \tilde{C}_{10}^{\text{B}\dagger} \left(\tilde{C}_{11}^{\text{B}}\right)^{-1} \tilde{C}_{10}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle} \overset{\text{(ii)}}{\leq} \sqrt{\sum_k \langle \dot{\tilde{K}}_{k,\theta}^{(m)}| \tilde{C}_{00}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle} = \\
&= \sqrt{\text{Tr}\left(\boldsymbol{\alpha}^{(m)} \tilde{C}_{00}^{\text{B}}\right)}
\end{aligned}$$

In (i) we used Cauchy-Schwarz inequality $|\sum_k \langle x_k|y_k\rangle| \leq \sqrt{\sum_k \langle x_k|x_k\rangle}\sqrt{\sum_k \langle y_k|y_k\rangle}$ with $|x_k\rangle = \left(\tilde{C}_{11}^{\text{B}}\right)^{\frac{1}{2}} \otimes \mathbb{1}_{\text{out}} |K_{k,\theta}^{(m)}\rangle$, $|y_k\rangle = \left(\tilde{C}_{11}^{\text{B}}\right)^{-\frac{1}{2}} \tilde{C}_{10}^{\text{B}} \otimes \mathbb{1}_{\text{out}} |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle$. In (ii) we used the fact that the expression under the first square root is $\text{Tr}\left(\overset{\leftrightarrow}{\Lambda}_{\theta}^{(m)} \tilde{C}_{11}^{\text{B}}\right) = 1$, since $\overset{\leftrightarrow}{\Lambda}_{\theta}^{(m)}$ and $\tilde{C}_{11}^{\text{B}}$ are two combs whose inputs and outputs are compatible. We also used the fact that the matrix $\tilde{C}^{B} = \begin{bmatrix} \tilde{C}_{00}^{\text{B}} & \tilde{C}_{01}^{\text{B}} \\ \tilde{C}_{10}^{\text{B}} & \tilde{C}_{11}^{\text{B}} \end{bmatrix}$ is hermitian and positive-semidefinite, which implies that $\tilde{C}_{01}^{\text{B}} = \tilde{C}_{10}^{\text{B}\dagger}$, and due to Schur's complement condition $\tilde{C}_{00}^{\text{B}} \succeq \tilde{C}_{10}^{\text{B}\dagger} \left(\tilde{C}_{11}^{\text{B}}\right)^{-1} \tilde{C}_{10}^{\text{B}}$, which means that $\langle u|\tilde{C}_{00}^{\text{B}} \otimes \mathbb{1}|u\rangle \geq \langle u|\tilde{C}_{10}^{\text{B}\dagger} \left(\tilde{C}_{11}^{\text{B}}\right)^{-1} \tilde{C}_{10}^{\text{B}} \otimes \mathbb{1}|u\rangle$ for any $|u\rangle$. ■

If we perform the minimization over h in (D7) for $F^{(l)} \to \infty$, then the term with $b^{(m)}$ dominates the term with $a^{(m)}$. When there is an h for which $b^{(m)} = 0$, then for large enough l the minimum over h in recursive steps is achieved when $b^{(m)} = 0$, and, according to Lemma 1 applied for a sequence $F^{(0)}, F^{(m)}, F^{(2m)}, ...$, we get

$$\lim_{N\to\infty} \frac{\mathcal{F}_{\text{AD}}^{(N)}}{N} \leq \frac{4}{m} \min_h a^{(m)} \text{ s.t. } b^{(m)} = 0, \tag{D10}$$

which is (9) from the main text. When there is no such h, in the large l limit the minimum over h in (D7) is achieved when $b^{(m)}$ is minimal. Therefore, using Lemmas 1 and 2 for a sequence $F^{(0)}, F^{(m)}, F^{(2m)}, ...$ we get

$$\lim_{N\to\infty} \frac{\mathcal{F}_{\mathrm{AD}}^{(N)}}{N^2} \leq \frac{4}{m^2} \min_h {b^{(m)}}^2. \tag{D11}$$

Notably, according to Lemma 1, we did not loose any tightness while going from the recursion (D7) to the asymptotic bounds (D10) and (D11). However, the tightness could be potentially lost when we replaced maximum of sum in (D6) with sum of maxima in (D7). Let us now demonstrate that this has no effect asymptotically, so after iterating (D2) many times we get an asymptotic behavior predicted by (D10) or (D11). To show this, let us consider two cases:

1. Heisenberg scaling is possible. Then, for $F^{(l)} \to \infty$ the term $\sqrt{F^{(l)}}\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}_{10}\right)\right)$ dominates over the term $\mathrm{Tr}\left(\boldsymbol{\alpha}^{(m)}\tilde{C}_{00}\right)$ in the RHS of (D6) . When we pick $\tilde{C}^{(l+m)}$ such that $\tilde{C}_{10}$ maximizes the dominating term, then the result of a maximization over $\tilde{C}^{(l+m)}$ is arbitrarily close to the result of independent maximization over $\tilde{C}^A$, $\tilde{C}^B$ for $F^{(l)} \to \infty$ (the ratio between the two results $\to 1$).

2. Heisenberg scaling is not possible. Then, to perform minimization over h in (D6) for $F^{(l)} \to \infty$, we should choose h for which $\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}_{10}\right)\right) = 0$ for any $\tilde{C}_{10}$. Then, we can choose $\tilde{C}_{10} = 0$ without affecting the result. This makes the condition $\tilde{C}^{(m)} \succeq 0$ equivalent to $\tilde{C}_{00}, \tilde{C}_{11} \succeq 0$, and the maximization over $\tilde{C}^{(m)}$ of the sum of terms is equivalent to maximization of the first term over $\tilde{C}_{00}$.

4. Simpler form of the asymptotic bound for standard scaling

The asymptotic bound (D10) is expressed as a minimization over all hermitian matrices h for which the constraint $b^{(m)} = 0$ is satisfied. This constraint is hard to deal with because $b^{(m)}$ is defined as a result of a non-trivial maximization, see (D7)—let us therefore try to formulate it in a more manageable way.

Let $\mathcal{X}_m$ be a linear subspace of space $\mathcal{X}_m^{\mathrm{tot}} = \mathcal{L}(\mathcal{R}_0 \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_{2m-1})$ such that $X \in \mathcal{X}_m$ iff

1. $\mathrm{Tr}(X) = 0$,

2. There exists a sequence of operators $X^{(1)}, X^{(2)}, ..., X^{(m-1)}, X^{(m)}$ for which $X = X^{(m)}$, $\forall_{2\leq k\leq m} \mathrm{Tr}_{\mathcal{H}_{2k-1}} X^{(k)} = X^{(k-1)} \otimes \mathbb{1}_{\mathcal{H}_{2k-2}}$, $X^{(1)} \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{R}_0)$.

Notice that condition 2. is similar to the linear comb conditions, a space $\mathcal{X}_m$ has a similar structure to the set of CJ operators of combs, but without positivity constraint and with zero trace. Let us also define $\mathcal{X}_m^\perp$ as an orthogonal complement of $\mathcal{X}_m$ in the space $\mathcal{X}_m^{\mathrm{tot}}$ with respect to Hilbert-Schmidt scalar product $(A|B) = \mathrm{Tr}(AB^\dagger)$. Since $\boldsymbol{\beta}^{(m)} \in \mathcal{X}_m^{\mathrm{tot}}$, we can uniquely decompose it as $\boldsymbol{\beta}^{(m)} = \boldsymbol{\beta}_1^{(m)} + \boldsymbol{\beta}_2^{(m)}$, where $\boldsymbol{\beta}_1^{(m)} \in \mathcal{X}_m$, $\boldsymbol{\beta}_2^{(m)} \in \mathcal{X}_m^\perp$.

We are now ready to prove the following statement.

Lemma 3 $b^{(m)} \geq 0$, moreover $b^{(m)} = 0$ iff $\boldsymbol{\beta}_1^{(m)} = 0$.

Proof The comb conditions for $\tilde{C}^{(m)}$ are equivalent to the following set of conditions for its blocks $\tilde{C}_{ij}^{(m)}$: (i) $\tilde{C}_{00}^{(m)}, \tilde{C}_{11}^{(m)} \in \mathrm{Comb}[(\emptyset, \mathcal{H}_1 \otimes \mathcal{R}_0), (\mathcal{H}_2, \mathcal{H}_3), ..., (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1})]$; (ii) $\tilde{C}_{01}^{(m)}, \tilde{C}_{10}^{(m)} \in \mathcal{X}_m$; (iii) $\tilde{C}_{01}^{(m)} = \tilde{C}_{10}^{(m)\dagger}$ and $\tilde{C}_{00}^{(m)} \succeq \tilde{C}_{10}^{(m)\dagger}\left(\tilde{C}_{11}^{(m)}\right)^{-1}\tilde{C}_{10}^{(m)}$. When $\tilde{C}_{11}^{(m)}$ is singular, then its inverse should be replaced with pseudo inverse. Conditions (i) and (ii) are a consequence of the linear constraints for the comb $\tilde{C}^{(m)}$, and condition (iii) is a consequence of the positivity constraint and Schur's complement condition. If we choose any $\tilde{C}_{00}^{(m)}$, $\tilde{C}_{11}^{(m)}$ satisfying (i) and set $\tilde{C}_{01}^{(m)} = \tilde{C}_{10}^{(m)} = 0$, then conditions (ii)-(iii) are also satisfied and $\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}_{10}^{(m)}\right) = 0$ which proves that $b^{(m)} \geq 0$. Since for all cases $\tilde{C}_{10}^{(m)\dagger} \in \mathcal{X}_m$ and $\boldsymbol{\beta}_2^{(m)} \in \mathcal{X}_m^\perp$, we have $\mathrm{Tr}\left(\tilde{C}_{10}^{(m)}\boldsymbol{\beta}_2^{(m)}\right) = 0$, so $b^{(m)} = 0$ when $\boldsymbol{\beta}_1^{(m)} = 0$. Let us now assume that $\boldsymbol{\beta}_1^{(m)} \neq 0$, and fix strictly positive definite $\tilde{C}_{00}^{(m)}$ and $\tilde{C}_{11}^{(m)}$ satisfying condition (i), and $\tilde{C}_{01}^{(m)} = \epsilon\boldsymbol{\beta}_1^{(m)}$, $\tilde{C}_{10}^{(m)} = \epsilon\boldsymbol{\beta}_1^{(m)\dagger}$. Then, condition (ii) is satisfied since $\boldsymbol{\beta}_1^{(m)} \in \mathcal{X}_m$, and for small enough $\epsilon > 0$ condition (iii) is also satisfied since $\tilde{C}_{00}^{(m)}$ is strictly positive. Therefore, there exists a comb $\tilde{C}^{(m)}$ for which $\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}^{(m)}\tilde{C}_{10}^{(m)}\right)\right) = \epsilon\mathrm{Re}\left(\mathrm{Tr}\left(\boldsymbol{\beta}_1^{(m)}\boldsymbol{\beta}_1^{(m)\dagger}\right)\right) > 0$, so $b^{(m)} > 0$, which finishes the proof of the second part of the lemma. ■

According to Lemma 3 the condition $b^{(m)}=0$ can be replaced with condition $\boldsymbol{\beta}_1^{(m)}=0$, or, equivalently, with the condition $\boldsymbol{\beta}^{(m)}\in\mathcal{X}_m^{\perp}$.

5. Formulations as SDPs

To formulate the bounds as SDPs, we will use the following Lemma.

Lemma 4 Let us consider the following primal SDP maximization problem:

$$\begin{array}{ll}\max_C \mathrm{Tr}(AC) & \\ \text{s.t.} & C\in \mathrm{Comb}[(\mathcal{K}_1,\mathcal{K}_2),...(\mathcal{K}_{2N-1},\mathcal{K}_{2N})]\end{array},$$

A is a hermitian matrix, here Comb[...] is the set of CJ operators of combs. Then the dual problem is

$$\begin{array}{ll}\min_{Q^{(1)},Q^{(2)},...,Q^{(N)}} \mathrm{Tr}\left(Q^{(1)}\right) & \\ \text{s.t.} & Q^{(N)}\otimes \mathbb{1}_{2N}\succeq A\\ & \mathrm{Tr}_{2k-1}Q^{(k)}=Q^{(k-1)}\otimes\mathbb{1}_{2k-2}\quad \text{for}\quad k\in\{2,3,...,N\}\end{array},\tag{D12}$$

where $Q^{(k)}$ are hermitian matrices, $Q^{(k)}\in\mathcal{L}(\mathcal{K}_1\otimes\mathcal{K}_2\otimes...\otimes\mathcal{K}_{2k-1})$. The optimal value of the dual problem is equal to optimal value of the primal problem (strong duality).

Proof A very similar statement was proven in the Supplementary Material of Ref. [6], see Section I.B, Lemmas 5 and 6. There, A was assumed to be equal to performance operator, but this assumption was not used, and the proof remains correct for any hermitian A. Moreover, the proof provided in Ref. [6] can be only directly applied to cases when $\mathcal{K}_1$ is a trivial one-dimensional space $\emptyset$. To generalize the proof for a general case, it is enough to apply it for a comb with an additional artificial empty teeth, so we consider $C\in\mathrm{Comb}[(\emptyset,\emptyset),(\mathcal{K}_1,\mathcal{K}_2),...,(\mathcal{K}_{2N-1},\mathcal{K}_{2N})]$ instead of $C\in\mathrm{Comb}[(\mathcal{K}_1,\mathcal{K}_2),...,(\mathcal{K}_{2N-1},\mathcal{K}_{2N})]$. Then, we end up with a dual problem in the form (D12).■

The iterative bound (D2) can be formulated as an SDP using Algorithm 1 from Ref. [6] directly applied for parameter-dependent comb $\Upsilon_\theta^{(l,m)}$. Let us use the decomposition (D3), (D4), and define $|\dot{c}_{k,j}^{(l,m)}(h)\rangle = {}_{\mathcal{H}_{2(l+m)}\otimes\mathcal{R}_{l+m}}\langle j|\dot{\tilde{L}}_{k,\theta}^{(l,m)}\rangle$, where $|j\rangle$ is some orthonormal basis of $\mathcal{H}_{2(l+m)}\otimes\mathcal{R}_{l+m}$ (notice that $|\dot{c}_{k,j}^{(l,m)}(h)\rangle$ depend linearly on the mixing hermitian matrix h). The recursive step (D2) can be calculated using the following SDP:

$$\begin{aligned}F^{(l+m)}=4\min_{h,\{Q^{(k)}\}_{k\in\{1,...,m\}}}&\ \mathrm{Tr}\left(Q^{(1)}\right),\\ &\text{subject to}\\ &A\succeq 0,\\ &\forall_{2\le k\le m-1}\ \mathrm{Tr}_{\mathcal{H}_{2(l+k)}}Q^{(k+1)}=Q^{(k)}\otimes\mathbb{1}_{\mathcal{H}_{2(l+k)-1}},\\ &\mathrm{Tr}_{\mathcal{H}_{2l+2}}Q^{(2)}=Q^{(1)}\otimes\mathbb{1}_{\mathcal{H}_{2l+1}\otimes\mathcal{R}_l},\end{aligned}$$

where

$$A=\left(\begin{array}{c|c}Q^{(m)}\otimes\mathbb{1}_{\mathcal{H}_{2(l+m)-1}} & |\dot{c}_{1,1}^{(l,m)}(h)\rangle\ \ \ldots\ \ |\dot{c}_{r,d}^{(l,m)}(h)\rangle\\ \hline \langle\dot{c}_{1,1}^{(l,m)}(h)| & \\ \vdots & \mathbb{1}_{dr}\\ \langle\dot{c}_{r,d}^{(l,m)}(h)| & \end{array}\right),$$

d is the dimension of $\mathcal{H}_{2(l+m)}\otimes\mathcal{R}_{l+m}$, r is the rank of $\Upsilon_\theta^{(l,m)}$, $Q^{(1)}\in\mathcal{L}(\mathcal{V}_l)$, notice that $|\dot{c}_{k,j}^{(l,m)}(h)\rangle$ depend on $F^{(l)}$.

The asymptotic bound (D10) can be written as a similar SDP with additional constraints for h coming from condition $b^{(m)}=0$. First, let us notice that the maximization over $\tilde{C}^{(m)}$ in $a^{(m)}$ in (D10) boils down to a maximization over $\tilde{C}_{00}^{(m)}\in\mathrm{Comb}[(\emptyset,\mathcal{H}_1\otimes\mathcal{R}_0),(\mathcal{H}_2,\mathcal{H}_3),...,(\mathcal{H}_{2m-2},\mathcal{H}_{2m-1})]$. Secondly, the condition $b^{(m)}=0$ can be written as $\boldsymbol{\beta}^{(m)}\in\mathcal{X}_m^{\perp}$ (see Lemma 3 and the remark below). Since the dual affine space to comb space is another comb space (with outputs and inputs interchanged) [7, 8], it can be shown that $\boldsymbol{\beta}^{(m)}\in\mathcal{X}_m^{\perp}$ iff there exists a sequence of operators $Y^{(1)},Y^{(2)},...,Y^{(m-1)}$ for which $\boldsymbol{\beta}^{(m)}=Y^{(m-1)}\otimes\mathbb{1}_{\mathcal{H}_{2m-1}}$, $\forall_{2\le k\le m-1}\mathrm{Tr}_{\mathcal{H}_{2k}}Y^{(k)}=Y^{(k-1)}\otimes\mathbb{1}_{\mathcal{H}_{2k-1}}$, $\mathrm{Tr}_{\mathcal{H}_2}Y^{(1)}=$

$Y^{(0)}\mathbb{1}_{\mathcal{H}_1\otimes\mathcal{R}_0}$, $Y^{(0)}\in\mathbb{C}$. By introducing the notation $|c^{(m)}_{k,j}(h)\rangle = {}_{\mathcal{H}_{2m}\otimes\mathcal{R}_m}\langle j|K^{(m)}_{k,\theta}\rangle$, $|\dot{c}^{(m)}_{k,j}(h)\rangle = {}_{\mathcal{H}_{2m}\otimes\mathcal{R}_m}\langle j|\dot{K}^{(m)}_{k,\theta}\rangle$, we get

$$\boldsymbol{\beta}^{(m)}(h) = \sum_{k,j} |\dot{c}^{(m)}_{k,j}(h)\rangle \langle c^{(m)}_{k,j}|\,, \tag{D13}$$

notice that $\boldsymbol{\beta}^{(m)}(h)$ depends linearly on h. Finally, after supplementing the SDP from Ref. [6] with the condition for $\boldsymbol{\beta}^{(m)}$, we get the following SDP for the asymptotic bound:

$$\begin{aligned}
\lim_{N\to\infty} \mathcal{F}^{(N)}_{\mathrm{AD}}/N \leq 4/m \min_{h,\{Q^{(k)},Y^{(k)}\}_{k\in\{0,1,\ldots,m-1\}}} & Q^{(0)},\\
&\text{subject to}\\
&A\succeq 0,\\
&\forall_{2\leq k\leq m-1} \mathrm{Tr}_{\mathcal{H}_{2k}} Q^{(k)} = Q^{(k-1)}\otimes\mathbb{1}_{\mathcal{H}_{2k-1}},\\
&\mathrm{Tr}_{\mathcal{H}_2} Q^{(1)} = Q^{(0)}\mathbb{1}_{\mathcal{H}_1\otimes\mathcal{R}_0},\quad Q^{(0)}\in\mathbb{R},\\
&\boldsymbol{\beta}^{(m)}(h) = Y^{(m-1)}\otimes\mathbb{1}_{\mathcal{H}_{2m-1}},\\
&\forall_{2\leq k\leq m-1} \mathrm{Tr}_{\mathcal{H}_{2k}} Y^{(k)} = Y^{(k-1)}\otimes\mathbb{1}_{\mathcal{H}_{2k-1}},\\
&\mathrm{Tr}_{\mathcal{H}_2} Y^{(1)} = Y^{(0)}\mathbb{1}_{\mathcal{H}_1\otimes\mathcal{R}_0},\quad Y^{(0)}\in\mathbb{C},
\end{aligned}$$

where

$$A = \left(\begin{array}{c|c} Q^{(m-1)}\otimes\mathbb{1}_{\mathcal{H}_{2m-1}} & |\dot{c}^{(m)}_{1,1}(h)\rangle \quad \ldots \quad |\dot{c}^{(m)}_{r,d}(h)\rangle \\ \hline \begin{array}{c}\langle\dot{c}^{(m)}_{1,1}(h)|\\ \vdots \\ \langle\dot{c}^{(m)}_{r,d}(h)|\end{array} & \mathbb{1}_{dr} \end{array}\right),$$

d is dimension of $\mathcal{H}_{2m}\otimes\mathcal{R}_m$, r is the rank of $\overset{\leftrightarrow(m)}{\Lambda}_{\theta}$.

The asymptotic bound in the presence of HS (D11) can be written as

$$\lim_{N\to\infty} \mathcal{F}^{(N)}_{\mathrm{AD}}/N^2 \leq \left[2/m \min_h \max_{\tilde{C}^{(m)}\in\mathrm{Comb}[(\mathcal{V}_0,\mathcal{H}_1\otimes\mathcal{R}_0),(\mathcal{H}_2,\mathcal{H}_3),\ldots,(\mathcal{H}_{2m-2},\mathcal{H}_{2m-1})]} \mathrm{Tr}\left(\tilde{C}^{(m)}\begin{bmatrix} 0 & \frac{1}{2}\boldsymbol{\beta}^{(m)} \\ \frac{1}{2}\boldsymbol{\beta}^{(m)\dagger} & 0\end{bmatrix}\right)\right]^2,$$

where we used the block decomposition $\tilde{C}^{(m)} = \begin{bmatrix}\tilde{C}^{(m)}_{00} & \tilde{C}^{(m)}_{01}\\ \tilde{C}^{(m)}_{10} & \tilde{C}^{(m)}_{11}\end{bmatrix}$ where $\tilde{C}^{(m)\dagger}_{01} = \tilde{C}^{(m)}_{10}$. After dualizing the maximization problem using Lemma 4, we can write the whole min max problem as a single SDP minimization:

$$\begin{aligned}
\sqrt{\lim_{N\to\infty} \mathcal{F}^{(N)}_{\mathrm{AD}}/N^2} \leq 2/m \min_{h,\{Q^{(k)}\}_{k\in\{1,\ldots,m\}}} & \mathrm{Tr}\left(Q^{(1)}\right),\\
&\text{subject to}\\
&Q^{(m)}\otimes\mathbb{1}_{\mathcal{H}_{2m-1}} \succeq \begin{bmatrix} 0 & \frac{1}{2}\boldsymbol{\beta}^{(m)} \\ \frac{1}{2}\boldsymbol{\beta}^{(m)\dagger} & 0\end{bmatrix},\\
&\forall_{2\leq k\leq m-1} \mathrm{Tr}_{\mathcal{H}_{2k}} Q^{(k+1)} = Q^{(k)}\otimes\mathbb{1}_{\mathcal{H}_{2k-1}},\\
&\mathrm{Tr}_{\mathcal{H}_2} Q^{(2)} = Q^{(1)}\otimes\mathbb{1}_{\mathcal{R}_0\otimes\mathcal{H}_1},
\end{aligned}$$

where $Q^{(1)}\in\mathcal{L}(\mathcal{V}_0)$, $\boldsymbol{\beta}^{(m)}$ is given by (D13).

Appendix E: Examples

The code we used to generated upper bounds introduced in this work and tensor-network based lower bounds introduced in [9, 10] is available in a public repository [11]. To generate upper bounds, we used the function

`ad_asym_bound_correlated` from file `qmetro/bounds.py`. To generate lower bounds for correlated dephasing examples, we used tensor-network based optimization over adaptive strategies with limited ancilla dimension $d_{\mathcal{A}}$, this was done using the function `iss_tnet_adaptive_qfi` from the file `qmetro/protocols/iss.py`. We calculated lower bounds for the collisional thermometry example by considering a parallel scheme, when input state and output measurement are represented using tensor networks—the optimization was performed using the function `iss_tnet_parallel_qfi` from the file `qmetro/protocols/iss.py`.

1. Correlated dephasing

All the details related to the correlated dephasing model are described in Ref. [9] (note that here we use notation S instead of T for a mixing map). The only additional difficulty is that the tightness of an upper bound depends on the way in which we cut the chain of correlated channels into pieces. As we checked numerically, for parallel dephasing the tightest bound is obtained when we decompose mixing map as $S = \sqrt{S} \circ \sqrt{S}$, and make cuts between two maps $\sqrt{S}$, see Fig. 2(a) in the main text. The Kraus operators of channel S implementing a stochastic map $S_{i|i-1}$ on the basis $|\pm\rangle$ are $S_{sr} = \sqrt{\frac{1+srC}{2}}\,|s\rangle\langle r|$ for $s, r \in \{+, -\}$. It can be shown by direct calculations, that the square root map for $C > 0$ is

$$\sqrt{S}_{i|i-1}(r_i|r_{i-1}) = (1 + r_i r_{i-1}\sqrt{|C|})/2,$$

the corresponding Kraus operators of a quantum map are $\sqrt{S}_{sr} = \sqrt{\frac{1+sr\sqrt{|C|}}{2}}\,|s\rangle\langle r|$for $s, r \in \{+, -\}$. When $C < 0$, we should construct a map like for positive C of the same absolute value, but then additionally add a flip between elements of basis $|+\rangle$, $|-\rangle$ between two maps $\sqrt{S}$.

2. Collisional thermometry

The collisional quantum thermometry model was introduced in Ref. [12]. The thermal bath of temperature T is probed by a qubit thermometer whose hamiltonian is $H = \hbar\Omega\sigma_z/2$. The thermalization of a thermometer is described by a quantum master equation in Lindblad form

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \mathcal{L}_T(\rho) = \gamma(\bar{n}+1)\mathcal{D}(\sigma_-) + \gamma\bar{n}\mathcal{D}(\sigma_+),$$

where $\mathcal{D}(X) = X\rho X^\dagger - \frac{1}{2}\{X^\dagger X, \rho\}$, $\sigma_\pm = (\sigma_x \pm i\sigma_y)$, $\bar{n} = (e^{\hbar\Omega/k_B T} - 1)^{-1}$, γ is a coupling constant between the thermometer and the thermal bath. The channel describing the thermalization for a time τ is given by integrating the above equation

$$\mathrm{S}_T^\tau(\rho) = e^{\mathcal{L}_T \tau}(\rho),$$

which is a generalized amplitude damping channel. When $\tau \to \infty$, then the output of a channel S_T^τ is a Gibbs state $\rho_{\mathrm{th},T} = e^{-H\varepsilon/k_B T}/Z$, the QFI of this state is a thermal Fisher information $F_{\mathrm{th},T} = F(\rho_{\mathrm{th},T})$. Interestingly, the channel QFI of S_T^τ can be larger than $F_{\mathrm{th},T}$. We numerically study a case when $k_B T/\hbar\Omega = 2$, which means that $\bar{n} = 1.514$ (the same value of $\bar{n}$ was chosen in Ref. [12]). We observed that the maximal channel QFI is achieved for $\tau_{\mathrm{opt}} \approx 0.417\gamma^{-1}$, then $\mathcal{F}(\mathrm{S}_T^\tau) \approx 5.66 F_{\mathrm{th},T}$, see Fig. 5(a). Notably, one needs to entangle the input state with ancilla to achieve this optimal channel QFI.

In a collisional scheme, there is no direct access to a thermometer, one only has access to mediator qubits that exchange energy with thermometer after each thermalization process, the interaction between thermometer and mediator is governed by the unitary $V = e^{-igt(\sigma_+\otimes\sigma_- + \sigma_-\otimes\sigma_+)}$, g is a coupling constant, t is an interaction time; a full energy swap occurs for $gt = 0.5\pi$, in this work we numerically analyze the case $gt = 0.35\pi$. In Fig. 4, we present a collisional thermometry scheme, in which all mediators are prepared in a product state, but each mediator is assisted with a qubit ancilla, which can be entangled with it. Mediators interact with thermometer, and at the end, all mediators and ancillary qubits are measured. Note, that the thermometer is initialized in a thermal state, whereas in Ref. [12] the initial state of a thermometer is a stationary state of concatenation of maps V and S_T^τ. This stationary state depends on the input mediator of V, so we cannot fix a stationary state when computing universal bounds. However, the choice of initial state of thermometer does not affect the asymptotic results.

Importantly, our upper bounds cover a much broader class of strategies, in which mediators can be prepared in an arbitrary entangled state, and adaptive control between subsequent probings can be applied.

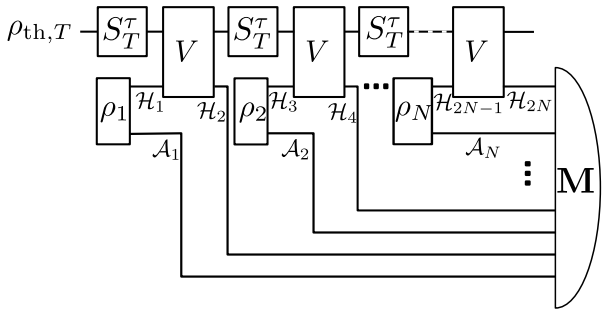


FIG. 4. Collisional quantum thermometry: thermalizing channels S_T^τ acting on thermometer are interwined with unitary channels V describing energy exchange between mediators ($\mathcal{H}$) and thermometer. At the end, the joint measurement $\mathbf{M}$ is performed on all mediators and ancillary qubits (initially entangled with mediators). Our upper bounds cover strategies with all possible entangled states of mediators and ancillas. We also calculate lower bounds for the strategy in which initially mediators are not entangled with each other, but each mediator ($\mathcal{H}_{2i-1}$) can be entangled with the corresponding ancilla ($\mathcal{A}_i$).

To complement these general upper bounds, we calculate lower bounds using tensor network techniques [10, 11], assuming that the input state of mediators is a product state, each mediator is entangled with an ancillary qubit, and that the final state of mediators and ancillas is measured using a collective measurement $\mathbf{M}$, which is described by a matrix product operator with bound dimension 4 (which means that some non-locality in a final measurement is allowed). As demonstrated in Fig. 2(c) in the main text, this strategy is optimal (up to numerical precision) for $\tau = 10\tau_{\mathrm{opt}}$, when correlations between subsequent channels are negligible. For $\tau = \tau_{\mathrm{opt}}$, there is some potential advantage from using entangled states—however, this potential advantage is relatively small.

To make the bounds as tight as possible, we need to judiciously cut the whole chain of channels V and S_T^τ ($\Lambda_\theta^{(N)}$) into smaller pieces $\overset{\leftrightarrow(m)}{\Lambda}_\theta$. More precisely, we need to find combs $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ such that their link product is $\Lambda_\theta^{(N)}$ (environmental spaces of neighbouring combs are linked). Therefore, we can concatenate $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ with some map X acting on its output environment and the inverse of this map X^{-1} acting on an input environment—then, after linking $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ with each other, each X will be concatenated with X^{-1}, so we will obtain the same resulting comb $\Lambda_\theta^{(N)}$. However, the bound tightness might be affected by the choice of X. Importantly, X must chosen such that the resulting $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ are still valid combs (in particular, they must be completely positive).

In our example we choose $\mathrm{X} = \mathrm{S}_{T_0}^{\tau_0}$, which is a map that does not differ from $\mathrm{S}_T^{\tau_0}$ at the value of temperature around which the estimation is performed, but $\frac{\mathrm{d}}{\mathrm{d}T}\mathrm{S}_{T_0}^{\tau_0} = 0$, which means that $\mathrm{S}_{T_0}^{\tau_0}$ is insensitive to the changes of T, so its channel QFI is 0. The intuition behind this choice is that such T-insensitive thermalization may hide some information that leaks from environment when we cut the whole chain of channels into pieces.

Obviously, the inverse map $\mathrm{S}_{T_0}^{-\tau_0}$ is not completely positive. However, the total map acting on an input environment is $\mathrm{S}_T^\tau \circ \mathrm{S}_{T_0}^{-\tau_0}$ is completely positive assuming that $\tau_0 < \tau$—the resulting channel at $T = T_0$ is a thermalization with time $\tau - \tau_0$, the channel is full-rank, so it will remain completely positive in the first order expansion irrespectively of its derivative.

The large values of τ_0 allow to reduce the role of environmental leakages at the outpus of $\overset{\leftrightarrow(m)}{\Lambda}_\theta$, but at the same time, extra anti-thermalization map $\mathrm{S}_{T_0}^{-\tau_0}$ at the beginning may increase the QFI. Therefore, one need to optimize over τ_0 (assuming $\tau_0 < \tau$) to get the tightest possible bound. Interestingly, for $\tau = \tau_{\mathrm{opt}}$ the optimal value is $\tau_0 = 0$, so it is not advantagous to use the described trick. The situation is completely different for $\tau = 10\tau_{\mathrm{opt}}$, see Fig. 5(b), then choosing $\tau_0 > 0$ can tighten the bound significantly. With the optimal value $\tau_0 \approx 0.62\gamma^{-1}$, we get an upper bound that coincides with lower bounds already for $m = 1$.

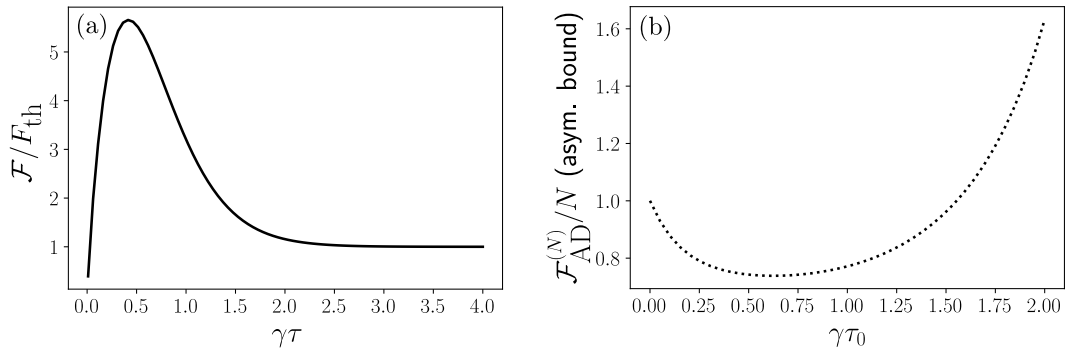


FIG. 5. (a): The ratio between channel QFI of a thermal channel S_T^{τ} and the QFI of a thermal state as a function of thermalization time multiplied by a coupling constant $\gamma\tau$. (b) The tightness of the bound is affected by the choice of time τ_0 in an inverse map $\mathrm{S}_T^{-\tau_0}$. Here we show how an asymptotic bound depends on $\gamma\tau_0$ for $\tau = 10\tau_{\text{opt}}$ $m = 1$.

[1] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, Phys. Rev. Lett. **101**, 060401 (2008).
[2] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory Effects in Quantum Channel Discrimination, Phys. Rev. Lett. **101**, 180501 (2008).
[3] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).
[4] S. Kurdziałek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, Phys. Rev. Lett. **131**, 090801 (2023).
[5] J. Kołodyński, *Precision Bounds in Noisy Quantum Metrology*, Ph.D. thesis, University of Warsaw (2014), arXiv:1409.0535.
[6] A. Altherr and Y. Yang, Quantum Metrology for Non-Markovian Processes, Phys. Rev. Lett. **127**, 060501 (2021).
[7] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal Strategies of Quantum Metrology with a Strict Hierarchy, Phys. Rev. Lett. **130**, 070803 (2023).
[8] J. Bavaresco, M. Murao, and M. T. Quintino, Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination, Phys. Rev. Lett. **127**, 200504 (2021).
[9] S. Kurdzialek, P. Dulian, J. Majsak, S. Chakraborty, and R. Demkowicz-Dobrzanski, Quantum metrology using quantum combs and tensor network formalism (2024), arXiv:2403.04854.
[10] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański, Tensor-network approach for quantum metrology in many-body quantum systems, Nat. Commun. **11**, 250 (2020).
[11] P. Dulian, S. Kurdziałek, and R. Demkowicz-Dobrzański, Qmetro++, `https://github.com/pdulian/qmetro` (2025).
[12] S. Seah, S. Nimmrichter, D. Grimmer, J. P. Santos, V. Scarani, and G. T. Landi, Collisional Quantum Thermometry, Phys. Rev. Lett. **123**, 180602 (2019).

Chapter 5

Measurement noise susceptibility in quantum estimation

Commentary

In this work, we propose a universal way to quantify how vulnerable the FI is to noise affecting the measurement. We introduce a quantity called Fisher information measurement noise susceptibility (FI MeNoS) and show how to compute it for general metrological scenarios. We thoroughly study the quantum-inspired superresolution imaging case, which was discussed in Sec. 1.5 of Chapter 1. Using FI MeNoS, we identify the QCRB-saturating measurement that is the least susceptible to noise. An alternative approach, which also addresses the general role of measurement noise in quantum estimation, has been presented in Ref. [Len22].

In Appendix F, we study the role of signal correlations arising from sub-Poissonian emitter statistics. As shown in Sec. 1.5, Chapter 1, such correlations do not increase the QFI, even though they increase the classical FI for DI measurements. However, as we demonstrate here, correlations enable construction of measurements less susceptible to noise, suggesting that non-Poissonianity is a valuable resource for imaging not only for DI, but also more fundamentally, when the measurement is optimized.

My contribution

I conceived and developed (supported by RDD) the main idea presented in this work, performed all the calculations and proofs, generated and analyzed all the numerical data, wrote the whole main text and all appendices, I am the corresponding author.

Measurement Noise Susceptibility in Quantum Estimation

Stanisław Kurdziałek* and Rafał Demkowicz-Dobrzański
Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland

(Received 11 July 2022; revised 2 December 2022; accepted 30 March 2023; published 21 April 2023)

Fisher information is a key notion in the whole field of quantum metrology. It allows for a direct quantification of the maximal achievable precision of the estimation of the parameters encoded in quantum states using the most general quantum measurement. It fails, however, to quantify the robustness of quantum estimation schemes against measurement imperfections, which are always present in any practical implementations. Here, we introduce a new concept of Fisher information measurement noise susceptibility that quantifies the potential loss of Fisher information due to small measurement disturbance. We derive an explicit formula for the quantity, and demonstrate its usefulness in the analysis of paradigmatic quantum estimation schemes, including interferometry and superresolution optical imaging.

DOI: 10.1103/PhysRevLett.130.160802

Introduction.—Noise, decoherence, and implementation imperfections are the main factors hindering the transfer of quantum enhanced technologies (e.g., quantum computing and communication) from proof-of-principle experiments to real-life applications [1]. These issues also affect the development of quantum metrology, whose goal is to utilize sophisticated properties of light and matter to enhance sensing instruments [2–5]. Quantum estimation theory [6,7] laid theoretical grounds for present-day quantum metrology—one of its greatest achievements is identification of protocols that perform optimally in the presence of noise [8–12].

One of the key elements affecting the precision of metrological protocols is the imperfect realization of the final measurement step, where information is being extracted from quantum sensors. In order to assess the effect of imperfect measurement implementation, the standard route is to characterize the type of noise present, e.g., detector dark counts, measurement output crosstalks, etc., and then analyze its impact on the relevant figures-of-merit [13]. A more systematic and general study of the effect of readout noise on the measurement precision is provided in Ref. [14].

Still, from a fundamental point of view, it would be much more advantageous to be able to determine the noise robustness of a given measurement scheme without specifying the actual form of the noise. A similar motivation lays behind measurement robustness considerations that can be found in the context of other quantum information tasks, see, e.g., [15,16], but have never been applied to quantum estimation theory. In this Letter, we focus on the most common figure of merit in quantum estimation theory—the Fisher information—and propose a quantity, Fisher information measurement noise susceptibility (FI MENOS), which characterizes the maximal relative decrease of FI due to small measurement disturbance of the most general type. This quantity allows us to obtain a deep insight into fundamental noise-robustness properties of different measurement schemes without assuming any particular noise form. We illustrate the fruitfulness of this approach by analyzing paradigmatic quantum enhanced metrological schemes including interferometry and superresolution imaging.

Quantum estimation theory preliminaries.—In a paradigmatic quantum estimation scenario, a continuous parameter θ is encoded in a state ρ_θ of a probe system with associated Hilbert space $\mathcal{H}_S$. In order to describe the process of extraction of information on the parameter θ from the state in full generality, one considers an external measuring device whose Hilbert space is $\mathcal{H}_M$. The device is initialized in a pure state $|0\rangle_M$ and the generalized measurement of ρ_θ consists of two stages: (i) interaction between S and M described by a unitary operation U_{SM} and (ii) projective measurement of the postevolution state of M, which returns an outcome $i \in \{1, 2, \ldots, K\}$ with a probability

$$p_\theta(i) = \mathrm{Tr}(\rho_\theta M_i), \qquad M_i = {}_M\langle 0|U_{SM}^\dagger|i\rangle_M \langle i|U_{SM}|0\rangle_M \tag{1}$$

where M_i are effective measurement operators acting on S—note that scalar products in the above formula are partial, they act on subsystem M only, leaving part S intact. A set $\boldsymbol{M} = \{M_i\}_i$ is called a positive operator-valued measure (POVM), where M_i satisfy (i) $\sum_i M_i = \mathbb{1}$ and (ii) $M_i \geq 0$. The set of all POVMs will be denoted as $\mathcal{M}$, so we will write $\boldsymbol{M} \in \mathcal{M}$. Different choices of U_{SM} lead to different POVMs, the number of possible outcomes is $K = \dim \mathcal{H}_M$. Each POVM can be physically implemented with the help of an appropriate choice of $\mathcal{H}_M$ and U_{SM}. The projective measurement in a basis $|i\rangle$ of $\mathcal{H}_S$ corresponds to $\boldsymbol{M}$ with $M_i = |i\rangle\langle i|$.

When N copies of ρ_θ are measured independently with the same POVM $\boldsymbol{M}$ this leads to N i.i.d. random variables

sampled from $p_\theta(i)$. According to the Cramér-Rao bound (CRB) [6,7], the mean squared error (MSE) of any (locally) unbiased estimator $\tilde{\theta}$, that estimates θ based on this data, will be lower bounded as

$$\Delta^2\tilde{\theta} \geq \frac{1}{NF_C}, \qquad F_C = \sum_i p_i l_i^2 \tag{2}$$

where F_C is the classical Fisher information (CFI), $p_i = p_\theta(i)$, and $l_i = \partial_\theta \log p_\theta(i)$ is the logarithmic derivative of $p_\theta(i)$. Intuitively, the CFI quantifies how sensitive $p_\theta(i)$ is to the change of θ—the larger l_i^2, the greater the CFI. The CRB is tight—it is always possible to find a locally unbiased estimator whose MSE saturates (2), and for $N \to \infty$ one can construct a globally unbiased CRB-saturating estimator [17].

For a fixed quantum state ρ_θ, the CFI depends only on the measurement $\boldsymbol{M}$, and in order to highlight this we will denote it as $F_C[\boldsymbol{M}]$. Combining (2) with (1) the explicit form of the CFI reads

$$F_C[\boldsymbol{M}] = \sum_i \mathrm{Tr}(\rho_\theta M_i) l_i^2, \qquad l_i = \frac{\mathrm{Tr}(\dot{\rho}_\theta M_i)}{\mathrm{Tr}(\rho_\theta M_i)}, \tag{3}$$

where the dot denotes the derivative over θ. It is natural to ask what the greatest possible CFI is for a given ρ_θ—the answer is given by the quantum Fisher information (QFI) [6,7], which is the maximum of the CFI over all POVMs $\boldsymbol{M}$, and can be computed as

$$F_Q = \max_{\boldsymbol{M}\in\mathcal{M}} F_C[\boldsymbol{M}] = \mathrm{Tr}(\rho_\theta \Lambda_\theta^2), \tag{4}$$

where Λ_θ is the symmetric logarithmic derivative matrix defined by the equation $\partial_\theta \rho_\theta = \frac{1}{2}(\rho_\theta \Lambda_\theta + \Lambda_\theta \rho_\theta)$. For a given ρ_θ, and arbitrary $\boldsymbol{M}$, the MSE of any locally unbiased estimator of θ is lower-bounded by the quantum Cramér-Rao bound (QCRB), which is similar to (2), but F_C is replaced with F_Q. The projective measurement on eigenstates of Λ_θ is always QCRB saturating (its CFI is equal to QFI), but sometimes there are many different QCRB sat. measurements, see Refs. [18] and [19] Section D for a detailed discussion.

Fisher information measurement noise susceptibility.—Let us assume, that due to a small disturbance, $\boldsymbol{M}$ changes to $\tilde{\boldsymbol{M}} = (1-\epsilon)\boldsymbol{M} + \epsilon \boldsymbol{N}$ (summation of two POVMs is done elementwise), which can be viewed as the replacement of desired POVM $\boldsymbol{M}$ with an unwanted one $\boldsymbol{N}$ with a probability $\epsilon \ll 1$. This type of noise may be caused by inaccurate initialization of a measuring device M in a mixed state $(1-\epsilon)|0\rangle\langle 0| + \epsilon\rho'_M$ instead of $|0\rangle\langle 0|$, or it may be the result of other small imperfections, such as signal losses, dark counts, crosstalks etc. (see Ref. [19] Sections A and B).

The measurement noise affects the CFI, the effect of which we quantify using

$$\chi[\boldsymbol{M},\boldsymbol{N}] = \lim_{\epsilon\to 0} \frac{F_C[\boldsymbol{M}] - F_C[(1-\epsilon)\boldsymbol{M} + \epsilon\boldsymbol{N}]}{\epsilon\, F_C[\boldsymbol{M}]}, \tag{5}$$

which can be understood as the relative *decrease* of the CFI under infinitesimally added noise $\boldsymbol{N}$—the effect of ϵ noise results in F_C in the CRB, Eq. (2), being replaced with $F_C[\boldsymbol{M}](1-\epsilon\chi[\boldsymbol{M},\boldsymbol{N}])$. After inserting (3) into (5), we obtain, after straightforward calculations,

$$\chi[\boldsymbol{M},\boldsymbol{N}] = 1 + F_C[\boldsymbol{M}]^{-1} G[\boldsymbol{N}], \tag{6}$$

where

$$G[\boldsymbol{N}] = \sum_i \mathrm{Tr}(A_i N_i), \qquad A_i = l_i^2 \rho_\theta - 2 l_i \dot{\rho}_\theta. \tag{7}$$

To get an intuition regarding this quantity, consider a simple example where $\boldsymbol{M} = (0, M_2, \ldots, M_K)$ and $\boldsymbol{N} = (\mathbb{1}, 0, \ldots, 0)$, so noise only activates a noninformative measurement outcome 1. Then, $G[\boldsymbol{N}] = 0$ and hence $\chi[\boldsymbol{M},\boldsymbol{N}] = 1$, which means that the relative decrease of the CFI is equal to the probability of obtaining a useless, noisy result. Clearly, the decrease of CFI will be more substantial, when the disturbance affects the statistics of informative outcomes, and noise cannot be separated from the signal easily. Our goal is to figure out, what is the maximal shrinkage rate of the CFI caused by an infinitesimal measurement noise described by an arbitrary POVM $\boldsymbol{N}$. The answer is given by a quantity

$$\chi[\boldsymbol{M}] = \max_{\boldsymbol{N}\in\mathcal{M}} \chi[\boldsymbol{M},\boldsymbol{N}], \tag{8}$$

which we call FI MENOS because it tells us how *susceptible* CFI is to small disturbances of the measurement. Note that the larger χ implies potentially stronger decrease of CFI as a result of measurement disturbance, so strictly speaking this is a *negative* susceptibility (cf. "menos" in Spanish). Notably, it does not depend on $\boldsymbol{N}$—this allows us to compare the robustness against noise of different measurements without invoking any specific noise model.

Explicit formula for FI MENOS.—We now present an explicit solution to the maximization problem from (8), which, according to (6), boils down to finding the maximum of $G[\boldsymbol{N}]$. Without loss of generality, we can relabel the elements of POVM $\boldsymbol{M}$ such that logarithmic derivatives satisfy $l_1 \leq l_2 \leq \ldots \leq l_K$. Let $\boldsymbol{N} = (N_1, \ldots, N_i, \ldots, N_K)$ be an arbitrary POVM while $\Gamma_i^1(\boldsymbol{N}) = (N_1 + N_i, \ldots, 0, \ldots, N_K)$ and $\Gamma_i^K(\boldsymbol{N}) = (N_1, \ldots, 0, \ldots, N_K + N_i)$ be two POVMs constructed from $\boldsymbol{N}$ with zeros at ith positions, $i \in \{2, \ldots, K-1\}$. Using (7), we obtain

$$G[\Gamma_i^1(\boldsymbol{N})] - G[\boldsymbol{N}] = f_i(l_1) - f_i(l_i), \tag{9}$$

$$G[\Gamma_i^K(\boldsymbol{N})] - G[\boldsymbol{N}] = f_i(l_k) - f_i(l_i), \tag{10}$$

where $f_i(x) = x^2\mathrm{Tr}(\rho_\theta N_i) - 2x\mathrm{Tr}(\dot{\rho}_\theta N_i)$ is a convex quadratic function, and therefore $f_i(l_1) \geq f_i(l_i) \vee f_i(l_K) \geq f_i(l_i)$, so from (9) and (10), $G[\Gamma_i^1(\boldsymbol{N})] \geq G[\boldsymbol{N}] \vee G[\Gamma_i^K(\boldsymbol{N})] \geq G[\boldsymbol{N}]$. Therefore, for each $\boldsymbol{N}$ and i, there is $j_i \in \{1, K\}$ such that $G[\Gamma_i^{j_i}(\boldsymbol{N})] \geq G[\boldsymbol{N}]$, so we can choose $i_2, \ldots, i_{K-1} \in \{1, K\}$ such that $\tilde{\boldsymbol{N}} = \Gamma_2^{j_2} \circ \ldots \circ \Gamma_{K-1}^{j_{K-1}}(\boldsymbol{N})$ satisfies $G[\tilde{\boldsymbol{N}}] \geq G[\boldsymbol{N}]$, and from the construction of $\tilde{\boldsymbol{N}}$, $\tilde{N}_i = 0$ for $i \in \{2, \ldots, K-1\}$. It means, that for arbitrary $\boldsymbol{N}$, it is possible to construct $\tilde{\boldsymbol{N}} = (\tilde{N}_1, 0, \ldots, 0, \mathbb{1} - \tilde{N}_1)$ satisfying $\chi[\boldsymbol{M}, \tilde{\boldsymbol{N}}] \geq \chi[\boldsymbol{M}, \boldsymbol{N}]$. Hence, the worst-case scenario noise will affect only the outcomes with the smallest and the largest logarithmic derivatives. This observation allows us to perform the maximization from (8) analytically, (see Ref. [19] Section C), and obtain the general expression for the FI MENOS:

$$\chi[\boldsymbol{M}] = 1 + \frac{1}{2F_C[\boldsymbol{M}]}(l_1^2 + l_K^2 + \|A_1 - A_K\|_1), \tag{11}$$

where $\|A\|_1 = \mathrm{Tr}\sqrt{AA^\dagger}$ is the trace norm. Notice, that χ depends on F_C, ρ_θ, $\dot{\rho}_\theta$, and the extremal logarithmic derivatives only. When an outcome i has a vanishing probability, $p_i \to 0$ but its contribution to the CFI, $p_i l_i^2$, remains finite and nonzero, then $l_i^2 \to \infty$, which implies that either l_1^2 or l_k^2 diverges, and hence, χ diverges as well according to (11). This reflects the fact, that the contribution to the CFI resulting from an outcome with a very low probability may be completely washed out by a very small measurement noise, and the chosen measurement is not likely to be practical.

When there are many measurements which lead to the same CFI, the FI MENOS may help to judge which one is more robust and hence more suitable for practical purposes—the one with lower χ. It is especially interesting to find the minimum of $\chi[\boldsymbol{M}]$ over all QCRB sat. measurements,

$$\chi_Q = \min_{\{\boldsymbol{M} \in \mathcal{M}, F_C[\boldsymbol{M}] = F_Q\}} \chi[\boldsymbol{M}], \tag{12}$$

since the corresponding measurement $\boldsymbol{M}$ should be regarded as the most robust among the most informative measurements. This task is tractable thanks to the exact formula (11)—we demonstrate exemplary solutions to this problem in the next two paragraphs.

Pure state models.—Let us start with a simple, yet important case when ρ_θ is pure, $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. We focus on the local estimation paradigm and assume θ is close to some known parameter value θ_0. For any θ_0 it is possible to fix orthonormal vectors $|0\rangle, |1\rangle \in \mathrm{span}(|\psi_{\theta_0}\rangle, |\dot{\psi}_{\theta_0}\rangle)$ such that $\rho_{\theta_0} = |+\rangle\langle+|$, $\dot{\rho}_{\theta_0} = \frac{1}{2}\sqrt{F_Q}\sigma_y$, where $|+\rangle = (1/\sqrt{2})(|0\rangle + |1\rangle)$, σ_y is a Pauli matrix, F_Q is the QFI. Then, the measurement $\boldsymbol{M}$ is QCRB sat. if and only if all its elements are of the form $M_i = \lambda_i|\phi_i\rangle\langle\phi_i|$, where $|\phi_i\rangle = (1/\sqrt{2})(|0\rangle + e^{i\varphi_i}|1\rangle)$ (see Ref. [19] Section E1). As we prove in [19] Section E1, $\chi[\boldsymbol{M}] \geq 4F_Q$ for all such measurements, and the inequality is only saturated for the projective measurement on the eigenstates of σ_y. Notice, that we used a qubit subspace to describe the evolution of any pure state locally even though $\mathcal{H}_S$ can be arbitrarily large.

This parametrization allows us to represent any pure state problem as a phase estimation in a Mach-Zender interferometer with a single photon input. When the phase θ between the upper ($|0\rangle$) and lower ($|1\rangle$) arm is acquired, then the photon state is $|\psi_\theta\rangle = (1/\sqrt{2})(|0\rangle + e^{i\theta}|1\rangle)$. After fixing $\theta_0 = 0$, our problem reduces to the one already defined with $F_Q = 1$. All QCRB sat. projective measurements can be implemented with the help of two single-photon detectors followed by a beam-splitter, and a well-controlled phase difference between two arms, φ (see Fig. 1). The upper and lower detectors click with probabilities p_+ and p_-, respectively, where $p_\pm = \frac{1}{2}[1 \pm \cos(\theta + \varphi)]$. Straightforward calculations confirm, that $F_C[p_\pm] = 1$ independently of φ [23]. However, the FI MENOS depends on φ, and for $\theta_0 = 0$ we have

$$\chi(\varphi) = 1 + \cos^{-2}(\varphi/2) + \tan^{-2}(\varphi/2), \tag{13}$$

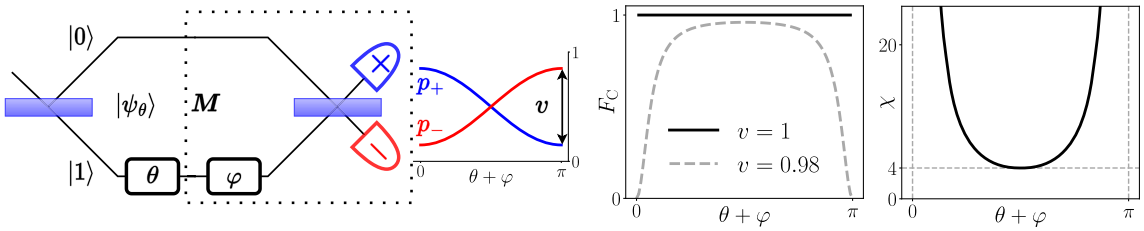


FIG. 1. Phase θ is measured using a Mach-Zender interferometer; φ is an extra, well-controlled phase. The resulting CFI (F_C) does not depend on φ when $v = 1$, which is never achieved in practice. For any smaller visibility (e.g., $v = 0.98$), the CFI is maximal for $\theta + \varphi = \pi/2$, and vanishes for $\theta + \varphi \in \{0, \pi\}$. We can deduce that $\theta + \varphi = \pi/2$ is the optimal working point without assuming nonunit visibility because FI MENOS (χ) is minimal there, so the precision of the estimation of θ is the least vulnerable to a general measurement noise.

which means that the optimal working point is at $\varphi = \pi/2$ (balanced interferometer), while our scheme is extremely sensitive to a measurement noise for $\varphi \to 0$ and $\varphi \to \pi$ (unbalanced interferometer) as in this case $\chi(\varphi) \to \infty$, see Fig. 1.

Similar conclusions follow from a standard analysis of the nonunit visibility ($v < 1$) interferometer model, where the detection probabilities are $p_\pm = \frac{1}{2}[1 \pm v\cos(\theta)]$. Then, the CFI is maximal for $\varphi = \pi/2$, and reaches 0 for $\varphi \in \{0, \pi\}$ even for v very close to 1, see Fig. 1. The advantage of the approach based on FI MENOS, is that one does not need to consider any particular noise model, and it is guaranteed that the worst case scenario has been taken into account.

Superresolution optical imaging.—Quantum estimation theory allows for a rigorous study of fundamental limits in optical microscopy, and serves as a tool for a systematic search for the most precise imaging schemes [24–30]. In the elementary scenario, two equally bright incoherent weak point sources are imaged using a translationally invariant system [25] (see discussion of limitations of this approximation in [31]). The state of a single photon in the image plane is

$$\rho_\theta = \frac{1}{2}(|u_{+,\theta}\rangle\langle u_{+,\theta}| + |u_{-,\theta}\rangle\langle u_{-,\theta}|), \tag{14}$$

where $\langle x|u_{\pm,\theta}\rangle = u(x \pm \theta/2)$, $\{|x\rangle\}$ is the position basis, $|u(x)|^2 = (2\pi\sigma^2)^{-1/2}e^{-x^2/2\sigma^2}$ is the system point spread function. The only unknown parameter is the separation between two sources, θ, the centroid of two points is known *a priori*. Intuitively, it should be hard to estimate θ when $\theta \ll \sigma$ because then images of two points overlap significantly. This is true for a standard measurement in the position basis $|x\rangle$ because $F_C[\{|x\rangle\langle x|\}] \to 0$ when $\theta \to 0$. Surprisingly, the QFI does not depend on θ at all, $F_Q[\rho_\theta] = 1/4\sigma^2$ [25]. Therefore, it seems to be no fundamental difference between small and large separations θ, when all quantum measurements are allowed. Unfortunately, this is a highly idealized statement since the estimation precision for small θ is highly affected by detection noise, system misalignment, crosstalk noise, and other imperfections [28–30]—in fact, even for the most clever choice of the measurement, the CFI vanishes with $\theta \to 0$ for all practical scenarios.

At this point, we want to demonstrate the fundamental difficulty of resolving two sources whose images overlap, without referring to any specific noise model, but rather employing the newly introduced FI MENOS figure of merit. In the most commonly studied superresolution protocol, the state in the image plane is measured in the basis of orthogonal Hermite-Gaussian modes $|\phi_q\rangle$ whose center lies in the centroid of two observed sources [25]—see Ref. [27] for an overview of implementations of this measurement based on holography, interferometry, etc. In most of these implementations, we can extract and separate first $K-1$ modes, and the rest of the signal is collected in the Kth outcome, such that our POVM consists of elements $M_i = |\phi_i\rangle\langle\phi_i|$ for $i \in \{1, \ldots, K-1\}$, $M_K = \mathbb{1} - \sum_{i=1}^{K-1} M_i$. The CFI increases with K, but for $\theta = 0$, it approaches the QFI already for $K = 2$. The QCRB is saturated in the full range of θ only for $K \to \infty$, but the precision is close to optimal for a wide range of θ already for $K = 4$—see Fig. 2. In the figure we also plot $\chi(\theta)$ for different values of K. Unfortunately, $\chi \to \infty$ for $\theta \to 0$ in all cases. Consequently, for $\theta \ll \sigma$, it is impossible to achieve the high CFI in the presence of noise using the family of measurements considered so far.

However, there are many other ways to saturate QCRB locally for any fixed value of θ, and some of them may be less susceptible to noise. As for the pure states model, we systematically study all QCRB sat. measurements to find $\chi_Q(\theta)$. Following the technique from [25], we reduce the problem to four-dimensional Hilbert space $\mathcal{H}^{(\theta)} = \mathrm{span}\{|u_{\pm,\theta}\rangle, \partial_\theta|u_{\pm,\theta}\rangle\}$, which is a direct sum of two orthogonal two-dimensional subspaces $\mathcal{H}_s$ and $\mathcal{H}_a$, containing symmetric and antisymmetric modes, respectively. We construct the orthonormal basis of both subspaces, $|0\rangle_s$, $|1\rangle_s$ and $|0\rangle_a$, $|1\rangle_a$ respectively, such that

$$\rho_\theta = \frac{1+\delta}{2}|0\rangle_s\langle 0| + \frac{1-\delta}{2}|0\rangle_a\langle 0|, \tag{15}$$

$$\dot{\rho}_\theta = \alpha|0\rangle_s\langle 0| + \beta_s\sigma_x^{(s)} - \alpha|0\rangle_a\langle 0| + \beta_a\sigma_x^{(a)}, \tag{16}$$

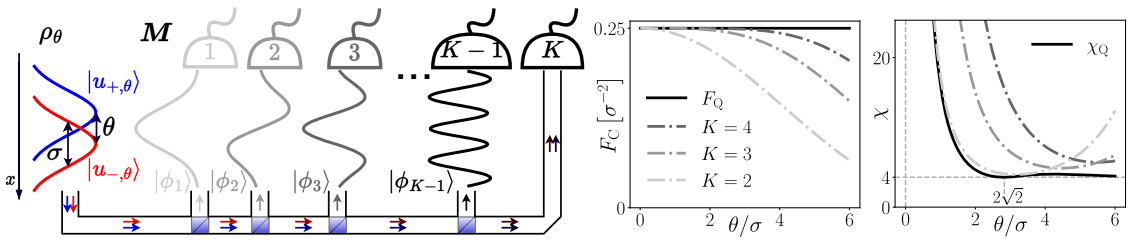


FIG. 2. The single photon state (ρ_θ) in an image plane is the mixture of two Gaussian wave functions of width σ. Their separation θ can be estimated accurately even when $\theta \ll \sigma$ by measuring ρ_θ in the basis of Hermite-Gaussian modes $|\phi_i\rangle$—in practice, it is enough to extract the first few $(K-1)$ of these modes. Unfortunately, the presented strategy is very sensitive to measurement noise for small θ, which is reflected by diverging FI MENOS (χ) for $\theta/\sigma \to 0$. No QCRB sat. measurement is free of this issue because χ_Q diverges as well for $\theta \to 0$.

where $\delta = \langle u_{+,\theta}|u_{-,\theta}\rangle$, real constants α, β_a, β_s are specified together with the exact construction of the basis in [19] Section E2. Matrices ρ_θ and $\dot{\rho}_\theta$ are both block diagonal with respect to $\mathcal{H}_s$ and $\mathcal{H}_a$. This means, that the QCRB sat. minimal susceptibility POVM contains only elements acting on $\mathcal{H}_s$ or $\mathcal{H}_a$ (see Ref. [19] Section D for proof). Consider the family of QCRB sat. POVMs

$$\boldsymbol{M}_{\varphi_s,\varphi_a} = \{P^{(s)}_{\varphi_s}, P^{(s)}_{\varphi_s+\pi}, P^{(a)}_{\varphi_a}, P^{(a)}_{\varphi_a+\pi}\}, \tag{17}$$

where $P^{s/a}_{\varphi}$ is a projector on $\cos(\varphi/2)|0\rangle_{s/a} + \sin(\varphi/2)|1\rangle_{s/a}$. We prove ([19] Section E2) that the minimal susceptibility QCRB sat. measurement is of the form $\boldsymbol{M}_{\varphi_s,\varphi_a}$. Then, we obtain χ_Q by minimizing numerically $\chi[\boldsymbol{M}_{\varphi_s,\varphi_a}]$ over φ_s and φ_a, the results are shown in Fig. 2. We observe, that $\chi_Q \to \infty$ when $\theta \to 0$, which means, that no QCRB sat. measurement is robust against noise in the region $\theta \ll \sigma$. Surprisingly, χ_Q does not decrease with θ everywhere—for example, a minimum $\chi = 4$ is achieved in $\theta = 2\sqrt{2}\sigma$. For $\theta \to \infty, \chi_Q \to 4$ again, and then the problem is equivalent to a single source localization both from the point of view of F_Q and χ_Q. Interestingly, it is possible to achieve noise susceptibility smaller than χ_Q, when correlations between subsequent photons are present—see Ref. [19] Section F for a discussion.

Outlook.—Computation of the FI MENOS should be regarded as a natural sanity check whenever any idealized quantum metrological protocol is proposed. If this quantity is large (or divergent) this should ring a bell that the performance of the proposed protocol will be significantly reduced by even a small imperfection in the measurement design. On the contrary, small values indicate that the measurement scheme is robust. The importance of this quantity stems also from the fact, that the FI is a local quantity (computed at single value of parameter) and therefore prone to reveal ephemeral effects that vanish in the presence of even infinitesimal noise—a property that haunts quantum metrology literature a lot. We envisage that our approach may be naturally extended to a multiparameter estimation framework as well as generalized to cover the Bayesian analysis as well. Still, we expect that in these cases it may be much harder or even impossible to obtain the explicit formula for FI MENOS analogous to (11). Moreover, we expect that the role of small measurement disturbances should be less substantial in the Bayesian approach, since such an approach is by construction applicable to more realistic scenarios, when the number of collected data samples is finite, and the protocols are expected to perform well beyond the "local estimation approach".

We thank Wojciech Górecki and Janek Kołodyński for fruitful discussions. This work was supported by the National Science Center (Poland) Grant No. 2020/37/B/ST2/02134.

*s.kurdzialek@student.uw.edu.pl

[1] J. Preskill, Quantum **2**, 79 (2018).
[2] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Phys. **5**, 222 (2011).
[3] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kołodyński, in *Progress in Optics*, edited by E. Wolf (Elsevier, New York, 2015), Vol. 60, pp. 345–435.
[4] C. L. Degen, F. Reinhard, and P. Cappellaro, Rev. Mod. Phys. **89**, 035002 (2017).
[5] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Rev. Mod. Phys. **90**, 035005 (2018).
[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[7] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[8] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nat. Phys. **7**, 406 (2011).
[9] R. Demkowicz-Dobrzański, M. Guţă, and J. Kołodyński, Nat. Commun. **3**, 1063 (2012).
[10] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Phys. Rev. X **7**, 041009 (2017).
[11] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Nat. Commun. **9**, 78 (2018).
[12] S. Zhou and L. Jiang, PRX Quantum **2**, 010343 (2021).
[13] A. Datta, L. Zhang, N. Thomas-Peter, U. Dorner, B. J. Smith, and I. A. Walmsley, Phys. Rev. A **83**, 063836 (2011).
[14] Y. L. Len, T. Gefen, A. Retzker, and J. Kołodyński, Nat. Commun. **13**, 6971 (2022).
[15] P. Skrzypczyk and N. Linden, Phys. Rev. Lett. **122**, 140403 (2019).
[16] M. Oszmaniec and T. Biswas, Quantum **3**, 133 (2019).
[17] A. W. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, Cambridge, England, 1998).
[18] S. Zhou, C.-L. Zou, and L. Jiang, Quantum Sci. Technol. **5**, 025005 (2020).
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.160802 for more information, which includes Refs. [20–22].
[20] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, Proc. Natl. Acad. Sci. U.S.A. **106**, 22287 (2009).
[21] S. Kurdziałek and R. Demkowicz-Dobrzański, J. Opt. **23**, 075701 (2021).
[22] O. Schwartz and D. Oron, Phys. Rev. A **85**, 033812 (2012).
[23] Strictly speaking, we cannot use (2) directly for $\varphi \in \{0, \pi\}$ because then $p_i = 0$ and $\dot{p}_i = 0$ for one term of the sum from (2). If we exclude such terms by convention, we obtain $F_C = 0$, but if we calculate limits $\varphi \to 0$ or $\varphi \to \pi$, we obtain $F_C = 1$.
[24] E. Bettens, D. V. Dyck, A. den Dekker, J. Sijbers, and A. van den Bos, Ultramicroscopy **77**, 37 (1999).
[25] M. Tsang, R. Nair, and X.-M. Lu, Phys. Rev. X **6**, 031033 (2016).

[26] C. Lupo and S. Pirandola, Phys. Rev. Lett. **117**, 190802 (2016).
[27] M. Tsang, Contemp. Phys. **60**, 279 (2019).
[28] Y. L. Len, C. Datta, M. Parniak, and K. Banaszek, Int. J. Quantum. Inform. **18**, 1941015 (2020).
[29] C. Oh, S. Zhou, Y. Wong, and L. Jiang, Phys. Rev. Lett. **126**, 120502 (2021).
[30] J. O. de Almeida, J. Kołodyński, C. Hirche, M. Lewenstein, and M. Skotiniotis, Phys. Rev. A **103**, 022406 (2021).
[31] S. Kurdzialek, Quantum **6**, 697 (2022).

Measurement noise susceptibility in quantum estimation: Supplementary Material

Stanisław Kurdziałek* and Rafał Demkowicz-Dobrzański
Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland

Appendix A: Noise due to inaccurate meter preparation

In this Section, we prove that an inaccurate initialization of a measuring device M in a state $\rho_M = (1-\epsilon)\left|0\right\rangle\!\left\langle 0\right| + \epsilon\rho'_M$ instead of $\left|0\right\rangle\!\left\langle 0\right|$ changes the POVM acting on S from $\boldsymbol{M}$ to $\tilde{\boldsymbol{M}} = (1-\epsilon)\boldsymbol{M} + \epsilon\boldsymbol{N}$, where the exact form of POVM $\boldsymbol{N}$ depends on ρ'_M. Initially, S and M are uncorrelated, the state of S is ρ_S. The joint state of S and M after unitary interaction U_{SM} is

$$\rho_{SM} = U_{SM}\rho_S \otimes \rho_M U_{SM}^\dagger = (1-\epsilon)U_{SM}\rho_S \otimes \left|0\right\rangle\!\left\langle 0\right| U_{SM}^\dagger + \epsilon U_{SM}\rho_S \otimes \rho'_M U_{SM}^\dagger. \tag{A1}$$

After interaction, the projective measurement of M in the basis $\left|i\right\rangle_{i\in\{1,\ldots,K\}}$ is performed. The probability of obtaining i-th outcome is

$$p(i) = \mathrm{Tr}\left({}_M\!\left\langle i\right|\rho_{SM}\left|i\right\rangle_M\right) = (1-\epsilon)p_0(i) + \epsilon p_1(i), \tag{A2}$$

where

$$p_0(i) = \mathrm{Tr}\left(\rho_S M_i\right), \quad M_i = {}_M\!\left\langle 0\right|U_{SM}^\dagger\left|i\right\rangle_M\!\left\langle i\right|U_{SM}\left|0\right\rangle_M \tag{A3}$$

$$p_1(i) = \mathrm{Tr}\left(\rho_S \otimes \mathbb{1}\cdot\mathbb{1} \otimes \rho'_M U_{SM}^\dagger \left|i\right\rangle_M\!\left\langle i\right| U_{SM}\right) = \mathrm{Tr}\left(\rho_S \mathrm{Tr}_M\left(\mathbb{1}\otimes\rho'_M U_{SM}^\dagger\left|i\right\rangle_M\!\left\langle i\right| U_{SM}\right)\right) = \mathrm{Tr}\left(\rho_S N_i\right). \tag{A4}$$

The partial scalar product notation was used—$\left|i\right\rangle_M \equiv \mathbb{1}\otimes\left|i\right\rangle$, $\left\langle i\right|_M \equiv \mathbb{1}\otimes\left\langle i\right|$, partial trace is defined as $\mathrm{Tr}_M(\bullet) = \sum_i {}_M\!\left\langle i\right|\bullet\left|i\right\rangle_M$. Operators $M_i \in \mathcal{L}\left(\mathcal{H}_S\right)$ are the same as those defined in (1), and form undisturbed POVM $\boldsymbol{M}$, by $\mathcal{L}\left(\mathcal{H}\right)$ we denote the set of all operators acting on $\mathcal{H}$. Let us now prove that $\boldsymbol{N} = \{N_1, \ldots, N_K\}$, where $N_i \in \mathcal{L}\left(\mathcal{H}_S\right)$ is also POVM. Firstly,

$$\sum_i N_i = \mathrm{Tr}_M\left(\mathbb{1}\otimes\rho'_M U_{SM}^\dagger \sum_i \left|i\right\rangle_M\!\left\langle i\right| U_{SM}\right) = \mathrm{Tr}_M\left(\mathbb{1}\otimes\rho'_M\right) = \mathbb{1}_S. \tag{A5}$$

Moreover,

$$N_i = \mathrm{Tr}_M\left(K_i^\dagger K_i\right), \quad K_i = {}_M\!\left\langle i\right| U_{SM}\ \mathbb{1}\otimes\sqrt{\rho'_M}, \tag{A6}$$

so $N_i \geq 0$ because partial trace of positive operator is positive. From (A2),

$$p(i) = (1-\epsilon)\mathrm{Tr}(\rho_S M_i) + \epsilon\mathrm{Tr}(\rho_S N_i) = \mathrm{Tr}(\rho_S \tilde{M}_i), \tag{A7}$$

so indeed $\boldsymbol{M}$ is replaced with $\tilde{\boldsymbol{M}}$ due to the considered disturbance of the initial state of a measuring device.

Appendix B: Different sources of noise—examples

In practice, the real source of noise may be different than the one described in the previous section. However, our general noise model $\boldsymbol{M} \to (1-\epsilon)\boldsymbol{M} + \epsilon\boldsymbol{N}$ encompasses many different types of noise. In this Section, we demonstrate a few practical examples of noise types together with corresponding noise POVMs $\boldsymbol{N}$.

* s.kurdzialek@student.uw.edu.pl

1. Signal losses

Let us consider arbitrary POVM with $K-1$ elements, $\boldsymbol{M'}=(M_2,...,M_K)$—clearly, the measurement described by $\boldsymbol{M'}$ is equivalent to the one described by $\boldsymbol{M}=(0,M_2,...,M_K)$ because the probability associated with the 1st element of $\boldsymbol{M}$ is always 0. When the noise of the form $\boldsymbol{N}=(\mathbb{1},0,...,0)$ acts with probability ϵ, then the resulting POVM is $\tilde{\boldsymbol{M}}=(\epsilon\mathbb{1},(1-\epsilon)M_2,...,(1-\epsilon)M_K)$. The 1st element of $\tilde{\boldsymbol{M}}$ corresponds to non-informative outcome because its probability doesn't depend on ρ and is always equal to ϵ. Therefore, $\tilde{\boldsymbol{M}}$ describes the scenario in which we loose our signal with probability ϵ, independently of the measured state. Then, $G[\boldsymbol{N}]=0$, and consequently $\chi[\boldsymbol{M},\boldsymbol{N}]=1$—for example, when $\epsilon=1\%$, then we loose 1% of the signal, which leads to $\epsilon\chi[\boldsymbol{M},\boldsymbol{N}]=1\%$ relative decrease of the Fisher Information.

2. Dark counts

Let us assume, that we have K detectors in our setup, undisturbed measurement is described by POVM $\boldsymbol{M}=(M_1,...,M_K)$. Detector may sometimes "click" even if no signal was sent—such events are called "dark counts", in each repetition real signal is observed with probability $1-\epsilon$, and a dark count with probability ϵ. Probabilities of dark counts do not depend on ρ, but may be different for different detectors, so the most general POVM describing dark counts is $\boldsymbol{N}=(q_1\mathbb{1},...,q_K\mathbb{1})$, $q_1+...+q_K=1$. Dark counts are generally more problematic than signal losses because noise is not separated from the signal. This is reflected by larger $\chi[\boldsymbol{M},\boldsymbol{N}]$—notice that $G[\boldsymbol{N}]=\sum_i q_i l_i^2>0$, where $l_i=\frac{\partial_\theta p_i}{p_i}$, $p_i=\mathrm{Tr}(\rho_\theta M_i)$. For a particular case $q_i=p_i$, we have $G[\boldsymbol{N}]=F_C[\boldsymbol{M}]$, so $\chi[\boldsymbol{M},\boldsymbol{N}]=2$—there are dark counts models, for which the decrease of FI is twice as large as for signal losses.

3. Cross-talks

This type of noise appears, when different measurement outcomes are sometimes confused with each other. In a noiseless scenario, i-th detector clicks with a probability $p_i=\mathrm{Tr}(\rho_\theta M_i)$ ($i\in\{1,...,K\}$). However, because of cross-talks, the result associated with M_j is interpreted as i with a probability $t(i|j)$, so the probability of obtaining i-th outcome becomes $\tilde{p}_i=\sum_{j=1}^K t(i|j)p_j=\mathrm{Tr}(\rho_\theta\tilde{M}_i)$, where operators $\tilde{M}_i=\sum_j t(i|j)M_j$ form a noise-affected POVM $\tilde{\boldsymbol{M}}$, $\sum_i t(i|j)=1$ for all j. As in all previous examples, we assume that the noise is ϵ-small, so $t(i|i)\geq 1-\epsilon$ for all i. Then we can express cross-talk probabilities as $t(i|j)=(1-\epsilon)\delta_{ij}+\epsilon\tilde{t}(i|j)$, where $\tilde{t}(i|j)>0$, $\sum_i\tilde{t}(i|j)=1$. Consequently, $\tilde{M}_i=(1-\epsilon)M_i+\epsilon N_i$, where operators $N_i=\sum_j\tilde{t}(i|j)M_j$ form a noise POVM $\boldsymbol{N}$.

4. Random rotations of measurement basis

Let us consider a projective measurement $\boldsymbol{M}=(|\phi_1\rangle\langle\phi_1|,...,|\phi_K\rangle\langle\phi_K|)$. In some cases, the misalignment of the measuring device may change the measurement basis. Let us assume, that the basis is rotated by unitary operation $\hat{U}_j$ with probability ϵq_j, $\sum_j q_j=1$. Then, the noise-affected POVM is $\tilde{\boldsymbol{M}}=(1-\epsilon)\boldsymbol{M}+\epsilon\boldsymbol{N}$, where $\boldsymbol{N}$ consists of elements $N_i=\sum_j q_j U_j|\phi_i\rangle\langle\phi_i|U_j^\dagger$.

Appendix C: Explicit Formula for FI MENOS

In this Section, the derivation of (11) is completed. As we proved in the main text, the worst-case scenario noise is of the form $\tilde{\boldsymbol{N}}=(\tilde{N}_1,0,...,0,\mathbb{1}-\tilde{N}_1)$, and the maximization of $G[\boldsymbol{N}]$ over all POVMs $\boldsymbol{N}$ simplifies to

$$\max_{\boldsymbol{N}\in\mathcal{M}} G[\boldsymbol{N}]=\max_{0\leq\tilde{N}_1\leq\mathbb{1}}\mathrm{Tr}\left[A_1\tilde{N}_1+A_K\left(\mathbb{1}-\tilde{N}_1\right)\right]=l_K^2+\max_{0\leq\tilde{N}_1\leq\mathbb{1}}\mathrm{Tr}\left[(A_1-A_K)\tilde{N}_1\right]. \tag{C1}$$

Let $|e_i\rangle$ be orthonormal basis diagonalizing A_1-A_K. We can express both A_1-A_K and $\tilde{N}_1$ in this basis:

$$A_1-A_K=\sum_{i=1}^R\lambda_i|e_i\rangle\langle e_i|,\quad \tilde{N}_1=\sum_{i,j=1}^R\tilde{N}_{ij}|i\rangle\langle j|, \tag{C2}$$

where $R = \dim\mathcal{H}_S$ is the size of the matrices. We choose the ordering of $|e_i\rangle$ such that for some $P \in \{1, ..., R\}$, $\lambda_1, ..., \lambda_P \geq 0$ and $\lambda_{P+1}, ..., \lambda_R < 0$. The maximization problem from the RHS of (C1) transforms to

$$\max_{0\leq\tilde{N}_1\leq\mathbb{1}} \mathrm{Tr}\left[(A_1 - A_K)\tilde{N}_1\right] = \max_{\tilde{N}_{ii}} \sum_i \tilde{N}_{ii}\lambda_i. \tag{C3}$$

From positivity of $\tilde{N}_1$ and of $\mathbb{1} - \tilde{N}_1$, $\tilde{N}_{ii} \geq 0$ and $\tilde{N}_{ii} \leq 1$, therefore

$$\max_{0\leq\tilde{N}_1\leq\mathbb{1}} \mathrm{Tr}\left[(A_1 - A_K)\tilde{N}_1\right] = \sum_{i=1}^{P} \lambda_i, \tag{C4}$$

which is the sum of positive eigenvalues of $A_1 - A_K$. Moreover, we have $\mathrm{Tr}(A_1 - A_K) = \sum_{i=1}^{P} \lambda_i + \sum_{i=P+1}^{R} \lambda_i$, $\|A_1 - A_K\|_1 = \sum_{i=1}^{P} \lambda_i - \sum_{i=P+1}^{R} \lambda_i$, which means that

$$\sum_{i=1}^{P} \lambda_i = \frac{1}{2}\left(\mathrm{Tr}(A_1 - A_K) + \|A_1 - A_K\|_1\right) = \frac{1}{2}(l_1^2 - l_K^2 + \|A_1 - A_K\|_1). \tag{C5}$$

Taking all this results together, we obtain

$$\max_{\boldsymbol{N}\in\mathcal{M}} G[\boldsymbol{N}] = \max_{0\leq\tilde{N}_1\leq\mathbb{1}} \mathrm{Tr}\left[A_1\tilde{N}_1 + A_K\left(\mathbb{1} - \tilde{N}_1\right)\right] = \frac{1}{2}(l_1^2 + l_K^2 + \|A_1 - A_K\|_1), \tag{C6}$$

which, after inserting into (6), leads to (11).

Appendix D: The most robust among the most informative measurements—Theorems

In this Section, we show how to approach the problem of finding a QCRB sat. measurement with the minimal FI MENOS. Before showing solutions for examples mentioned in the main text, let us introduce some theorems that help to systematically describe potential candidates for such a measurement. Let us start with quoting a condition for QCRB saturability, firstly introduced in [1], and then reformulated in [2] in a following form:

Theorem 1. POVM $\boldsymbol{M} = \{M_i\}_i$ is QCRB sat. for the family of states ρ_θ around $\theta = \theta_0$ iff

$$M_i^{1/2} L_{jk} M_i^{1/2} = 0, \ \forall i, j, k \tag{D1}$$

and

$$\forall i \text{ s.t. } \mathrm{Tr}\left(M_i\rho_{\theta_0}\right) = 0, \quad M_i^{1/2}\Lambda_{\theta_0}\left|\psi_{\theta_0,j}\right\rangle = 0, \ \forall j. \tag{D2}$$

Here $\rho_\theta = \sum_j p_{\theta,j}\left|\psi_{\theta,j}\right\rangle\!\left\langle\psi_{\theta,j}\right|$ is the diagonalization of ρ_θ, $p_{\theta,j} > 0$, Λ_θ is the symmetric logarithmic derivative matrix, and $L_{jk} = \left|\psi_{\theta_0,j}\right\rangle\!\left\langle\psi_{\theta_0,k}\right|\Lambda_{\theta_0} - \Lambda_{\theta_0}\left|\psi_{\theta_0,j}\right\rangle\!\left\langle\psi_{\theta_0,k}\right|$.

Proof. See Ref.[2] for a proof. ■

Next two lemmas will help us to reduce the set of potential candidates for the minimal MENOS QCRB sat. measurements to those containing only rank-one elements M_i (see Theorem 2).

Lemma 1 Let $p_\theta(i)$ be the family of probability distributions indexed by θ, $i \in \{1, ..., K\}$. Let $q_\theta(j) = \sum_i p_\theta(i)t(j|i)$ be another probability distribution, $j \in \{1, ..., L\}$, $t(j|i) \geq 0$, $\sum_j t(j|i) = 1$. Then

$$1.\ F_C[p_\theta(i)] \geq F_C[q_\theta(j)], \tag{D3}$$

$$2.\ F_C[p_\theta(i)] = F_C[q_\theta(j)] \text{ iff } \forall_{j,i_1,i_2}\left(t(j|i_1) \neq 0 \wedge t(j|i_2) \neq 0\right) \Rightarrow l_{i_1} = l_{i_2}, \tag{D4}$$

where $l_i = \frac{\dot{p}_\theta(i)}{p_\theta(i)}$. Point 2. means, that random mixing of results does not decrease the CFI iff only results with the same logarithmic derivatives are mixed with each other.

Proof. Let us adopt a short-hand notation $p_i \equiv p_\theta(i)$, $q_j \equiv q_\theta(j)$. The CFI associated with the distribution $q_\theta(j)$ can be written as

$$F_C\left[q_\theta(j)\right] = \sum_{j=1}^{L} F_j, \text{ where } F_j = \frac{\dot{q}_j^2}{q_j}. \tag{D5}$$

Now we want to prove, that

$$F_j = \frac{\left(\sum_{i=1}^{K} t(j|i)\dot{p}_i\right)^2}{q_j} \leq \sum_{i=1}^{K} t(j|i)\frac{\dot{p}_i^2}{p_i}. \tag{D6}$$

For a fixed j, let us introduce the notation $\alpha_i = \frac{t(j|i)p_i}{q_j}$. From the definition of q_j follows that $\sum_i \alpha_i = 1$, moreover $\alpha_i \geq 0$. Taking this into account, we have

$$\left|\sum_{i=1}^{K} t(j|i)\dot{p}_i\right| = \left|\sum_{i=1}^{K} \alpha_i \frac{t(j|i)\dot{p}_i}{\alpha_i}\right| \overset{(1)}{\leq} \sum_{i=1}^{K} \alpha_i \left|\frac{t(j|i)\dot{p}_i}{\alpha_i}\right| \overset{(2)}{\leq} \sqrt{\sum_{i=1}^{K} \alpha_i \left(\frac{t(j|i)\dot{p}_i}{\alpha_i}\right)^2}, \tag{D7}$$

where (1) is triangle inequality, and (2) follows from the weighted power mean inequality (between arithmetic and quadratic mean, AM-QM in short). After squaring both sides of D7, substituting the definition of α_i in the RHS, and dividing both sides by q_j we obtain inequality D6. Subsequently, using D6 and D5 , we obtain

$$F_C\left[q_\theta(j)\right] \leq \sum_{j=1}^{L}\sum_{i=1}^{K} t(j|i)\frac{\dot{p}_i^2}{p_i} = \sum_{i=1}^{K} \frac{\dot{p}_i^2}{p_i} = F_C\left[p_\theta(i)\right], \tag{D8}$$

which is exactly part 1. of our Lemma. In order to prove part 2., let us notice, that inequality D8 is saturated iff inequalities (1) and (2) in D7 are saturated for all j. For a given j, part (1) of D7 becomes equality iff for all i s.t. $t(j|i) \neq 0$, terms $\frac{t(j|i)\dot{p}_i}{\alpha_i} = q_j\frac{\dot{p}_i}{p_i}$ have the same sign. Part (2), which is AM-QM inequality becomes saturated, when all terms in the mean with non-zero weights are equal, which implies that for a given j, for all all i s.t. $t(j|i) \neq 0$ terms $\left|q_j\frac{\dot{p}_i}{p_i}\right|$ are equal. From these two saturability conditions, we obtain part (2) of Lemma 1. ■

The next lemma states that random mixing of the elements of POVM either decreases the CFI or keeps both the CFI and FI MENOS unaffected.

Lemma 2. Let $\boldsymbol{M}, \boldsymbol{M'} \in \mathcal{M}$, $\boldsymbol{M} = \{M_i\}_{i\in\{1,\ldots,K\}}$, $\boldsymbol{M'} = \{M'_j\}_{j\in\{1,\ldots,L\}}$, $M'_j = \sum_i t(j|i)M_i$, where $t(j|i) \geq 0$, $\sum_j t(j|i) = 1$. The family of quantum states ρ_θ is fixed. Then

$$1.\ F_C[\boldsymbol{M}] \geq F_C[\boldsymbol{M'}], \tag{D9}$$

$$2.\ F_C[\boldsymbol{M}] = F_C[\boldsymbol{M'}] \Rightarrow \chi\left[\boldsymbol{M}\right] = \chi\left[\boldsymbol{M'}\right]. \tag{D10}$$

Proof. The classical probability distributions associated with measurements $\boldsymbol{M}$ and $\boldsymbol{M'}$ are $p_\theta(i) = \mathrm{Tr}(\rho_\theta M_i)$, $q_\theta(j) = \mathrm{Tr}(\rho_\theta M_j)$. From the definition of $\boldsymbol{M'}$, we have $q_\theta(j) = \sum_i t(j|i)p_\theta(i)$. Therefore, part 1. of Lemma 2 is a direct consequence of part 1. of Lemma 1.

The logarithmic derivative of q_j is

$$l'_j = \frac{\dot{q}_j}{q_j} = \frac{\sum_i t(j|i)\dot{p}_i}{\sum_i t(j|i)p_i} = \frac{\sum_i t(j|i)l_i p_i}{\sum_i t(j|i)p_i} \tag{D11}$$

When $F_C[\boldsymbol{M}] = F_C[\boldsymbol{M'}]$, then, according to part 2. of Lemma 1.,

$$l_{j_1} = l_{j_2} = \ldots = l_{j_{K_j}} = l'_j, \tag{D12}$$

where $j_1, j_2, \ldots j_{K_j}$ are the indices satisfying $t(j|j_i) \neq 0$. That means, that the sequence of logarithmic derivatives l'_j consist off exactly the same elements as the sequence l_i (some elements may repeat), and the minimal and maximal logarithmic derivatives for $\boldsymbol{M}$ and $\boldsymbol{M'}$ are the same, so according to (11), $\chi\left[\boldsymbol{M}\right] = \chi\left[\boldsymbol{M'}\right]$. ■

Theorem 2. For any family of quantum states ρ_θ, there exists a QCRB sat. measurement with a minimal FI MENOS whose all elements are rank-one matrices.

Proof. Let $\boldsymbol{M'} = \{M'_i\}_i$ be any QCRB sat. POVM with a minimal FI MENOS, $F_C\left[\boldsymbol{M'}\right] = F_Q$, $\chi[\boldsymbol{M'}] = \chi_Q$. In a diagonal form, $M'_i = \sum_j \lambda_{i,j}\left|\phi_{i,j}\right\rangle\!\left\langle\phi_{i,j}\right|$, $\lambda_{i,j} \geq 0$ because M'_i are positive-semidefinite. Let us consider a POVM $\boldsymbol{M} = \{\lambda_{i,j}\left|\phi_{i,j}\right\rangle\!\left\langle\phi_{i,j}\right|\}_{i,j}$, whose elements are all rank-one matrices. According to Lemma 2, $F_C\left[\boldsymbol{M}\right] \geq F_C\left[\boldsymbol{M'}\right] = F_Q$, but on the other hand, the CFI cannot be greater then the QFI, so $F_C\left[\boldsymbol{M}\right] = F_C\left[\boldsymbol{M'}\right] = F_Q$. Therefore, from part 2.

of Lemma 2., $\chi[\boldsymbol{M}] = \chi[\boldsymbol{M}'] = \chi_Q$, which means, that $\boldsymbol{M}$ is a QCRB sat. POVM with a minimal MENOS consisting of rank-one matrices only. ■

The next theorem is helpful for problems involving states with block-diagonal structure, it will be used in the next section to study super-resolution imaging.

Theorem 3. Let $\rho_\theta, \dot{\rho}_\theta \in \mathcal{L}(\mathcal{H})$ be a density matrix and its derivative, which can be decomposed as

$$\rho_\theta = \rho_\theta^{(1)} + \rho_\theta^{(2)} + ... + \rho_\theta^{(L)}, \tag{D13}$$

$$\dot{\rho}_\theta = \dot{\rho}_\theta^{(1)} + \dot{\rho}_\theta^{(2)} + ... + \dot{\rho}_\theta^{(L)}, \tag{D14}$$

where $\rho_\theta^{(l)}, \dot{\rho}_\theta^{(l)}$ act on a subspace $\mathcal{H}^{(l)} \subset \mathcal{H}$ (it has zeros outside this subspace), $\mathcal{H} = \mathcal{H}^{(1)} \oplus \mathcal{H}^{(2)} \oplus ... \oplus \mathcal{H}^{(L)}$, subspaces $\mathcal{H}^{(l)}$ are orthogonal to each other. Then, there exists a QCRB sat. measurement with a minimal MENOS, which can be written as

$$\boldsymbol{M} = \left\{ M_1^{(1)}, ..., M_{K_1}^{(1)}, M_1^{(2)}, ..., M_{K_2}^{(2)}, ..., M_1^{(L)}, ..., M_{K_L}^{(L)}, \right\} \tag{D15}$$

where for each $l \in \{1, ..., L\}$ matrices $M_1^{(l)}, ..., M_{K_l}^{(l)}$ act on $\mathcal{H}^{(l)}$ and form a QCRB sat. POVM for a family of normalized states $\tilde{\rho}_\theta^{(l)} = \rho_\theta^{(l)} / \mathrm{Tr}\left(\rho_\theta^{(l)}\right)$.

Proof. Let us denote an orthogonal projector on $\mathcal{H}^{(l)}$ by $P^{(l)}$. Let $\boldsymbol{M}' = \{M_i'\}_{i \in \{1,...,K\}}$ be any QCRB sat. POVM with a minimal FI MENOS. Let us construct another POVM $\boldsymbol{M}'' = \{M_i''\}_{i \in \{1,...,K\}}$, where $M_i'' = P^{(1)} M_i' P^{(1)} + P^{(2)} M_i' P^{(2)} + ... + P^{(L)} M_i' P^{(L)}$.

We have

$$\mathrm{Tr}(\rho_\theta M_i') = \mathrm{Tr}\left(\left(\rho_\theta^{(1)} + \rho_\theta^{(2)} + ... + \rho_\theta^{(L)}\right) M_i'\right) = \mathrm{Tr}\left(\rho_\theta \left(P^{(1)} M_i' P^{(1)} + ... + P^{(L)} M_i' P^{(L)}\right)\right) = \mathrm{Tr}\left(\rho_\theta M_i''\right), \tag{D16}$$

where we used the fact that $\rho_\theta^{(l)} = P^{(l)} \rho_\theta P^{(l)}$. Similarly,

$$\mathrm{Tr}(\dot{\rho}_\theta M_i') = \mathrm{Tr}\left(\left(\dot{\rho}_\theta^{(1)} + \dot{\rho}_\theta^{(2)} + ... + \dot{\rho}_\theta^{(L)}\right) M_i'\right) = \mathrm{Tr}\left(\dot{\rho}_\theta M_i''\right). \tag{D17}$$

Therefore, POVMs $\boldsymbol{M}'$ and $\boldsymbol{M}''$ are fully equivalent—probabilities of different outcomes $p_\theta(i)$ and their derivatives $\dot{p}_\theta(i)$ are the same for $\boldsymbol{M}'$ and $\boldsymbol{M}''$. Consequently, $F_C[\boldsymbol{M}'] = F_C[\boldsymbol{M}'']$ and $\chi[\boldsymbol{M}'] = \chi[\boldsymbol{M}'']$.

Let us finally construct a POVM

$$\boldsymbol{M} = \left\{ P^{(1)} M_1' P^{(1)}, P^{(1)} M_2' P^{(1)}, ..., P^{(1)} M_K' P^{(1)}, ..., P^{(L)} M_1' P^{(L)}, P^{(L)} M_2' P^{(L)}, ..., P^{(L)} M_K' P^{(L)} \right\}, \tag{D18}$$

which is clearly of the form from (D15) because $P^{(l)} M_i' P^{(l)}$ act on $\mathcal{H}^{(l)}$. Moreover, elements of $\boldsymbol{M}''$ are linear combinations of the elements of $\boldsymbol{M}$, so, according to Lemma 2, $F_C[\boldsymbol{M}] = F_Q$ and $\chi[\boldsymbol{M}] = \chi_Q$. Therefore, we constructed a QCRB sat. POVM with a minimal MENOS which has a form defined in (D15). Let us now prove, that when $\boldsymbol{M}$ from (D15) is QCRB sat., then for any $l \in \{1, ..., L\}$, POVM $\boldsymbol{M}^{(l)} = \{P^{(l)} M_1' P^{(l)}, P^{(l)} M_2' P^{(l)}, ..., P^{(l)} M_K' P^{(l)}\}$ is QCRB sat. for $\tilde{\rho}_\theta^{(l)}$. It is straightforward to show that

$$F_C[\rho_\theta, \boldsymbol{M}] = \sum_{l=1}^{L} F_C\left[\tilde{\rho}_\theta^{(l)}, \boldsymbol{M}^{(l)}\right] + F_C[P_\theta(l)], \tag{D19}$$

where $P_\theta(l) = \mathrm{Tr}\left(\rho_\theta^{(l)}\right)$ is the probability of measuring ρ_θ in a subspace $\mathcal{H}^{(l)}$. When for some l, $\boldsymbol{M}^{(l)}$ is not QCRB sat., then there exists a POVM $\boldsymbol{M}_2^{(l)}$ for which $F_C\left[\tilde{\rho}_\theta^{(l)}, \boldsymbol{M}_2^{(l)}\right] > F_C\left[\tilde{\rho}_\theta^{(l)}, \boldsymbol{M}^{(l)}\right]$. Then, according to (D19), $F_C[\rho_\theta, \boldsymbol{M}_2] > F_C[\rho_\theta, \boldsymbol{M}]$, where $\boldsymbol{M}_2 = \left\{\boldsymbol{M}^{(1)}, ..., \boldsymbol{M}_2^{(l)}, ..., \boldsymbol{M}^{(K)}\right\}$. This is a contradiction with the assumption that $\boldsymbol{M}$ is QCRB sat.—therefore $\boldsymbol{M}^{(l)}$ is QCRB sat. for $\tilde{\rho}_\theta^{(l)}$. ■

Theorem 4. Function $\chi[\boldsymbol{M}]$ defined in (11) is non-decreasing with l_K and non-increasing with l_1 for fixed $F_C[\boldsymbol{M}]$ (when l_1, l_K, $F_C[\boldsymbol{M}]$ are treated as independent variables).

Proof. Let us introduce a notation

$$G[\boldsymbol{N}, l_1, ..., l_K] = \sum_{i=1}^{K} \mathrm{Tr}(A_i N_i), \quad A_i = l_i^2 \rho_\theta - 2 l_i \dot{\rho}_\theta, \tag{D20}$$

which is a more precise version of (7). We assume that ρ_θ and $\dot{\rho}_\theta$ are fixed, so we do not write them down explicitly as arguments of G. Let $l_1 \leq ... \leq l_K \leq l_{K+1}$ be an arbitrary ascending sequence of real numbers. For any K-element POVM $\boldsymbol{N} = \{N_1, ..., N_K\}$ it is possible to construct a $K+1$-element POVM $\tilde{\boldsymbol{N}}$ satisfying

$$G\left[\boldsymbol{N}, l_1, ..., l_K\right] = G\left[\tilde{\boldsymbol{N}}, l_1, ..., l_K, l_{K+1}\right], \tag{D21}$$

just by defining $\tilde{\boldsymbol{N}} = \{N_1, ..., N_K, 0\}$. Consequently,

$$\max_{\boldsymbol{N} \in \mathcal{M}} G\left[\boldsymbol{N}, l_1, ..., l_K\right] \leq \max_{\tilde{\boldsymbol{N}} \in \mathcal{M}} G\left[\tilde{\boldsymbol{N}}, l_1, ..., l_K, l_{K+1}\right]. \tag{D22}$$

After applying (C6) to both sides of this inequality, we obtain

$$\frac{1}{2}(l_1^2 + l_K^2 + \|A_1 - A_K\|_1) \leq \frac{1}{2}(l_1^2 + l_{K+1}^2 + \|A_1 - A_{K+1}\|_1), \tag{D23}$$

which means that for a fixed $F_C[\boldsymbol{M}]$, $\chi[\boldsymbol{M}]$ is non-decreasing with l_{K+1} because real number $l_K \leq l_{K+1}$ can be arbitrary. Analogously, we can prove that for any pair $l_0 \leq l_1$,

$$\frac{1}{2}(l_1^2 + l_K^2 + \|A_1 - A_K\|_1) \leq \frac{1}{2}(l_0^2 + l_K^2 + \|A_0 - A_K\|_1), \tag{D24}$$

from which follows the 2nd part of our thesis. ■

Appendix E: The most robust among the most informative measurements—Examples

1. Pure state models

Let $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, $\dot{\rho}_\theta = |\psi_\theta\rangle\langle\dot{\psi}_\theta| + |\dot{\psi}_\theta\rangle\langle\psi_\theta|$. Then, the symmetric logarithmic derivative matrix is $\Lambda_\theta = 2\left(|\psi_\theta\rangle\langle\dot{\psi}_\theta| + |\dot{\psi}_\theta\rangle\langle\psi_\theta|\right)$, and the QFI is

$$F_Q = \mathrm{Tr}(\rho_\theta \Lambda_\theta^2) = 4\left(\langle\dot{\psi}_\theta|\dot{\psi}_\theta\rangle - \left|\langle\dot{\psi}_\theta|\psi_\theta\rangle\right|^2\right). \tag{E1}$$

For a fixed value of a parameter $\theta = \theta_0$ we define

$$|0\rangle = \frac{1}{\sqrt{2}}|\psi_{\theta_0}\rangle + i\sqrt{2F_Q^{-1}}\left(|\dot{\psi}_{\theta_0}\rangle - \langle\psi_{\theta_0}|\dot{\psi}_{\theta_0}\rangle|\psi_{\theta_0}\rangle\right), \tag{E2}$$

$$|1\rangle = \frac{1}{\sqrt{2}}|\psi_{\theta_0}\rangle - i\sqrt{2F_Q^{-1}}\left(|\dot{\psi}_{\theta_0}\rangle - \langle\psi_{\theta_0}|\dot{\psi}_{\theta_0}\rangle|\psi_{\theta_0}\rangle\right). \tag{E3}$$

It is straightforward to check that $|0\rangle$ and $|1\rangle$ form o.-n. basis of span$\left\{|\psi_{\theta_0}\rangle, |\dot{\psi}_{\theta_0}\rangle\right\}$, the representations of ρ_{θ_0} , $\dot{\rho}_{\theta_0}$ and Λ_{θ_0} in this basis are:

$$\rho_{\theta_0} = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = |+\rangle\langle+|, \quad \dot{\rho}_{\theta_0} = \frac{1}{2}\sqrt{F_Q}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = \frac{1}{2}\sqrt{F_Q}\sigma_y, \quad \Lambda_{\theta_0} = \sqrt{F_Q}\sigma_y. \tag{E4}$$

The condition (D1) from Theorem 1 says, that all elements M_i of any QCRB sat. POVM must satisfy

$$M_i^{1/2} L_{11} M_i^{1/2} = 0, \tag{E5}$$

where

$$L_{11} = |\psi_{\theta_0}\rangle\langle\psi_{\theta_0}|\,\Lambda_{\theta_0} - \Lambda_{\theta_0}\,|\psi_{\theta_0}\rangle\langle\psi_{\theta_0}| = i\sqrt{F_Q}\sigma_z \tag{E6}$$

in the basis $|0\rangle, |1\rangle$. After taking the trace of n-th power of both sides of E5, we obtain

$$\mathrm{Tr}\left((M_i L_{11})^n\right) = 0. \tag{E7}$$

The most general parametrization of M_i is

$$M_i = \begin{bmatrix} \alpha_i & \beta_i - i\gamma_i \\ \beta_i + i\gamma_i & \delta_i \end{bmatrix}, \tag{E8}$$

where $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$. Using E7 with $n = 1$ and $n = 2$, we obtain the following conditions:

$$n = 1 : \alpha_i = \delta_i, \quad n = 2 : \det M_i = 0. \tag{E9}$$

From the 2nd condition, one eigenvalue of M_i must be 0, so M_i is rank-one, and can be written as

$$M_i = \lambda_i |\phi_i\rangle\langle\phi_i|, \quad |\phi_i\rangle = \begin{bmatrix} a_i \\ b_i \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ e^{i\varphi_i} \end{bmatrix}, \tag{E10}$$

$0 \leq \lambda_i \leq 1$, $|a_i| = |b_i| = 1/\sqrt{2}$ because $\alpha_i = \delta_i$ and because of normalization of $|\phi_i\rangle$. We obtained necessary conditions for QCRB saturation, but it turns out, that they are also sufficient—all POVMs, whose elements are of the described form, saturate QCRB. To show this, let us compute the CFI explicitly:

$$F_C = \sum_i \lambda_i \frac{\left(\langle\phi_i|\dot{\rho}_{\theta_0}|\phi_i\rangle\right)^2}{\langle\phi_i|\rho_{\theta_0}|\phi_i\rangle} = \frac{F_Q}{2} \sum_i \lambda_i \left(1 - a_i\bar{b}_i - \bar{a}_i b_i\right) = F_Q, \tag{E11}$$

where we used the identity $\sum_i M_i = \mathbb{1}$, from which follows that $\sum_i \lambda_i = \mathrm{Tr}(\mathbb{1}) = 2$, and $\sum_i \lambda_i \bar{a}_i b_i = \sum_i \lambda_i a_i \bar{b}_i = 0$. To conclude, for a given parametrization of a family of pure states, a POVM is QCRB sat. iff it contains only elements proportional to projectors on states from the Bloch sphere's equator.

At this point, we are ready to compare different QCRB sat. measurements according to their FI MENOS. Let us start with projective measurements, i.e. those for which $K = 2$, $\lambda_1 = \lambda_2 = 1$, $\varphi_1 = -\varphi$, $\varphi_2 = \pi - \varphi$, we refer to notation from (E10), $\varphi \in [0, 2\pi]$ corresponds to phase φ which has to be added to the lower arm of Mach-Zender interferometer from Fig. 1 in order to implement the described projective measurement. Logarithmic derivatives associated with this measurement are

$$l_1 = \frac{\mathrm{Tr}(\dot{\rho}_{\theta_0} M_1)}{\mathrm{Tr}(\rho_{\theta_0} M_1)} = \sqrt{F_Q}\tan(\varphi/2), \quad l_2 = \frac{\mathrm{Tr}(\dot{\rho}_{\theta_0} M_2)}{\mathrm{Tr}(\rho_{\theta_0} M_2)} = -\sqrt{F_Q}\tan^{-1}(\varphi/2). \tag{E12}$$

There are only two logarithmic derivatives, so one of them is maximal, and another is minimal. After inserting the values of l_1, l_2 into (11) and computing eigenavalues of $A_1 - A_2$ we obtain (13), from which it follows that $\chi[\boldsymbol{M}] \geq 4$ for all QCRB sat. projective measurements, and the inequality is saturated only for $\varphi = \pi/2$. Let us now show, that $\chi[\boldsymbol{M}] \geq 4$ for all QCRB sat. measurements (not only projective). For $\boldsymbol{M}$ with K elements of the form defined in (E10), we construct a noise POVM

$$\boldsymbol{N} = \{N_1, ..., N_k\}, \quad N_i = \lambda_i \left(\mathbb{1} - |\phi_i\rangle\langle\phi_i|\right). \tag{E13}$$

Then

$$G[\boldsymbol{N}] = \sum_i \mathrm{Tr}(A_i N_i) = \sum_i \lambda_i l_i^2 - p_i l_i^2 + 2\dot{p}_i l_i = F_C + \sum_i \lambda_i l_i^2, \tag{E14}$$

and therefore

$$\chi[\boldsymbol{M}, \boldsymbol{N}] = 2 + F_Q^{-1} \sum_i \lambda_i l_i^2, \tag{E15}$$

where we substituted F_C with F_Q because $\boldsymbol{M}$ is QCRB sat. In our case, $l_i = -\sqrt{F_Q}\tan(\varphi_i/2)$, so

$$\chi[\boldsymbol{M}, \boldsymbol{N}] = 2 - \sum_i \lambda_i + \sum_i \lambda_i \cos^{-2}(\varphi_i/2) = 2\sum_i \frac{\lambda_i}{2} \left|\cos^{-1}(\varphi_i/2)\right|^2, \tag{E16}$$

where we used the identity $\sum_i \lambda_i = \mathrm{Tr}(\mathbb{1}) = 2$. Using power mean inequality (between powers 2 and -2, $\lambda_i/2$ are weights), we obtain

$$\sqrt{\sum_i \frac{\lambda_i}{2} \left|\cos^{-1}(\varphi_i/2)\right|^2} \geq \left(\sum_i \frac{\lambda_i}{2} \left|\cos^{-1}(\varphi_i/2)\right|^{-2}\right)^{-1/2} = \sqrt{2}, \tag{E17}$$

the last equality follows from the fact that $\sum_i \lambda_i \cos^2(\varphi_i/2) = \frac{1}{4}\sum_i \lambda_i (2 + e^{i\varphi_i} + e^{-i\varphi_i}) = 1$ because the sum over i of anti-diagonal terms of M_i, which are $\lambda_i e^{\pm i\varphi_i}$, is equal to 0. After inserting (E17) into (E16), and using an inequality $\chi[\boldsymbol{M}] \geq \chi[\boldsymbol{M}, \boldsymbol{N}]$, we obtain that $\chi[\boldsymbol{M}] \geq 4$ for any QCRB sat. $\boldsymbol{M}$. This inequality can only be saturated if (E17) is saturated, which means that all terms in the average are equal, and therefore $\cos^2(\varphi_1/2) = ... = \cos^2(\varphi_K/2)$. Then we have $1 = \sum_i \lambda_i \cos^2(\varphi_i/2) = \cos^2(\varphi_1/2)\sum_i \lambda_i = 2\cos^2(\varphi_1/2)$, so $\cos^2(\varphi_1/2) = ... = \cos^2(\varphi_K/2) = 1/2$, and consequently $\varphi_i = \pm\pi/2$ for each i. Therefore, $\chi[\boldsymbol{M}] = 4$ iff the measurement is a projective measurement on eigenstates of σ_y.

2. Super-resolution optical imaging

Let us firstly follow the construction from [3] of vectors $|0\rangle_s$,$|1\rangle_s$,$|0\rangle_a$,$|0\rangle_a$, where $|0/1\rangle_s$ form o.-n. basis of $\mathcal{H}_s$, $|0/1\rangle_a$ form o.-n. basis of $\mathcal{H}_a$, $\mathcal{H}^{(\theta)} = \mathcal{H}_s \oplus \mathcal{H}_a$, for a fixed value of θ, $\rho_\theta, \dot{\rho}_\theta \in \mathcal{L}\left(\mathcal{H}^{(\theta)}\right)$. We define

$$|0\rangle_s = \frac{1}{\sqrt{2(1+\delta)}}\left(|u_{+,\theta}\rangle + |u_{-,\theta}\rangle\right), \tag{E18}$$

$$|1\rangle_s = \frac{1}{c_4}\left[-\sqrt{2}\left(|\dot{u}_{+,\theta}\rangle + |\dot{u}_{-,\theta}\rangle\right) + \frac{\gamma}{\sqrt{1+\delta}}|0\rangle_s\right], \tag{E19}$$

$$|0\rangle_a = \frac{1}{\sqrt{2(1-\delta)}}\left(|u_{+,\theta}\rangle - |u_{-,\theta}\rangle\right), \tag{E20}$$

$$|1\rangle_a = \frac{1}{c_3}\left[-\sqrt{2}\left(|\dot{u}_{+,\theta}\rangle - |\dot{u}_{-,\theta}\rangle\right) - \frac{\gamma}{\sqrt{1-\delta}}|0\rangle_a\right], \tag{E21}$$

where dot over vector denotes its derivative over θ, and

$$\delta = \langle u_{+,\theta}|u_{-,\theta}\rangle = e^{-\theta^2/8\sigma^2}, \tag{E22}$$

$$\gamma = 2\langle u_{+,\theta}|\dot{u}_{-,\theta}\rangle = -\frac{\theta e^{-\frac{\theta^2}{8\sigma^2}}}{4\sigma^2}, \tag{E23}$$

$$c_3 = \sqrt{4\langle \dot{u}_{-,\theta}|\dot{u}_{-,\theta}\rangle - 4\langle \dot{u}_{+,\theta}|\dot{u}_{-,\theta}\rangle - \frac{\gamma^2}{1-\delta}} = \frac{1}{4}\sqrt{\frac{8\sigma^2\sinh\left(\frac{\theta^2}{8\sigma^2}\right) - \theta^2}{\sigma^4\left(e^{\frac{\theta^2}{8\sigma^2}} - 1\right)}}, \tag{E24}$$

$$c_4 = \sqrt{4\langle \dot{u}_{-,\theta}|\dot{u}_{-,\theta}\rangle + 4\langle \dot{u}_{+,\theta}|\dot{u}_{-,\theta}\rangle - \frac{\gamma^2}{1+\delta}} = \frac{1}{4}\sqrt{\frac{8\sigma^2\sinh\left(\frac{\theta^2}{8\sigma^2}\right) + \theta^2}{\sigma^4\left(e^{\frac{\theta^2}{8\sigma^2}} + 1\right)}}. \tag{E25}$$

Direct computations show, that $|0\rangle_s$, $|1\rangle_s$,$|0\rangle_a$,$|1\rangle_a$ indeed form o.-n. basis, and ρ_θ, $\dot{\rho}_\theta$ can be expressed in this basis as

$$\rho_\theta = \frac{1}{2}\begin{bmatrix} 1+\delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\delta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ \dot{\rho}_\theta = \frac{1}{2}\begin{bmatrix} \gamma & -\frac{c_4}{2}\sqrt{1+\delta} & 0 & 0 \\ -\frac{c_4}{2}\sqrt{1+\delta} & 0 & 0 & 0 \\ 0 & 0 & -\gamma & -\frac{c_3}{2}\sqrt{1-\delta} \\ 0 & 0 & -\frac{c_3}{2}\sqrt{1-\delta} & 0, \end{bmatrix} \tag{E26}$$

which is equivalent to (15), (16) after substituting

$$\alpha = \frac{\gamma}{2}, \ \beta_s = -\frac{c_4}{4}\sqrt{1+\delta}, \ \beta_a = -\frac{c_3}{4}\sqrt{1-\delta}. \tag{E27}$$

Both ρ_θ and $\dot{\rho}_\theta$ have block-diagonal structure (as in Theorem 3):

$$\rho_\theta = \rho_\theta^{(s)} + \rho_\theta^{(a)}, \ \dot{\rho}_\theta = \dot{\rho}_\theta^{(s)} + \dot{\rho}_\theta^{(a)}, \tag{E28}$$

normalized components of ρ_θ are $\tilde{\rho}_\theta^{(s/a)} = |0\rangle_{s/a}\langle 0|$, their derivatives: $\dot{\tilde{\rho}}_\theta^{(s)} = \frac{2\beta_s}{1+\delta}\sigma_x^{(s)}$, $\dot{\tilde{\rho}}_\theta^{(a)} = \frac{2\beta_a}{1-\delta}\sigma_x^{(a)}$. Therefore, the evolution of the normalized density matrix in each subspace is locally equivalent to rotation of a pure state $|0\rangle$ around y axis of a Bloch sphere. Consequently, taking into account Theorem 3 and the characterization of QCRB sat. measurements for pure states from Section E 1, we conclude that potential candidates for minimal MENOS QCRB sat. measurements are of the form

$$\boldsymbol{M} = \left\{\boldsymbol{M}^{(s)}, \boldsymbol{M}^{(a)}\right\}, \ \boldsymbol{M}^{(s)} = \left\{M_1^{(s)}, ..., M_{K_1}^{(s)}\right\}, \ \boldsymbol{M}^{(a)} = \left\{M_1^{(a)}, ..., M_{K_2}^{(a)}\right\} \tag{E29}$$

where $M_i^{(s/a)} = \lambda_{i,s/a} |\varphi_{i,s/a}\rangle_{s/a}\langle\varphi_{i,s/a}|$, $|\varphi\rangle_{s/a} = \cos(\varphi/2)|0\rangle_{s/a} + \sin(\varphi/2)|1\rangle_{s/a}$. In this case, QCRB sat. measurements in each subspace consist on projectors on states from a meridian of a Bloch sphere, not from equator, because parametrization is different from the one used in Section E 1. As the last step, let us prove that $\boldsymbol{M}$ with a minimal MENOS is of the form (17), which means that $K_1 = K_2 = 2$. We will prove,that if $K_1 \geq 3$, then it is possible to construct a QCRB sat. measurement $\tilde{\boldsymbol{M}} = \left\{\tilde{\boldsymbol{M}}^{(s)}, \boldsymbol{M}^{(a)}\right\}$ such that $\chi[\tilde{\boldsymbol{M}}] \leq \chi[\boldsymbol{M}]$ and $\tilde{\boldsymbol{M}}^{(s)}$ contains $K_1 - 1$ elements. Let us start with choosing two indices $i, j \in \{1, ..., K_1\}$ such that

$$\varphi_{i,s} \leq \varphi_{j,s} \ \wedge \ (\varphi_{i,s}, \varphi_{j,s} \in [0, \pi] \vee \varphi_{i,s}, \varphi_{j,s} \in [\pi, 2\pi]). \tag{E30}$$

Such a choice is always possible for $K_1 \geq 3$ because of pigeonhole principle. After some algebra, we obtain the following identity

$$M_i^{(s)} + M_j^{(s)} = A\mathbb{1}_s + B|\tilde{\varphi}\rangle_s\langle\tilde{\varphi}|, \tag{E31}$$

where

$$B = \sqrt{\lambda_{i,s}^2 + \lambda_{j,s}^2 + 2\lambda_{i,s}\lambda_{j,s}\cos(\varphi_{j,s} - \varphi_{i,s})}, \quad A = \frac{1}{2}(\lambda_{i,s} + \lambda_{j,s} - B) \tag{E32}$$

are positive constants, $A < 1$ because $\lambda_{i,s} + \lambda_{j,s} < \mathrm{Tr}(\mathbb{1}_s) = 2$, and

$$\tan(\tilde{\varphi}) = \frac{C\tan(\varphi_{i,s}) + D\tan(\varphi_{j,s})}{C + D}, \quad C = \lambda_{i,s}\cos^{-1}(\varphi_{j,s}), \quad D = \lambda_{j,s}\cos^{-1}(\varphi_{i,s}). \tag{E33}$$

From the above formula it is clear that $\tan(\tilde{\varphi})$ lies between $\tan(\varphi_{i,s})$ and $\tan(\varphi_{j,s})$, so using (E30), we obtain $\tilde{\varphi} \in [\varphi_{i,s}, \varphi_{j,s}]$. After removing $M_i^{(s)}$ and $M_j^{(s)}$ from $\boldsymbol{M}^{(s)}$, and replacing them with $B|\tilde{\varphi}\rangle_s\langle\tilde{\varphi}|$, the sum of all elements is decreased from $\mathbb{1}_s$ to $(1 - A)\mathbb{1}_s$, according to (E31). Therefore, to obtain a valid POVM we need to rescale all elements after such a replacement, and then we end up with

$$\tilde{\boldsymbol{M}}^{(s)} = \left\{(1 - A)^{-1} M_k^{(s)}\Big|_{k\in\{1,...,K_1\}\setminus\{i,j\}}, (1 - A)^{-1} B|\tilde{\varphi}\rangle_s\langle\tilde{\varphi}|\right\}, \tag{E34}$$

which is a POVM containing $K_1 - 1$ elements. From the construction, $\tilde{\boldsymbol{M}}^{(s)}$ contains only elements proportional to projectors on a Bloch sphere meridian, so it is QCRB sat. Let us now prove, that for a given construction, $\chi[\tilde{\boldsymbol{M}}] \leq \chi[\boldsymbol{M}]$. The logarithmic derivative of $p_\theta(k, s) = \mathrm{Tr}\left(\rho_\theta M_k^{(s)}\right)$ is

$$l_{k,s} = \frac{\mathrm{Tr}\left(\dot{\rho}_\theta M_k^{(s)}\right)}{\mathrm{Tr}\left(\rho_\theta M_k^{(s)}\right)} = f(\varphi_{k,s}), \text{ where } f(\varphi) = \frac{2}{1+\delta}(\alpha + 2\beta_s\tan(\varphi/2)), \tag{E35}$$

notice that $l_{k,s}$ does not depend on $\lambda_{k,s}$, and that $f(\varphi)$ is monotonic in range $[0, \pi]$ and in range $[\pi, 2\pi]$. That means, that after replacing $\boldsymbol{M}^{(s)}$ with $\tilde{\boldsymbol{M}}^{(s)}$, logarithmic derivatives of outcomes with indices $k \in \{1, ..., K_1\} \setminus \{i, j\}$ are not affected, and $l_{i,s}$, $l_{j,s}$ are replaced with a logarithmic derivative $\tilde{l} = f(\tilde{\varphi})$. From inclusion $\tilde{\varphi} \in [\varphi_{i,s}, \varphi_{j,s}]$, we conclude that $\tilde{l}$ lies between $l_{i,s}$ and $l_{j,s}$. This means, that the described modification of $\boldsymbol{M}$ neither increased the largest logarithmic derivative, nor decreased the smallest one. Therefore, according to Theorem 4, $\chi[\tilde{\boldsymbol{M}}] \leq \chi[\boldsymbol{M}]$. It means, that we can lower the value of K_1 without affecting the CFI, and without increasing MENOS, as long as $K_1 > 2$. The same reasoning can be applied to the assymetric subspace, so we can also lower K_2. Therefore, we can always find a QCRB sat. measurement with a minimal MENOS with $K_1 = K_2 = 2$, and such a measurement must be of the form defined in (17).

In the main text, we also considered the family of measurements in Hermite-Gaussian modes, $M_i = |\phi_i\rangle\langle\phi_i|$ for $i \in \{1, 2, ..., K-1\}$, $M_K = \mathbb{1} - M_1 - ... - M_{K-1}$, where the representation of $|\phi_i\rangle$ in a position basis is

$$\langle x|\phi_i\rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{1/4} \frac{1}{\sqrt{2^i i!}} H_q\left(\frac{x}{\sqrt{2}\sigma}\right) \exp\left(-\frac{x^2}{4\sigma^2}\right). \tag{E36}$$

According to Ref.[3], the probability of obtaining i-th outcome for an input state defined in (14) is

$$p_\theta(i) = \langle\phi_i|\rho_\theta|\phi_i\rangle = \exp\left(-\theta^2/16\sigma^2\right) \frac{\left(\theta^2/16\sigma^2\right)^i}{i!} \quad \text{for } i \in \{1, ..., K-1\}, \quad p_\theta(K) = 1 - p_1 - ... - p_{K-1}, \tag{E37}$$

corresponding logarithmic derivatives are

$$l_i = -\frac{\theta}{8\sigma^2} + \frac{2i}{\theta} \quad \text{for } i \in \{1, ..., K-1\}, \quad l_K = -\frac{l_1 p_1 + ... + l_{K-1} p_{K-1}}{p_K}. \tag{E38}$$

Both CFI and FI MENOS plotted in Fig. 2 are computed directly using above relations and formulas (3) and (11).

Appendix F: The role of correlations in super-resolution imaging

For some types of sources, subsequent photon emissions are not independent, and then many-photon density matrix is not separable, i.e. $\rho_\theta^{(N)} \neq \left(\rho_\theta^{(1)}\right)^{\otimes N}$, where $\rho_\theta^{(1)}$ is defined in (14). Then, the CFI and QFI for $\rho_\theta^{(N)}$ are not necessarily N times larger than for $\rho_\theta^{(1)}$, so the results obtained using the one-photon model cannot be applied directly. It was shown, that correlations between subsequent photons, resulting from super-bunching or anti-bunching phenomena, can be used to increase the imaging resolution within the *direct imaging* paradigm, where each photon in the image plane is measured in the position basis [4–6].

It is natural to ask, if the advantage from correlations can be also seen if we do not restrict ourselves to direct imaging, but rather assume, that any quantum measurement can be performed. To get some intuition regarding this problem, let us consider a simple toy-model involving correlations—we assume, that sources always emit photons in pairs and the time between subsequent pairs is much longer than the time between emissions of two photons within one pair. Consequently, each pair comes from one source, but we do not know from which one, and the state of a photon pair is

$$\tilde{\rho}_\theta^{(2)} = \frac{1}{2}\left(|u_{+,\theta}^{(2)}\rangle\langle u_{+,\theta}^{(2)}| + |u_{-,\theta}^{(2)}\rangle\langle u_{-,\theta}^{(2)}|\right), \tag{F1}$$

where $|u_{\pm,\theta}^{(2)}\rangle = |u_{\pm,\theta}\rangle^{\otimes 2}$. This density matrix has the same structure as $\rho_\theta^{(1)}$, so we can construct 4-element o.-n. basis spanning $\tilde{\rho}_\theta^{(2)}$ and $\partial_\theta\tilde{\rho}_\theta^{(2)}$, as in (E18–E21):

$$|\tilde{0}\rangle_s = \frac{1}{\sqrt{2(1+\tilde{\delta})}}\left(|u_{+,\theta}^{(2)}\rangle + |u_{-,\theta}^{(2)}\rangle\right), \tag{F2}$$

$$|\tilde{1}\rangle_s = \frac{1}{\tilde{c}_4}\left[-\sqrt{2}\left(|\dot{u}_{+,\theta}^{(2)}\rangle + |\dot{u}_{-,\theta}^{(2)}\rangle\right) + \frac{\tilde{\gamma}}{\sqrt{1+\tilde{\delta}}}|\tilde{0}\rangle_s\right], \tag{F3}$$

$$|\tilde{0}\rangle_a = \frac{1}{\sqrt{2(1-\tilde{\delta})}}\left(|u_{+,\theta}^{(2)}\rangle - |u_{-,\theta}^{(2)}\rangle\right), \tag{F4}$$

$$|\tilde{1}\rangle_a = \frac{1}{\tilde{c}_3}\left[-\sqrt{2}\left(|\dot{u}_{+,\theta}^{(2)}\rangle - |\dot{u}_{-,\theta}^{(2)}\rangle\right) - \frac{\tilde{\gamma}}{\sqrt{1-\tilde{\delta}}}|\tilde{0}\rangle_a\right], \tag{F5}$$

where

$$\tilde{\delta} = \langle u^{(2)}_{+,\theta} | u^{(2)}_{-,\theta} \rangle = e^{-\theta^2/4\sigma^2}, \tag{F6}$$

$$\tilde{\gamma} = 2\,\langle u^{(2)}_{+,\theta} | \dot{u}^{(2)}_{-,\theta} \rangle = -\frac{\theta e^{-\frac{\theta^2}{4\sigma^2}}}{2\sigma^2}, \tag{F7}$$

$$\tilde{c}_3 = \sqrt{4\,\langle \dot{u}^{(2)}_{-,\theta} | \dot{u}^{(2)}_{-,\theta} \rangle - 4\,\langle \dot{u}^{(2)}_{+,\theta} | \dot{u}^{(2)}_{-,\theta} \rangle - \frac{\tilde{\gamma}^2}{1-\tilde{\delta}}} = \frac{1}{2}\sqrt{\frac{4\sigma^2 \sinh\left(\frac{\theta^2}{4\sigma^2}\right) - \theta^2}{\sigma^4\left(e^{\frac{\theta^2}{4\sigma^2}} - 1\right)}}, \tag{F8}$$

$$\tilde{c}_4 = \sqrt{4\,\langle \dot{u}^{(2)}_{-,\theta} | \dot{u}^{(2)}_{-,\theta} \rangle + 4\,\langle \dot{u}^{(2)}_{+,\theta} | \dot{u}^{(2)}_{-,\theta} \rangle - \frac{\tilde{\gamma}^2}{1+\tilde{\delta}}} = \frac{1}{2}\sqrt{\frac{4\sigma^2 \sinh\left(\frac{\theta^2}{4\sigma^2}\right) + \theta^2}{\sigma^4\left(e^{\frac{\theta^2}{4\sigma^2}} + 1\right)}}. \tag{F9}$$

The expressions for $\tilde{\rho}^{(2)}_\theta$ and $\partial_\theta \tilde{\rho}^{(2)}_\theta$ in the introduced basis are the same as the expressions for $\rho^{(1)}_\theta$ and $\dot{\rho}^{(1)}_\theta$ (E26), the only difference is that δ, γ, c_3, c_4 are replaced with $\tilde{\delta}$, $\tilde{\gamma}$, $\tilde{c}_3$, $\tilde{c}_4$. After comparing Eqs. (E22–E25) with Eqs. (F6–F9) we see, that formulas for $\tilde{\rho}^{(2)}_\theta$ and $\partial_\theta \tilde{\rho}^{(2)}_\theta$ are obtained by replacing σ with $\sigma/\sqrt{2}$ in formulas for $\rho^{(1)}_\theta$ and $\dot{\rho}^{(1)}_\theta$. Therefore, two-photon model described by $\tilde{\rho}^{(2)}_\theta$ is equivalent to one-photon model $\rho^{(1)}_\theta$ with σ narrowed by a factor $\sqrt{2}$. Using this observation, and the fact that $F_Q[\rho^{(1)}_\theta] = \frac{1}{4\sigma^2}$, we obtain

$$F_Q[\tilde{\rho}^{(2)}_\theta] = \frac{1}{4(\sigma/\sqrt{2})^2} = 2F_Q[\rho^{(1)}_\theta] \tag{F10}$$

The density matrix describing two uncorrelated photons is $\rho^{(2)}_\theta = \rho^{(1)}_\theta \otimes \rho^{(1)}_\theta$, and from additivity of QFI, we have $F_Q[\rho^{(2)}_\theta] = 2F_Q[\rho^{(1)}_\theta] = F_Q[\tilde{\rho}^{(2)}_\theta]$. This may suggest, that super-bunching based correlations of the considered type do not provide any increase of the imaging precision, when all quantum measurements are allowed. However, this is only true for a noiseless scenario—let us check, how small measurement noise affects the esimation precision in both cases—with and without correlations.

For the correlated model described by $\tilde{\rho}^{(2)}_\theta$, we consider arbitrary (possibly correlated) measurement $\boldsymbol{M}^{(2)}$, which can be affected by arbitrary noise $\boldsymbol{N}^{(2)}$, such that noise-affected two-photon POVM is $\tilde{\boldsymbol{M}}^{(2)} = (1-\epsilon)\boldsymbol{M}^{(2)} + \epsilon \boldsymbol{N}^{(2)}$, and then the resulting noise-affected CFI is, up to terms of the order of ϵ,

$$F_C[\tilde{\rho}^{(2)}_\theta, \tilde{\boldsymbol{M}}^{(2)}] = F_C[\tilde{\rho}^{(2)}_\theta, \boldsymbol{M}^{(2)}]\left(1 - \epsilon\chi[\boldsymbol{M}^{(2)}, \boldsymbol{N}^{(2)}]\right). \tag{F11}$$

The lowest possible FI MENOS of QCRB sat. measurement for this model is

$$\tilde{\chi}^{(2)}_Q\left(\frac{\theta}{\sigma}\right) = \min_{\left\{\boldsymbol{M}^{(2)} \in \mathcal{M}, F_C\left[\boldsymbol{M}^{(2)}\right] = F_Q\right\}} \chi\left[\boldsymbol{M}^{(2)}\right]. \tag{F12}$$

Fortunately, we do not need to repeat the whole procedure from Section E 2 to calculate $\tilde{\chi}^{(2)}_Q\left(\frac{\theta}{\sigma}\right)$ because models described by $\rho^{(1)}_\theta$ and $\tilde{\rho}^{(2)}_\theta$ are isomorphic, and the only difference between them is that parameter σ is decreased by a factor $\sqrt{2}$ in the latter one. Consequently,

$$\tilde{\chi}^{(2)}_Q\left(\frac{\theta}{\sigma}\right) = \chi^{(1)}_Q\left(\frac{\theta}{\sigma\sqrt{2}}\right), \tag{F13}$$

where $\chi^{(1)}_Q$ is χ_Q computed for $\rho^{(1)}_\theta$.

Let us now examine the scenario with uncorrelated photons ($\rho^{(2)}_\theta = \rho^{(1)}_\theta \otimes \rho^{(1)}_\theta$) and uncorrelated one-photon measurements ($\boldsymbol{M}^{(2)} = \boldsymbol{M}^{(1)} \otimes \boldsymbol{M}^{(1)}$). We choose $\boldsymbol{M}^{(1)}$ to be the lowest susceptibility QCRB sat. measurement for $\rho^{(1)}_\theta$, found in Section E 2, and compute $\chi^{(2)}_{\text{loc}} = \chi[\boldsymbol{M}^{(1)} \otimes \boldsymbol{M}^{(1)}]$ directly from (11). As we see in Fig. 1, $\chi^{(2)}_{\text{loc}} > \chi^{(1)}$. This is because $\chi[\boldsymbol{M}^{(1)} \otimes \boldsymbol{M}^{(1)}]$ is computed by maximizing $\chi[\boldsymbol{M}^{(1)} \otimes \boldsymbol{M}^{(1)}, \boldsymbol{N}^{(2)}]$ over all noise POVMs $\boldsymbol{N}^{(2)}$, also non-local ones. Therefore, the maximization is taken over a larger class of noise than for one-photon model.

It is often reasonable to assume, that for uncorrelated measurements on uncorrelated systems, the same noise affects each one-photon POVM independently, such that $\boldsymbol{M}^{(1)}$ becomes $\tilde{\boldsymbol{M}}^{(1)} = (1-\epsilon)\boldsymbol{M}^{(1)} + \epsilon\boldsymbol{N}^{(1)}$, and consequently $\boldsymbol{M}^{(2)}$ becomes $\tilde{\boldsymbol{M}}^{(2)} = \tilde{\boldsymbol{M}}^{(1)} \otimes \tilde{\boldsymbol{M}}^{(1)}$. Then, the noise affected CFI is

$$F_C[\rho^{(2)}_\theta, \tilde{\boldsymbol{M}}^{(2)}] = 2F_C[\rho^{(1)}_\theta, \tilde{\boldsymbol{M}}^{(1)}] = 2F_C[\rho^{(1)}_\theta, \boldsymbol{M}^{(1)}]\left(1 - \epsilon\chi[\boldsymbol{M}^{(1)}, \boldsymbol{N}^{(1)}]\right) = F_C[\rho^{(2)}_\theta, \boldsymbol{M}^{(2)}]\left(1 - \epsilon\chi[\boldsymbol{M}^{(1)}, \boldsymbol{N}^{(1)}]\right), \tag{F14}$$

terms of the order of ϵ^2 were omitted. When $\boldsymbol{M}^{(1)}$ is the lowest susceptibility QCRB sat. measurement, then

$$F_C[\rho_\theta^{(2)}, \tilde{\boldsymbol{M}}^{(1)} \otimes \tilde{\boldsymbol{M}}^{(1)}] \geq F_Q[\rho^{(2)}](1 - \epsilon\chi_Q^{(1)}) \tag{F15}$$

We see, that the optimal noise susceptibility for two non-correlated photons is $\chi_Q^{(1)}$, assuming non-correlated QCRB sat. measurement and non-correlated noise. Notice, that when ϵ-small noise $\boldsymbol{N}^{(1)}$ acts on each photon, then two photons are affected by 2ϵ-small noise because

$$\tilde{\boldsymbol{M}}^{(1)} \otimes \tilde{\boldsymbol{M}}^{(1)} = (1 - 2\epsilon)\boldsymbol{M}^{(1)} \otimes \boldsymbol{M}^{(1)} + 2\epsilon\boldsymbol{N}^{(2)}, \tag{F16}$$

where $\boldsymbol{N}^{(2)} = \frac{1}{2}(\boldsymbol{M}^{(1)} \otimes \boldsymbol{N}^{(1)} + \boldsymbol{N}^{(1)} \otimes \boldsymbol{M}^{(1)})$ is a two-photon noise POVM. However, from comparison between $\chi_Q^{(1)}$ and $\chi_{\text{loc}}^{(2)}$ follows, that the effect of non-correlated 2ϵ-noise is weaker than the effect of ϵ-small correlated noise.

As we see in Fig. 1, we have $\tilde{\chi}_Q^{(2)} < \chi_{\text{loc}}^{(2)}$. Consequently, even though correlations do not increase the QFI, they allow to obtain better robustness against noise, and consequently better precision of the estimation of θ in a noisy environment. The advantage for $\theta < 2\sigma$ is present even if we assume that for the non-correlated model, noise also has to be non-correlated, because then $\tilde{\chi}_Q^{(2)} < \chi_Q^{(1)}$.

To sum up, we considered a simple model of two-photon correlations, and demonstrated their usefulness in achieving better precision of the separation estimation between two sources in the presence of noise. Further research is required to study more realistic correlation models, and search for practical, noise-robust protocols utilizing correlations between photons and correlated measurements.

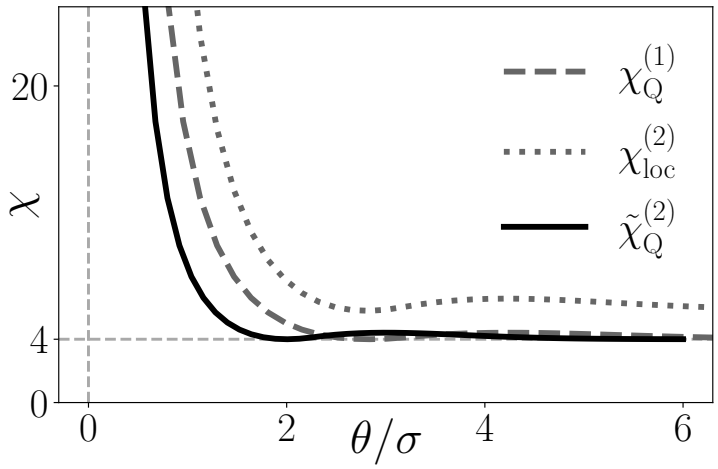


FIG. 1. Two-photon correlations allow to achieve a lower noise susceptibility ($\tilde{\chi}_Q^{(2)}$) than for non-correlated case ($\chi_{\text{loc}}^{(2)}$). Optimal one-photon noise susceptibility ($\chi_Q^{(1)}$) can be interpreted as two-photon noise susceptibility for non-correlated case, when we restrict ourselves to non-correlated noise.

[1] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[2] S. Zhou, C.-L. Zou, and L. Jiang, Quantum Science and Technology **5**, 025005 (2020).
[3] M. Tsang, R. Nair, and X.-M. Lu, Phys. Rev. X **6**, 031033 (2016).
[4] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, Proceedings of the National Academy of Sciences **106**, 22287 (2009).
[5] S. Kurdziałek and R. Demkowicz-Dobrzański, Journal of Optics **23**, 075701 (2021).
[6] O. Schwartz and D. Oron, Physical Review A **85**, 033812 (2012).

Chapter 6

Super-resolution microscopy via fluctuation-enhanced spatial mode demultiplexing

Commentary

This is the least quantum-focused and most experiment-oriented work in this thesis.

Signal correlations resulting from super-Poissonian [Rus06; Der09] and sub-Poissonian [Sch12; Sch13] photon statistics have been proven to enhance direct imaging resolution—this fact is exploited in practical microscopy techniques. Nevertheless, the calculations in Sec. 1.5 of Chapter 1 may suggest that this advantage exists only for direct imaging and vanishes for optimal measurements.

However, we only proved that correlations do not increase the QFI for a simple case of estimating the separation between two equally bright sources with a known centroid. Moreover, no noise was present in our model. The fundamental advantage from correlations may still exist in the presence of noise (as suggested by results in Chapter 5) or for more complex objects. Even for binary objects, correlations allow to slightly increase the QFI for unknown centroid and unequally bright sources [Hor25].

In this work, optical signal correlations are shown to significantly enhance the estimation precision of spatial moments of complex objects. The analysis is done for the H-G modes (1.30) measurement proposed by Tsang [Tsa16]. The higher the estimated spatial moment, the larger the gain from correlations. This aligns with the absence of correlation-induced improvements for the estimation of separation between two points—correlations become important for estimating more refined object features that are absent in simple binary objects.

Even more importantly, correlations resulting from emitter fluctuations enable simplified measurements for extracting information about all spatial moments of an analyzed object. The information contained in higher H-G modes can alternatively be extracted from higher-order temporal statistics of lower modes.

My contribution

Single-author work.

Super-resolution microscopy via fluctuation-enhanced spatial mode demultiplexing

Stanisław Kurdziałek
Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland

We introduce a superresolution technique that combines spatial mode demultiplexing (SPADE) with emitter blinking. We show that temporal fluctuations not only enhance the precision of SPADE imaging, but also drastically simplify the measurement required to recover full object information—in the presence of fluctuations, SPADE can be replaced by the much simpler image inversion interferometry. Both gains are enabled by exploiting temporal cumulants of the detected signal.

I. INTRODUCTION

Light diffraction limits the resolution of conventional light microscopy to $\sim$ 200 nm (Rayleigh criterion [1]). Recently developed super-resolution methods [2–27] overcome this barrier and, under certain conditions, reveal details at nanometers scale[28–30].

Traditional imaging assumes (i) uniform and classical illumination, (ii) a linear, time-invariant and classical response of imaged object, and (iii) measurement of output field with a camera placed in an image plane. Superresolution techniques break at least one of assumptions (i)–(iii), yielding three broad classes.

Category (i) covers structured illumination microscopy (SIM) [22] or image scanning microscopy (ISM) [23], as well as approaches with nonclassical light sources [15, 17, 18].

Category (ii) exploits emitter dynamics [3, 4, 6], non-Poissonian photon statistics [24, 25, 31] or nonlinearity [26, 27]. Stochastic optical reconstruction microscopy (STORM)[4] and photoactivated localization microscopy (PALM) [3], both awarded Nobel Prize in Chemistry in 2014, exploit *blinking*: under uniform excitation, fluorophores randomly switch on and off, leaving only a few active at a time, which enables precise localization. Stochastic optical fluctuation imaging (SOFI) [6, 32] also uses blinking but does not require isolating single emitters; instead, it analyzes temporal cumulants of image sequences. Photon antibunching, which yields sub-Poissonian statistics, can likewise enable super-resolution imaging [24, 25].

Category (iii) is the youngest family of techniques, inspired by Tsang et al. [10], who showed that resolution can be enhanced by replacing intensity measurements with more informative detection of the optical field. In particular, measuring in the Hermite–Gaussian basis is optimal for resolving two closely spaced point sources. This idea is realized by spatial mode demultiplexing (SPADE) measurement, which we use to refer to Tsang's imaging scheme. A closely related, less powerful but often more accessible variant is image inversion interferometry (III) [33, 34], which sorts the field into even and odd spatial modes [this technique is also sometimes called Super Localization by Image inVERsion (SLIVER)]. Superresolution can also be achieved via homodyne or heterodyne detection of the image-plane field [12, 35, 36] or by using Hong-Ou-Mandel effect [14].

Several techniques combine categories (i) and (ii), for example the Nobel Prize-winning stimulated emission depletion (STED) [5] microscopy or SOFISM [37, 38], which merges SOFI and ISM. Yet no method has so far combined the resolution gains of non-Poissonian sources [category (ii)] with measurement strategies of category (iii) such as SPADE.

In this work we introduce stochastic optical fluctuation SPADE (SOFSPADE), which for the first time combines emitter fluctuations with SPADE detection to realize joint resolution gain. Equally important, temporal fluctuations allow to significantly simplify the measurement: instead of full SPADE, one can perform the much simpler III measurement while still recovering information about complex object structure. We will call this technique stochastic optical fluctuation image inversion interferometry (SOFIII).

The paper is organized as follows. In Sec. II we introduce general formalism describing linear optical imaging, and discuss direct imaging (DI) and known superresolution techniques such as SPADE, III, iSPADE and SOFI using this formalism. Section III contains the main result of this work—the working principle of novel SOFSPADE and SOFIII superresolution techniques, with practical recipe for object's spatial moments estimator. Then, in Sec. IV, we show numerical results demonstrating the practical advantage from fluctuations and the optimality of proposed estimator. Additional practical aspects of proposed techniques are discussed in Sec. V, Sec. VI contains the summary and the discussion of the results.

II. OPTICAL IMAGING MODELS

Any linear imaging system is characterized by a transfer function $T(j|\boldsymbol{r})$, which is a probability of detecting a photon emitted from object's position $\boldsymbol{r}$ at detector labeled by j. Consider an object consisting of multiple independent, incoherent, weak ($\ll 1$ photons per temporal mode) sources with positions $\boldsymbol{r}_i$ and constant brightnesses q_i. Each source emits $n_i \sim \mathrm{Poiss}(q_i)$ photons per unit time, where $\mathrm{Poiss}(\mu)$ denotes Poissonian random variable with mean μ. The detectors photon counts n_j are then independent and

$$n_j \sim \mathrm{Poiss}(I_j), \quad I_j = \sum_i T(j|\boldsymbol{r}_i) q_i, \tag{1}$$

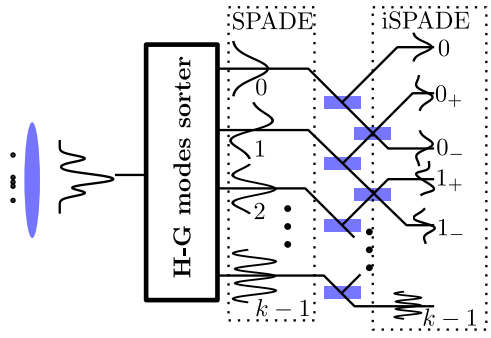


FIG. 1. In SPADE technique, image plane field is decomposed into k H-G modes, which allows to estimate lowest k object's even spatial moments. To estimate odd spatial moments, one needs to interfere neighbouring H-G modes, which leads to iSPADE technique.

where I_j are detectors intensities. Note that we cannot measure I_j exactly, we only have access to shot-noise-affected photon counts n_j. In direct imaging (DI), detectors correspond to camera pixels at positions $\boldsymbol{r}_j$. The resulting transfer function is

$$T^{\mathrm{DI}}(\boldsymbol{r}_j|\boldsymbol{r}) = \eta U(\boldsymbol{r}_j - \boldsymbol{r}), \tag{2}$$

where U is the point spread function (PSF). Coordinates were rescaled such that system magnification is unity, and η denotes signal attenuation factor. From now on we set $\eta = 1$, so that q_i represent total brightnesses observed in the image plane rather than at the object.

We also assume a 1D object, and write positions as scalars x instead of vectors $\boldsymbol{r}$. This simplification highlights the core ideas of introduced imaging techniques, all the results extend to 2D, as we will show in Sec. V A. Furthermore, we set a Gaussian PSF $U(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$, extensions to more realistic PSFs are straightforward [39].

A. Spatial mode demultiplexing imaging (SPADE)

In SPADE imaging technique, we replace the position measurement with the measurement in Hermite-Gaussian (H-G) modes basis [10]

$$\phi_j(x') = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \frac{1}{\sqrt{2^j j!}} H_j\left(\frac{x'}{\sigma\sqrt{2}}\right) \exp\left(-\frac{x'^2}{4\sigma^2}\right), \tag{3}$$

$j \in \{0, 1, 2, ...\}$, H_j are Hermite polynomials, x' is a position in the image plane. This can be implemented by commercially available multi-plane light converters [40–42] or by heterodyning image field with local oscillators with varying space profiles [12, 43–45]—this latter method, however, causes additional signal losses. The resulting transfer function is

$$T^{\mathrm{S}}(j|x) = \frac{1}{4^j j!} e^{-\frac{x^2}{4\sigma^2}} \left(\frac{x}{\sigma}\right)^{2j}, \tag{4}$$

where x is measured respectively to the center of a mode sorting device. Notably, $T(0|0) = 1$, $T(j|0) = 0$ for $j > 0$, so all H-G modes except ϕ_0 are dark for a point source at $x = 0$. Small changes of position and shape then produces signal in these initially dark modes. This is the reason for resolution enhancement achieved by SPADE—it is easier to distinguish zero from nonzero than to resolve small changes atop a large background. This mechanism can be used for exoplanet detection[46]—higher H–G modes are dark for a lone star but become excited for a star with an orbiting planet.

The superresolving power of SPADE can be demonstrated using a minimal example of two identical sources of unknown separation s and a known centroid. For DI, the mean squared error (MSE) of any unbiased separation estimator $\hat{s}$, satisfies, for a fixed photon budget, $\delta^2\hat{s} \overset{s\to 0}{=} \langle(s - \hat{s})^2\rangle \longrightarrow \infty$—this is called a *Rayleigh's curse* [10]. Surprisingly, with SPADE one can construct an estimator for which $\delta^2\hat{s}$ remains finite and constant even for $s \to 0$, enabling precise separation estimation even far below Rayleigh's limit. While noise and mode sorter misalignment reduce SPADE precision [47, 48], it still outperforms DI in sub-Rayleigh regime, which was demonstrated both theoretically [49–52] and experimentally [40–42].

After Taylor expanding (4) and inserting it into (1), we get light intensities on SPADE detectors for a general multi-emitter object,

$$I_j^{\mathrm{S}} = \sum_{\mu=j}^{\infty} A_{\mu,j}^{\mathrm{S}} \theta_{2\mu}, \quad A_{\mu,j}^{\mathrm{S}} = \frac{(-1)^{\mu-j}}{4^\mu j!(\mu - j)!}, \tag{5}$$

where $\theta_\mu = \sum_i q_i (x_i/\sigma)^\mu$ are object's dimensionless spatial moments.

Consequently, object's even spatial moments can be reconstructed by performing SPADE measurement and inverting linear relation (5). In a subdiffraction limit $\Delta = \max_i |x_i/\sigma| \ll 1$, (5) reduces to [53]

$$I_j^{\mathrm{S}} = \frac{1}{4^j j!}\theta_{2j} + \mathcal{O}(\Delta^{2j+2}), \tag{6}$$

so $2j$th moment estimator is simply $\hat{\theta}_{2j} = 4^j j! n_j^{\mathrm{S}}$.

Unfortunately, SPADE only allows to estimate even moments because I_j^{S} do not depend on odd moments. To remedy this, interferometric SPADE (iSPADE) technique was introduced [53], in which a mode sorting device is followed by beam splitters mixing neighbouring modes, as shown in Fig. 1. When the SPADE sorter distinguishes k modes $\phi_0, \phi_1, ..., \phi_{k-1}$, then there are $2k$ iSPADE outputs labeled by $0, j_\pm$ for $j \in \{0, 1, ..., k-2\}$, $k-1$; outputs $j_\pm$ contain interfered modes j and $j+1$. The corresponding transfer functions are

$$T^{\mathrm{iS}}(j|x) = \frac{1}{2} T^{\mathrm{S}}(j|x) \quad \text{for} \quad j \in \{0, k\} \tag{7}$$

$$T^{\mathrm{iS}}(j_\pm|x) = \frac{1}{4^{j+1} j!} e^{-\frac{x^2}{4\sigma^2}} \left(\frac{x}{\sigma}\right)^{2j} \left(1 \pm \frac{1}{2\sqrt{j+1}} \frac{x}{\sigma}\right)^2. \tag{8}$$

These transfer functions depend on odd powers of x, which allows to estimate odd moments—in particular, in the subdiffraction limit [53]

$$I^{\mathrm{iS}}_{j+} + I^{\mathrm{iS}}_{j-} = \frac{1}{2}\frac{1}{4^j j!}\theta_{2j} + \mathcal{O}(\Delta^{2j+2}), \tag{9}$$

$$I^{\mathrm{iS}}_{j+} - I^{\mathrm{iS}}_{j-} = \frac{1}{2}\frac{1}{4^j j!\sqrt{j+1}}\theta_{2j+1} + \mathcal{O}(\Delta^{2j+3}), \tag{10}$$

so $\hat{\theta}_{2j} = 2\cdot 4^j j!(n^{\mathrm{iS}}_{j+} + n^{\mathrm{iS}}_{j-})$, $\hat{\theta}_{2j+1} = 2\cdot 4^j j!\sqrt{j+1}(n^{\mathrm{iS}}_{j+} - n^{\mathrm{iS}}_{j-})$.

Relative errors of moments estimation for SPADE and iSPADE techniques satisfy [53, 54] $\delta\hat{\theta}_\mu/\theta_\mu = \mathcal{O}(\Delta^{-\lceil\frac{\mu}{2}\rceil})$ —this is the optimal scaling with Δ over all measurements allowed by quantum mechanics[55–57]. For DI, $\delta\hat{\theta}_\mu/\theta_\mu = \mathcal{O}(\Delta^{-\mu})$ [53], so (i)SPADE gives a better scaling for $\mu \geq 2$. While these results are derived for $\Delta \to 0$, numerical studies confirm (i)SPADE's optimality in a more realistic $\Delta \approx 1$ regime [58].

In principle, iSPADE enables any object reconstruction by estimating spatial moments and expressing intensity distribution in terms of them. This was demonstrated in a proof-of-principle experiment, with neural networks assisting in processing noisy data [43, 44] .

However, this approach is impractical for objects much larger than the PSF ($\Delta \gg 1$) since large number of moments is needed, and the advantage of (i)SPADE over DI is not clear in this regime [59]. It is therefore much more promising to use (i)SPADE in scanning, confocal microscopy, where the illuminated part of an object is of the size of diffraction spot, so $\Delta \approx 1$. The object can then be scanned, (i)SPADE measurements taken locally, and the full image reconstructed from multiple spatial moment maps using generalized Richardson-Lucy deconvolution algorithm [60] or neural networks [61].

Applying (i)SPADE in practical microscopy with fluorescent emitters and high-NA objectives is technically and conceptually demanding [62]. A simpler and much easier to implement [63, 64] alternative is image inversion interferometry (III), which has already demonstrated superresolution of two fluorescent emitters in a high-NA setup [65]. III sorts the optical field into even and odd modes by interfering the image with its spatial inversion. The corresponding transfer function is

$$T^{\mathrm{III}}(\pm|x) = \frac{1}{2}\left(1 \pm e^{-\frac{x^2}{2\sigma^2}}\right), \tag{11}$$

$+/-$ correspond to even/odd modes. For two point sources, III and SPADE achieve the same precision for $s \to 0$ [34]; however, III becomes worse for larger separations. More importantly, III provides limited information about complex objects; with only two outcomes, only two independent object properties can be inferred. In the subdiffraction regime, intensities of even and odd modes are

$$I^{\mathrm{III}}_{+} = \theta_0 + \mathcal{O}(\Delta^2), \quad I^{\mathrm{III}}_{-} = \frac{\theta_2}{4} + \mathcal{O}(\Delta^4), \tag{12}$$

so only 0th and 2nd spatial moments can be estimated.

An important result of this work, demonstrated in Sec. III A, is that temporal fluctuations enable extraction of all even moments from III, providing the full information of SPADE with a much simpler measurement.

B. Stochastic optical fluctuation imaging (SOFI)

Let us consider super-Poissonian sources with time-fluctuating brightnesses, such as quantum dots [66] or dyes [67] used in fluorescence microscopy. We divide observation time into *frames*. In each frame, the photon number from ith source follows $n_i \sim \mathrm{Poiss}(q_i)$, where q_i is a random variable representing the frame-integrated source brightness. Thus, fluctuations of n_i arise both from variability of q_i and and from shot noise. For multiple sources at positions x_i, the pixel intensity for DI is $I(x_j) = \sum_i q_i U(x_j - x_i)$, which also fluctuates in time. Recording multiple frames yields different samples of a random variable $I(x_j)$.

SOFI [6, 32] exploits these fluctuations to enhance resolution. The rth cumulant of a random variable X is defined as

$$\kappa^{(r)}(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t^r}\right|_{t=0} \log\langle e^{tX}\rangle, \tag{13}$$

for example, $\kappa^{(1)}(X) = \langle X\rangle$, $\kappa^{(2)}(X) = \langle X^2\rangle - \langle X\rangle^2$, $\langle\bullet\rangle$ denotes expected value. For independent q_js

$$\kappa^{(r)}(I(x_j)) = \sum_i \kappa^{(r)}(q_i)U^r(x_j - x_i). \tag{14}$$

This is the core equation of SOFI—in the rth cumulant, the PSF is effectively replaced with its rth power, and consequently narrowed by a factor $\sqrt{r}$. In principle, one can arbitrarily narrow the PSF using cumulants of large orders; however, in practice higher cumulants become very noisy, which limits the resolution gain [68].

Importantly, in practice one measures shot noise-affected photon counts $n(x_j)$, not the noiseless intensities $I(x_j)$, and $\kappa^{(r)}(I(x_j)) \neq \kappa^{(r)}(n(x_j))$ for $r \geq 2$. Therefore, (14) does not hold for experimentally measured cumulants—the discrepancy is especially significant at low photon numbers. This issue was recently addressed in Ref. [31], which showed that resolution enhancement by $\sqrt{r}$ requires using an appropriate linear combination of measured cumulants $\kappa^{(1)}, \kappa^{(2)}, ..., \kappa^{(r)}$ rather than $\kappa^{(r)}$ alone. We will return to this point in Sec. III B when discussing the techniques introduced in the present work.

III. CONNECTING SPADE AND SOFI

A. General principle

It is natural to ask whether intensity fluctuations help to enhance resolution beyond DI measurement. We answer this question positively and introduce SOFSPADE

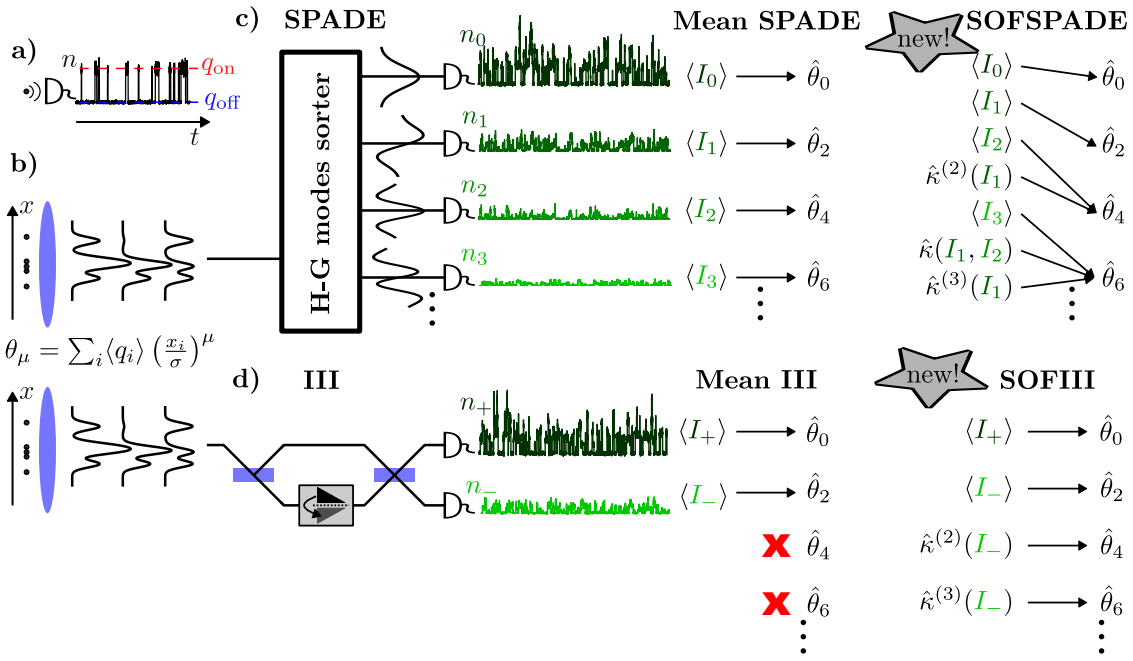


FIG. 2. Photon emission from quantum dots or dyes used to label an object fluctuates in time—typically, an emitter switches between two states characterized by brightnesses $q_{\rm off}$ and $q_{\rm on}$ (a). The resulting image also varies over time (b). In SPADE, the image plane field is sorted into H-G modes (c). Fluctuations of photon counts n_j in different modes provide additional information about object's higher ($\mu \geq 4$) spatial moments through suitable intensity cumulants—this leads to a proposed SOFSPADE technique. In III, the image-plane field is interfered with its spatial inversion (d), yielding a much simpler setup than SPADE, but providing access only to two outcomes corresponding to 0th and 2nd spatial moments. However, temporal fluctuations allow all even moments to be recovered from temporal cumulants of the odd-mode intensity I_-, which forms the basis of the proposed SOFIII technique.

technique which combines resolution gains from SOFI and SPADE.

Let us consider a 1D object consisting of independent fluctuating emitters imaged by a general linear system with a transfer function $T(j|x)$—(1) then applies with q_i and I_j interpreted as random variables.

Emitters fluctuations introduce correlations between detectors, which can be characterized by joint cumulants, defined as

$$\kappa(X_1,...,X_n) = \frac{\mathrm{d}^r}{\mathrm{d}t_1\mathrm{d}t_2...\mathrm{d}t_r}\bigg|_{t_1,...,t_r=0} \log\langle e^{\sum_i t_i X_i}\rangle, \tag{15}$$

for example $\kappa(X_1,X_2) = \langle X_1X_2\rangle - \langle X_1\rangle\langle X_2\rangle$. The general formula for multivariate cumulants is

$$\kappa(X_1,...,X_r) = \sum_\pi \Omega_\pi \prod_{B\in\pi}\left\langle \prod_{i\in B} X_i\right\rangle, \tag{16}$$

where the sum runs over all partitions π of $\{1,...,r\}$ into non-empty blocks B, $\Omega_\pi = (|\pi|-1)!(-1)^{|\pi|-1}$, where $|\pi|$ denotes the number of blocks in a given partition; see appendix A for explicit formulas for low-order cases and further cumulants properties. We denote by $\kappa^{(r_1,...,r_l)}(X_1,...,X_l)$ the joint cumulant in which X_i appears r_i times.

The joint cumulant of detectors intensities is

$$\begin{aligned}
&\kappa(I_{j_1},...,I_{j_r}) = \\
&\quad = \kappa\left(\sum_{i_1} T(j_1|x_{i_1})q_{i_1},...,\sum_{i_r} T(j_l|x_{i_r})q_{i_r}\right) = \\
&\quad \overset{\text{(I)}}{=} \sum_{i_1,...i_r} T(j_1|x_{i_1})...T(j_l|x_{i_r})\kappa(q_{i_1},...,q_{i_r}) = \\
&\quad \overset{\text{(II)}}{=} \sum_i T(j_1|x_i)T(j_2|x_i)...T(j_r|x_i)\kappa^{(r)}(q_i),
\end{aligned} \tag{17}$$

in (I) we used cumulants linearity, in (II) we used the fact that cumulants of independent random variables vanish. We got extension of (14) for arbitrary measurement and arbitrary multi-detectors correlations; for cumulants, effective transfer functions are products of intensity transfer functions.

To proceed further, let us assume that the ratio $\tilde{\kappa}_r = \kappa^{(r)}(q_i)/\langle q_i\rangle$ is the same for all emitters. This assumption is true for statistically identical sources, but also more generally, when each single source consists of multiple independent, identical emitters stick together. Importantly, if blinking dynamics is not known, coefficients $\tilde{\kappa}_r$ can be estimated based on measured signal, see V B for more details. Let $T(j|x) = \sum_\mu A_{\mu,j}(x/\sigma)^\mu$ be Taylor

expansions of transfer functions, then (17) becomes

$$\kappa(I_{j_1}, I_{j_2}, ..., I_{j_r}) = \tilde{\kappa}_r \sum_{\mu} A_{\mu,\vec{j}} \theta_\mu, \quad \vec{j} = (j_1, j_2, ..., j_r),$$
$$A_{\mu,\vec{j}} = \sum_{\mu_1+\mu_2+...+\mu_r=\mu} A_{\mu_1,j_1} A_{\mu_2,j_2} ... A_{\mu_r,j_r}, \quad (18)$$

where $\theta_\mu = \sum_i \langle q_i \rangle (x_i/\sigma)^\mu$ are object's dimensionless spatial moments, previous definition was generalized to fluctuating sources.

Equation (18) shows that intensity cumulants provide additional information about the object's spatial moments. For SPADE measurement, in a subdiffraction limit, (18) reduces to

$$\kappa(I^{\mathrm{S}}_{j_1}, ..., I^{\mathrm{S}}_{j_r}) = \frac{\tilde{\kappa}_r}{\prod_{l=1}^{r} 4^{j_l} j_l!} \theta_{2(j_1+...+j_r)} + \mathcal{O}(\Delta^{2(j_1+...+j_r)+2}), \quad (19)$$

so higher-order moments can be inferred from lower-order H-G modes by analyzing suitable cumulants. For example, $\kappa^{(2)}(I_1) = \frac{\tilde{\kappa}_2}{16}\theta_4 + \mathcal{O}(\Delta^6)$, so the variance of intensity in 1st H-G mode encodes the fourth spatial moment.

This allows to enhance estimation precision by combining information from different cumulants: θ_4 can be estimated using both $\kappa(I_2) = \langle I_2 \rangle$ and $\kappa^{(2)}(I_1)$, θ_6 can be estimated using $\kappa(I_3) = \langle I_3 \rangle$, $\kappa(I_1, I_2)$, $\kappa^{(3)}(I_1)$, etc., see Fig. 2c. This extra information sources lead to higher estimation precision, as we will demonstrate in Sec. IV

Moreover, one can simplify the measurement and reduce the number of distinguished H-G modes while retaining the ability to estimate higher moments; for example, θ_4 can be estimated *only* based on $\kappa^{(2)}(I_1)$, θ_6 based on $\kappa^{(3)}(I_1)$, etc. Therefore, measurement of just two H-G modes, ϕ_0 and ϕ_1, allows to estimate all even moments.

Pursuing this idea further, one can use a significantly simpler III measurement instead of SPADE, and still recover all even moments since cumulants of intensity of odd modes satisfy

$$\kappa^{(r)\mathrm{III}}(I_-) = \frac{\tilde{\kappa}_r}{4^r}\theta_{2r} + \mathcal{O}(\Delta^{2r+2}), \quad (20)$$

so rth cumulant encodes $2r$th spatial moment, see Fig. 2d. This is the basis of our newly introduced SOFIII technique.

For odd moments, one needs to use iSPADE, for which cumulants also allow to gain more information about higher moments. In particular, iSPADE with just four different outcomes—0, 0_+, 0_- and 1—allows to reconstruct all moments, for example, using equations

$$\kappa^{(r)}(I^{\mathrm{iS}}_1) = \frac{\tilde{\kappa}_r}{8^r}\theta_{2r} + \mathcal{O}(\Delta^{2r+2})$$
$$\kappa(I^{\mathrm{iS}(r)}_1, 0_+) - \kappa(I^{\mathrm{iS}(r)}_1, 0_-) = \frac{\tilde{\kappa}_r}{2^{3r+1}}\theta_{2r+1} + \mathcal{O}(\Delta^{2r+3}). \quad (21)$$

So far we focused on subdiffraction limit and provided formulas for cumulants in the leading orders of Δ to underly the most important aspects of fluctuation enhanced SPADE. Nevertheless, (18) is valid beyond this approximation and allows to express any intensity cumulant as linear combination of spatial moments. This allows to estimate moments of any object with increased precision, as we will demonstrate in what follows.

B. Estimator construction

Let us now analyze the statistics of data obtained in a general imaging scheme of fluctuating sources. The observation time is divided into M frames (as was done to describe SOFI in Sec. II B), the number of photon counts at detector j in frame m is $n_{j,m}$. To simplify the analysis, we assume that subsequent frames are statistically independent, but we will comment how to go beyond this assumption at the end of this section.

Using collected data, we compute sample temporal moments

$$\hat{\mu}(n_{j_1}, ...n_{j_r}) = \frac{1}{M}\sum_{m=1}^{M} n_{j_1,m} ... n_{j_r,m}, \quad (22)$$

which are efficient and unbiased estimators of theoretical temporal moments $\mu(n_{j_1}, ..., n_{j_r}) = \langle n_{j_1} ... n_{j_r} \rangle$. Then, we construct cumulants estimators using (16),

$$\hat{\kappa}(n_{j_1}, ...n_{j_r}) = \sum_{\pi} \Omega_\pi \prod_{B\in\pi} \hat{\mu}(n_{j_i} : i \in B), \quad (23)$$

these estimators are asymptotically unbiased for $M \to \infty$.

In order to make practical use of (18), we should transform estimators of photon count cumulants into estimators of intensity cumulants using

$$\hat{\kappa}^{(r_1,...,r_l)}(I_{j_1}, ..., I_{j_l}) =$$
$$= \sum_{k_1=0}^{r_1} ... \sum_{k_l=0}^{r_l} s(r_1, k_1) ... s(r_l, k_l) \hat{\kappa}^{(k_1,...,k_l)}(n_{j_1}, ..., n_{j_l}), \quad (24)$$

where $s(r, k)$ denote Stirling numbers of the 1st type, j_1,...j_l are non-repeating indices—see appendix B for the definition of Stirling numbers and proof of (24). Formula similar to (24) was derived in Ref. [31] for more general photon statistics—we provide an alternative proof valid for super-poissonian sources only, but extend the applicability to many detectors cases.

Let vector $\hat{\vec{\kappa}}^{(I)}$ collect estimators of all intensity cumulants we want to use. For example, with two detectors and cumulants up to 2nd order, $\hat{\vec{\kappa}}^{(I)} = \left[\hat{\kappa}(I_1), \hat{\kappa}(I_2), \hat{\kappa}^{(2)}(I_1), \hat{\kappa}^{(2)}(I_2), \hat{\kappa}(I_1, I_2)\right]^T$.

Using frames independce and central limit theorem, we prove (in appendix C) that cumulants estimators calculated using (23) and (24) are asymptotically normally distributed,

$$\hat{\vec{\kappa}}^{(I)} \overset{M\to\infty}{\longrightarrow} \mathcal{N}\left(\boldsymbol{A}\vec{\theta}, \boldsymbol{\Sigma}^{(\kappa)}/M\right), \tag{25}$$

where $\boldsymbol{\Sigma}^{(\kappa)}/M$ is a covariance matrix, $\vec{\kappa}^{(I)} = \boldsymbol{A}\vec{\theta}$ is the vector of theoretical cumulants, depending linearly on objects spatial moments collected in $\vec{\theta}$. The matrix $\boldsymbol{A}$, whose elements are determined via (18), specifies this linear relation. Strictly speaking, $\vec{\theta}$ is usually infinite, since elements of $\vec{\kappa}^{(I)}$ generally depend on infinitely many spatial moments. In practice, the series is truncated by neglecting higher moments contributions, making $\vec{\theta}$ finite.

To estimate spatial moments, we use maximum likelihood (ML) estimator, which, for our linear Gaussian model, is given by [69]

$$\hat{\vec{\theta}} = \left(\boldsymbol{A}^T \left(\hat{\boldsymbol{\Sigma}}^{(\kappa)}\right)^{-1} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^T \left(\hat{\boldsymbol{\Sigma}}^{(\kappa)}\right)^{-1} \hat{\vec{\kappa}}^{(I)}, \tag{26}$$

where $\hat{\boldsymbol{\Sigma}}^{(\kappa)}$ is a covariance matrix estimated from collected data—the exact procedure of covariance matrix estimation is described in appendix C.

To summarize, we provided spatial moments estimation procedure valid for any object size (also beyond subdiffraction limit). One should firstly compute relevant temporal moments estimators directly from collected data using (22), and then transform them to intensity cumulants estimators using (23) and (24). Then, spatial moments estimators are linear combinations of intensity cumulants estimators, with coefficients given by (26).

This reasoning can be extended to temporally correlated frames case provided that correlations decay fast enough, so that generalized versions of central limit theorem [70] can be used to justify (25). This can be done, for example, when emitters blinking is governed by a Markov process, which leads to exponentially decaying temporal correlations between frames [68]. In such a case, one needs to take into account these correlations when computing covariance matrix estimator.

C. Fundamental bounds

When consecutive frames are independent and identically distributed, all relevant information is contained in the empirical temporal moments of photon counts, $\hat{\mu}(n_{j_1}, \ldots, n_{j_r})$, provided that a sufficiently high moment order is used. However, the question remains, whether the procedure described in the previous section optimally estimates object's spatial moments θ_μ, especially for finite M, when (25) does not hold exactly, ML estimator (26) is not guaranteed to be optimal, and covariance matrix estimator $\hat{\Sigma}^{(\kappa)}$ might be inaccurate.

To answer this question, we will compare the MSE of the described estimation procedure with the fundamental Cramér-Rao (C-R) bound. Let $\boldsymbol{C}$ be a covariance matrix of spatial moments estimators, $\boldsymbol{C}_{\mu\nu} = \mathrm{Cov}(\hat{\theta}_\mu, \hat{\theta}_\nu)$. The C-R bound states that for any unbiased estimators [69]

$$\boldsymbol{C} \geq \boldsymbol{F}^{-1}, \quad \boldsymbol{F}_{\mu\nu} = \int d\vec{x} \frac{\partial_\mu p(\vec{x}|\vec{\theta}) \partial_\nu p(\vec{x}|\vec{\theta})}{p(\vec{x}|\vec{\theta})} \tag{27}$$

where $\boldsymbol{F}$ is a Fisher Information matrix, $\boldsymbol{A} \geq \boldsymbol{B}$ means that matrix $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite, $\vec{x}$ collects all relevant measurement outcomes, $p(\vec{x}|\vec{\theta})$ is a probability of obtaining outcome $\vec{x}$ for parameters values $\vec{\theta}$, $\partial_\mu = \frac{\partial}{\partial\theta_\mu}$. In our case, $\vec{\theta}$ contains the spatial moments to be estimated, and $\vec{x}$ corresponds to the empirical temporal moments of photon counts, $\vec{x} = \hat{\vec{\mu}}^{(n)}$. When bounding the precision achievable with a chosen set of cumulants, $\vec{x}$ should include only the temporal moments required to reconstruct those cumulants.

Using (27) and asymptotic normality of $\hat{\vec{\mu}}^{(n)}$, we obtain the following bound for MSE of estimating θ_μ from M frames:

$$\delta^2\hat{\theta}_\mu \geq \frac{1}{M}\left[\boldsymbol{F}^{-1}\right]_{\mu\mu}, \quad \boldsymbol{F}_{\mu\nu} = \frac{\partial\vec{\mu}^{(n)T}}{\partial\theta_\mu} {\boldsymbol{\Sigma}^{(n)}}^{-1} \frac{\partial\vec{\mu}^{(n)}}{\partial\theta_\nu}, \tag{28}$$

where $\vec{\mu}$ collects the *theoretical* temporal moments (the expected values of photon counts products), $\boldsymbol{\Sigma}^{(n)}$ is a covariance matrix of these photon counts products, see appendix C for more details. Importantly, $\mathbf{F}$ does not depend on M and can be computed exactly from the statistical model of a single frame, without random sampling. While the bound above is valid for any M, it becomes tight only in the asymptotic limit $M \to \infty$. Nevertheless, we show that even for finite M, the simulated MSEs of our estimators closely follow the bound—this demonstrates both the bound relevance and the estimator optimality.

IV. EXPERIMENT SIMULATION

We now demonstrate the performance of the proposed imaging techniques, SOFSPADE and SOFIII, in estimating spatial moments of blinking objects. We do not discuss the subsequent reconstruction of the object from its moments, as this has already been demonstrated in SPADE imaging simulations [60] and experiments [43–45], both in confocal and wide-field setups. It is clear that improved accuracy in estimating spatial moments—especially higher-order ones—directly translates into higher image resolution and reconstruction fidelity.

Consider an object consisting of $k = 20$ independent blinking point emitters, distributed within a subdiffraction region ($\Delta = 0.3$), as shown in Fig. 3a. Each emitter switches between "on" and "off" states with brightnesses $q_{\mathrm{on}} = 100\frac{\mathrm{photons}}{\mathrm{frame}}$ and $q_{\mathrm{off}} = 5\frac{\mathrm{photons}}{\mathrm{frame}}$, respectively, the probability of "on" state is $p_{\mathrm{on}} = 0.1$—these are reasonable assumptions for SOFI experiments [71–73]. For each frame, the brightness q_i of every emitter is drawn

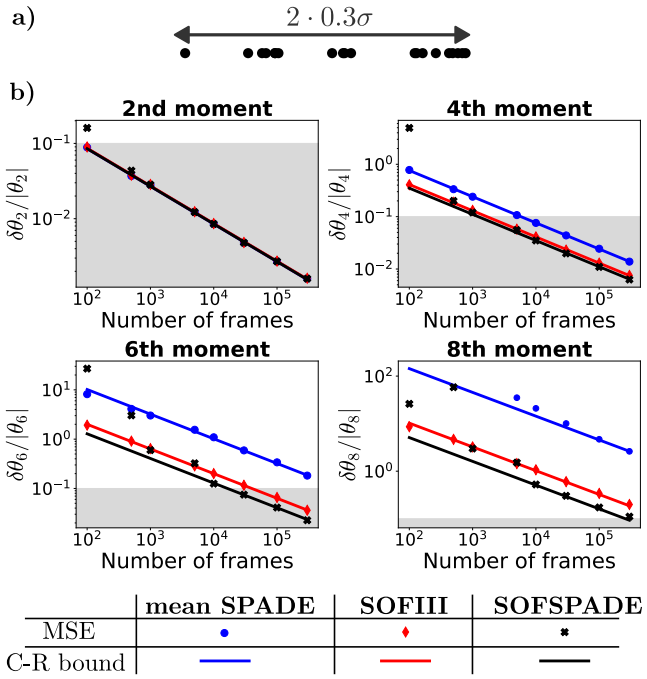


FIG. 3. Simulations are performed for a subdiffraction object ($\Delta = 0.3$) composed of $k = 20$ emitters distributed as in (a). Panel (b) shows the relative errors of even spatial-moment estimates as functions of the number of frames (markers: simulated MSEs; solid lines: C-R bounds) The area corresponding to accuracy 10% or better is shaded in each chart to facilitate comparison. Higher-order moments are harder to estimate, and the benefit of fluctuation-based methods (SOFSPADE, SOFIII) increases with moment order.

independently, after which photon counts at the detectors are sampled from a Poisson distribution with mean given by (1).

This process is repeated independently for M frames. From resulting time series of photon counts, empirical temporal moments are computed using (22), and the corresponding spatial-moment estimators $\hat{\theta}_\mu$ are obtained via (23), (24), (26). The entire estimation procedure is repeated $N = 1000$ times for independently generated datasets of M frames, and for each run we evaluate the squared error $(\hat{\theta}_\mu - \theta_\mu)^2$. Averaging these errors over all repetitions yields the MSE for each estimated moment. We compare these MSEs with C-R bounds obtained using (28).

Let us firstly consider estimation of 5 leading even moments, $\vec{\theta} = [\theta_0, \theta_2, \theta_4, \theta_6, \theta_8]^T$. This can be done using mean intensities on 5 SPADE detectors (mean SPADE), which corresponds to $\hat{\vec{\kappa}} = [\hat{\kappa}(I_0^{\mathrm{S}}), \hat{\kappa}(I_1^{\mathrm{S}}), \hat{\kappa}(I_2^{\mathrm{S}}), \hat{\kappa}(I_3^{\mathrm{S}}), \hat{\kappa}(I_4^{\mathrm{S}})]^T$. Alternatively, one can significantly simplify the measurement and use SOFIII technique, which requires just two detectors and temporal cumulants up to 4th order, $\hat{\vec{\kappa}} = [\hat{\kappa}(I_+^{\mathrm{III}}), \hat{\kappa}(I_-^{\mathrm{III}}), \hat{\kappa}^{(2)}(I_-^{\mathrm{III}}), \hat{\kappa}^{(3)}(I_-^{\mathrm{III}}), \hat{\kappa}^{(4)}(I_-^{\mathrm{III}})]$. Finally, to get maximal precision, one can use SOFSPADE, in which temporal correlations between multiple detectors are used,

$$\hat{\vec{\kappa}} = [\hat{\kappa}(I_0^{\mathrm{S}}), \hat{\kappa}(I_1^{\mathrm{S}}), \hat{\kappa}(I_2^{\mathrm{S}}), \hat{\kappa}(I_3^{\mathrm{S}}), \hat{\kappa}(I_4^{\mathrm{S}}), \hat{\kappa}^{(2)}(I_1^{\mathrm{S}}), \hat{\kappa}(I_1^{\mathrm{S}}, I_2^{\mathrm{S}}), \hat{\kappa}(I_1^{\mathrm{S}}, I_3^{\mathrm{S}}), \hat{\kappa}^{(2)}(I_2^{\mathrm{S}}), \hat{\kappa}^{(3)}(I_1^{\mathrm{S}}), \hat{\kappa}^{(2,1)}(I_1^{\mathrm{S}}, I_2^{\mathrm{S}}), \hat{\kappa}^{(4)}(I_1^{\mathrm{S}})]^T,$$

note that we neglected cumulants with sum of detectors labels greater than 4 since they do not give any information about spatial moments of orders up to 8. The calculated MSEs and corresponding C-R bounds as functions of the number of frames M are shown in Fig. 3b. These results demonstrate that higher-order spatial moments can be reliably estimated using the experimentally accessible SOFIII technique. Remarkably, SOFIII achieves even higher precision than full SPADE with mean signal only. This advantage diminishes for weaker fluctuations, where SOFIII may become less precise than mean SPADE, though still far easier to implement. However, SOFSPADE consistently outperforms both methods, as it combines the information from multiple detectors and temporal correlations. Our simulations show that SOFSPADE advantage grows with spatial moment order since higher spatial moments receive contributions from multiple cumulants. For θ_2, SOFSPADE gives no advantage, as only the first-order cumulant $\langle I_1 \rangle$ carries information about this parameter. In contrast, higher-order cumulants dramatically enhance higher-order moments estimates: SOFSPADE achieves the same precision for the 8th moment with nearly 1000 times fewer frames than mean-signal SPADE.

To estimate odd moments, we need to use iSPADE measurement. If we want to know moments up to order 6, then $\vec{\theta} = [\theta_0, \theta_1, ..., \theta_6]$. Let us consider two estimation strategies: (i) mean iSPADE, for which we use mean intensities on 8 detectors, $\hat{\vec{\kappa}} = [\kappa(I_0^{\mathrm{iS}}), \kappa(I_{0_\pm}^{\mathrm{iS}}), \kappa(I_{1_\pm}^{\mathrm{iS}}), \kappa(I_{2_\pm}^{\mathrm{iS}}), \kappa(I_3^{\mathrm{iS}})]^T$; (ii) binary, stochastic optical fluctuation iSPADE (bin SOFiSPADE), for which 2 H-G modes are interfered, which leads to 4 detectors 0, 0_+, 0_-, 1, for which temporal cumulants from (21) are computed, $\hat{\vec{\kappa}} = [\hat{\kappa}(I_0), \{\hat{\kappa}^{(r)}(I_1), \hat{\kappa}^{(1,r-1)}(I_{0_\pm}, I_1) : r \in \{1, 2, 3\}\}]^T$ The estimation MSEs again follow the C-R bound for sufficiently large M, see Fig. 4. Interestingly, fluctuation-based method with simplified measurement again outperforms mean iSPADE, especially for high even moments and small values of M, for which C-R bound for mean iSPADE is not saturated.

V. PRACTICAL ASPECTS

A. Generalization to 2D objects

To illustrate the working principle of SOFSPADE and SOFIII superresolution techniques without introducing extra mathematical complications, we focused on 1D imaging. Let us now discuss how these results generalize

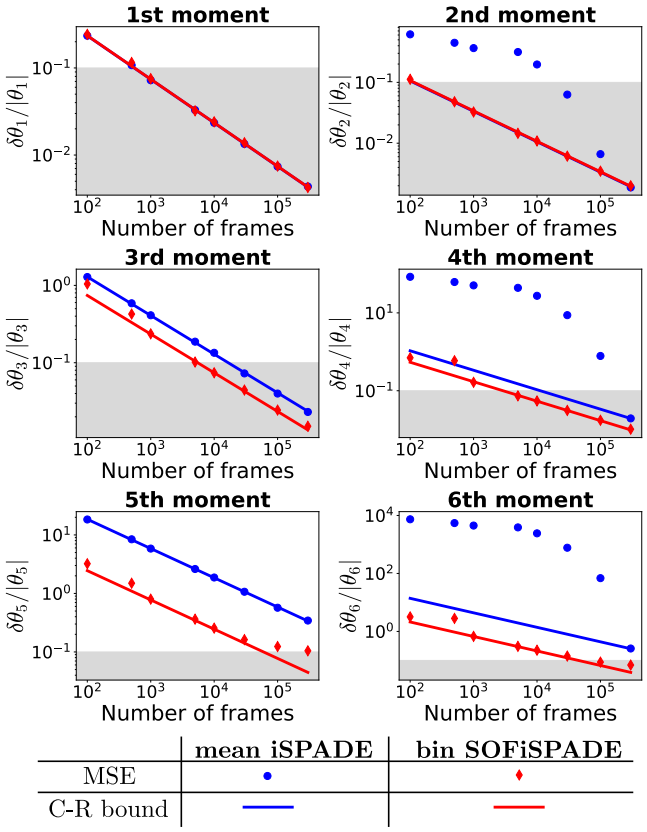


FIG. 4. Odd and even moments of an object shown in Fig. 3a are estimated using mean iSPADE method or bin SOFiSPADE, where measurement is simplified, and temporal fluctuations are used. The simulated MSEs and corresponding C-R bounds of both methods are plotted as functions of the number of frames. For even moments, the estimator based on mean iSPADE only becomes efficient for large number of frames, but bin SOFiSPADE method is free of this issue.

to a more practical, 2D case. The SPADE measurement is then performed in the basis of 2D H-G modes defined as $\phi_{jk}(x',y') = \phi_j(x')\phi_k(y')$, $j,k \in \{0,1,2,...\}$. The corresponding transfer function is a product of 1D transfer functions,

$$T^{2S}(j,k|x,y) = T^S(j|x)T^S(k|y), \tag{29}$$

where x, y are coordinates of a light emitter in the object plane. A 2D object consisting of point sources with positions (x_i, y_i) and fluctuating brightnesses q_i can be characterized by its 2D moments defined as $\theta_{\mu_x,\mu_y} = \sum_i \langle q_i \rangle x_i^{\mu_x} y_i^{\mu_y}$. As in 1D case, objects moments are directly related to mean intensities $I_{j,k}$ of different H-G modes—for subdiffraction objects

$$\langle I^{2S}_{j,k} \rangle = \frac{1}{4^{j+k} j!k!}\theta_{2j,2k} + \mathcal{O}(\Delta^{2j+2k+2}), \tag{30}$$

where $\Delta = \max_i \sqrt{x_i^2 + y_i^2}/\sigma$ characterizes the size of 2D object. Again, one can go beyond subdiffraction limit and express $\langle I_{j,k} \rangle$ as a combination of infinite number of moments with leading contribution from $\theta_{2j,2k}$, as in (5).

The 2D analogue of (19) is

$$\kappa(I^{2S}_{j_1,k_1},...,I^{2S}_{j_r,k_r}) = \\ = \frac{\tilde{\kappa}_r}{\prod_{l=1}^r 4^{j_l+k_l} j_l!k_l!}\theta_{2(j_1+...+j_r),2(k_1+...+k_r)}, \tag{31}$$

so as in 1D case, cumulants of intensities in lower modes provide information about higher-order moments; for example, $\kappa(I^{2S}_{0,1}, I^{2S}_{1,0}) = \frac{\tilde{\kappa}_2}{16}\theta_{2,2} + \mathcal{O}(\Delta^6)$ $\kappa^{(2)}(I^{2S}_{1,1}) = \frac{\tilde{\kappa}_2}{256}\theta_{4,4} + \mathcal{O}(\Delta^{10})$, etc. Therefore, fluctuations help to estimate even 2D moments $\theta_{j,k}$ for which $j+k \geq 4$; the general estimator construction is the same as for 1D case. To estimate odd moments, one should use 2D version of iSPADE measurement [53]; there is also 2D version of III technique [74], which, when enhanced by fluctuations, may be used to estimate all 2D even moments.

B. Blinking parameters estimation

To apply SOFSPADE or SOFIII techniques in practice, one needs to know $\tilde{\kappa}_r$ coefficients characterizing blinking dynamics of emitters. These coefficients can be easily computed when blinking parameters q_{on}, q_{off} and p_{on} are known. If this is not true, one can estimate $\tilde{\kappa}_r$ from experimental data. Let $n_{\text{all},m} = \sum_j n_{j,m}$ be the total number of photon counts collected from all detectors in a given frame m. Then, one can calculate sample moments $\hat{\mu}^{(r)}(n_{\text{all}}) = \frac{1}{M}\sum_{m=1}^M n^r_{\text{all},m}$, and, further on, intensity cumulants of total signal $\hat{\kappa}^{(r)}(I_{\text{all}})$ using (16) and (24). Then, coefficients $\tilde{\kappa}_r$ can be estimated as $\tilde{\kappa}_r = \hat{\kappa}^{(r)}(I_{\text{all}})/\langle I_{\text{all}} \rangle$—because of cumulants additivity, the ratio between rth cumulant and mean of total signal is the same as the corresponding ratio for a single emitter.

VI. CONCLUSIONS

To summarize, we have shown that blinking can substantially enhance the performance of SPADE: the proposed SOFSPADE method exploits temporal fluctuations to achieve markedly improved precision, especially for higher moments estimation.

More importantly, emitter fluctuations enable the far simpler III measurement to recover the full object information. The proposed SOFIII technique is experimentally much easier to implement than SPADE, yet it still allows recovery of all even spatial moments of the object. The only practical limitation comes from the fact that higher-order cumulants become hard to estimate.

It is especially promising to implement SOFIII in a confocal microscope—on one hand, confocal versions of SPADE have already demonstrated the ability to reconstruct complex objects with enhanced resolution; on the

other hand, SOFIII delivers the same information content as SPADE, while the III measurement has already been realized for fluorescent sources and high-NA objectives. Incorporating blinking fluorophores and time-resolved detectors should not significantly increase the experimental complexity of an III setup.

Acknowledgements I would like to thank Rafał Demkowicz-Dobrzański for suggestions, discussions, and reading the initial version of my manuscript. I would also like to thank Konrad Banaszek for his support, and Wojciech Górecki for discussions. This work is a part of the project "Quantum Optical Technologies" carried out within the International Research Agendas programme of the Foundation for Polish Science cofinanced by the European Union under the European Regional Development Fund and was supported by the National Science Center (Poland) grant No.2020/37/B/ST2/02134. The author is a recipient of the Foundation for Polish Science START 2025 scholarship.

[1] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).
[2] E. Betzig, Proposed method for molecular optical imaging, Opt. Lett. **20**, 237 (1995).
[3] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution, Science **313**, 1642 (2006), https://science.sciencemag.org/content/313/5793/1642.full.pdf.
[4] M. J. Rust, M. Bates, and X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm), Nature Methods **3**, 793 (2006).
[5] S. W. Hell, Far-field optical nanoscopy, Science **316**, 1153 (2007), https://science.sciencemag.org/content/316/5828/1153.full.pdf.
[6] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, Fast, background-free, 3d super-resolution optical fluctuation imaging (sofi), Proceedings of the National Academy of Sciences **106**, 22287 (2009), https://www.pnas.org/content/106/52/22287.full.pdf.
[7] B. O. Leung and K. C. Chou, Review of super-resolution fluorescence microscopy for biology, Applied Spectroscopy **65**, 967 (2011), pMID: 21929850, https://doi.org/10.1366/11-06398.
[8] D. Gatto Monticone, K. Katamadze, P. Traina, E. Moreva, J. Forneris, I. Ruo-Berchera, P. Olivero, I. P. Degiovanni, G. Brida, and M. Genovese, Beating the abbe diffraction limit in confocal microscopy via nonclassical photon statistics, Phys. Rev. Lett. **113**, 143602 (2014).
[9] W. E. Moerner, Nobel lecture: Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy, Rev. Mod. Phys. **87**, 1183 (2015).
[10] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, Phys. Rev. X **6**, 031033 (2016).
[11] M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto, and J. Rehacek, Achieving the ultimate optical resolution, Optica **3**, 1144 (2016).
[12] F. Yang, A. Tashchilina, E. S. Moiseev, C. Simon, and A. I. Lvovsky, Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode, Optica **3**, 1148 (2016).
[13] W.-K. Tham, H. Ferretti, and A. M. Steinberg, Beating rayleigh's curse by imaging using phase information, Phys. Rev. Lett. **118**, 070801 (2017).
[14] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, Beating the rayleigh limit using two-photon interference, Phys. Rev. Lett. **121**, 250503 (2018).
[15] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum illumination with gaussian states, Phys. Rev. Lett. **101**, 253601 (2008).
[16] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, Quantum interferometric optical lithography: Exploiting entanglement to beat the diffraction limit, Phys. Rev. Lett. **85**, 2733 (2000).
[17] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, Subdiffraction-limited quantum imaging within a living cell, Phys. Rev. X **4**, 011017 (2014).
[18] L. A. Rozema, J. D. Bateman, D. H. Mahler, R. Okamoto, A. Feizpour, A. Hayat, and A. M. Steinberg, Scalable Spatial Superresolution Using Entangled Photons, Phys. Rev. Lett. **112**, 223602 (2014).
[19] J. P. Dowling and K. P. Seshadreesan, Quantum optical technologies for metrology, sensing, and imaging, Journal of Lightwave Technology **33**, 2359 (2015).
[20] M. Genovese, Real applications of quantum imaging, Journal of Optics **18**, 073002 (2016).
[21] C. Schnell, Quantum imaging in biological samples, Nature Methods **16**, 214 (2019).
[22] M. G. L. Gustafsson, Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy: Short communication, Journal of Microscopy **198**, 82–87 (2000).
[23] C. B. Müller and J. Enderlein, Image scanning microscopy, Phys. Rev. Lett. **104**, 198101 (2010).
[24] O. Schwartz and D. Oron, Improved resolution in fluorescence microscopy using quantum correlations, Phys. Rev. A **85**, 033812 (2012).
[25] O. Schwartz, J. M. Levitt, R. Tenne, S. Itzhakov, Z. Deutsch, and D. Oron, Superresolution microscopy with quantum emitters, Nano letters **13**, 5832 (2013).
[26] K. Fujita, M. Kobayashi, S. Kawano, M. Yamanaka, and S. Kawata, High-resolution confocal microscopy by saturated excitation of fluorescence, Phys. Rev. Lett. **99**, 228105 (2007).
[27] M. Yamanaka, S. Kawano, K. Fujita, N. I. Smith, and S. Kawata, Beyond the diffraction-limit biological imaging by saturated excitation microscopy, Journal of Biomedical Optics **13**, 050507 (2008).
[28] F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynnå, V. Westphal, F. D. Stefani, J. Elf,

and S. W. Hell, Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes, Science **355**, 606 (2017), https://www.science.org/doi/pdf/10.1126/science.aak9913.

[29] T. Moosmayer, K. A. Kiszka, V. Westphal, J. K. Pape, M. Leutenegger, H. Steffens, S. G. N. Grant, S. J. Sahl, and S. W. Hell, Minflux fluorescence nanoscopy in biological tissue, Proceedings of the National Academy of Sciences **121**, e2422020121 (2024), https://www.pnas.org/doi/pdf/10.1073/pnas.2422020121.

[30] S. Wang, T. W. Shin, H. B. Yoder, R. B. McMillan, H. Su, Y. Liu, C. Zhang, K. S. Leung, P. Yin, L. L. Kiessling, and E. S. Boyden, Single-shot 20-fold expansion microscopy, Nature Methods **21**, 2128–2134 (2024).

[31] F. Picariello, E. Losero, S. D. Tchernij, P. Boucher, M. Genovese, I. Ruo-Berchera, and I. P. Degiovanni, Quantum super-resolution microscopy by photon statistics and structured light, Optica **12**, 490 (2025).

[32] T. Dertinger, R. Colyer, R. Vogel, J. Enderlein, and S. Weiss, Achieving increased resolution and more pixels with superresolution optical fluctuation imaging (sofi), Opt. Express **18**, 18875 (2010).

[33] N. Sandeau and H. Giovannini, Increasing the lateral resolution of 4pi fluorescence microscopes, J. Opt. Soc. Am. A **23**, 1089 (2006).

[34] R. Nair and M. Tsang, Interferometric superlocalization of two incoherent optical point sources, Opt. Express **24**, 3684 (2016).

[35] F. Yang, R. Nair, M. Tsang, C. Simon, and A. I. Lvovsky, Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection in a higher-order local oscillator mode, Physical Review A **96**, 10.1103/physreva.96.063829 (2017).

[36] C. Datta, M. Jarzyna, Y. L. Len, K. Łukanowski, J. Kołodyński, and K. Banaszek, Sub-rayleigh resolution of two incoherent sources by array homodyning, Phys. Rev. A **102**, 063526 (2020).

[37] A. Classen, J. von Zanthier, M. O. Scully, and G. S. Agarwal, Superresolution via structured illumination quantum correlation microscopy, Optica **4**, 580 (2017).

[38] A. Sroda, A. Makowski, R. Tenne, U. Rossman, G. Lubin, D. Oron, and R. Lapkiewicz, Sofism: Super-resolution optical fluctuation image scanning microscopy, Optica **7**, 1308 (2020).

[39] J. Rehacek, M. Paúr, B. Stoklasa, Z. Hradil, and L. L. Sánchez-Soto, Optimal measurements for resolution beyond the rayleigh limit, Opt. Lett. **42**, 231 (2017).

[40] X.-J. Tan, L. Qi, L. Chen, A. J. Danner, P. Kanchanawong, and M. Tsang, Quantum-inspired superresolution for incoherent imaging, Optica **10**, 1189 (2023).

[41] C. Rouvière, D. Barral, A. Grateau, I. Karuseichyk, G. Sorelli, M. Walschaers, and N. Treps, Ultra-sensitive separation estimation of optical sources, Optica **11**, 166 (2024).

[42] L. Santamaria, F. Sgobba, and C. Lupo, Single-photon sub-rayleigh precision measurements of a pair of incoherent sources of unequal intensity, Optica Quantum **2**, 46 (2024).

[43] A. A. Pushkina, G. Maltese, J. I. Costa-Filho, P. Patel, and A. I. Lvovsky, Superresolution linear optical imaging in the far field, Phys. Rev. Lett. **127**, 253602 (2021).

[44] J. Frank, A. Duplinskiy, K. Bearne, and A. I. Lvovsky, Passive superresolution imaging of incoherent objects, Optica **10**, 1147 (2023).

[45] A. Duplinskiy, J. Frank, K. Bearne, and A. I. Lvovsky, Tsang's resolution enhancement method for imaging with focused illumination, Light: Science & Applications **14**, 10.1038/s41377-025-01791-4 (2025).

[46] Z. Huang and C. Lupo, Quantum hypothesis testing for exoplanet detection, Phys. Rev. Lett. **127**, 130502 (2021).

[47] S. Zhou and L. Jiang, Modern description of rayleigh's criterion, Phys. Rev. A **99**, 013808 (2019).

[48] C. Oh, S. Zhou, Y. Wong, and L. Jiang, Quantum limits of superresolution in a noisy environment, Phys. Rev. Lett. **126**, 120502 (2021).

[49] Y. L. Len, C. Datta, M. Parniak, and K. Banaszek, Resolution limits of spatial mode demultiplexing with noisy detection, International Journal of Quantum Information **18**, 1941015 (2020), https://doi.org/10.1142/S0219749919410156.

[50] C. Lupo, Subwavelength quantum imaging with noisy detectors, Phys. Rev. A **101**, 022323 (2020).

[51] M. Gessner, C. Fabre, and N. Treps, Superresolution limits from measurement crosstalk, Phys. Rev. Lett. **125**, 100501 (2020).

[52] J. O. de Almeida, J. Kołodyński, C. Hirche, M. Lewenstein, and M. Skotiniotis, Discrimination and estimation of incoherent sources under misalignment, Phys. Rev. A **103**, 022406 (2021).

[53] M. Tsang, Subdiffraction incoherent optical imaging via spatial-mode demultiplexing, New Journal of Physics **19**, 023054 (2017).

[54] M. Tsang, Subdiffraction incoherent optical imaging via spatial-mode demultiplexing: Semiclassical treatment, Phys. Rev. A **97**, 023830 (2018).

[55] M. Tsang, Quantum limit to subdiffraction incoherent optical imaging, Phys. Rev. A **99**, 012305 (2019).

[56] M. Tsang, Semiparametric estimation for incoherent optical imaging, Phys. Rev. Res. **1**, 033006 (2019).

[57] M. Tsang, Quantum limit to subdiffraction incoherent optical imaging. ii. a parametric-submodel approach, Phys. Rev. A **104**, 052411 (2021).

[58] X.-J. Tan and M. Tsang, Quantum limit to subdiffraction incoherent optical imaging. iii. numerical analysis, Phys. Rev. A **108**, 052416 (2023).

[59] Y. Wang and V. O. Lorenz, Fundamental limit of bandwidth-extrapolation-based superresolution, Phys. Rev. A **108**, 012602 (2023).

[60] K. K. M. Bearne, Y. Zhou, B. Braverman, J. Yang, S. A. Wadood, A. N. Jordan, A. N. Vamivakas, Z. Shi, and R. W. Boyd, Confocal super-resolution microscopy based on a spatial mode sorter, Opt. Express **29**, 11784 (2021).

[61] S. Zhang, A. Zhang, I. M. d. B. Wenniger, P. M. Burdekin, S. Sagona-Stophel, A. Rastogi, S. E. Thomas, and I. A. Walmsley, Super-resolving frequency measurement with mode-selective quantum memory (2025).

[62] A. B. Greenwood, R. Oulton, and H. Gersen, On the impact of realistic point sources in spatial mode demultiplexing super resolution imaging, Quantum Science and Technology **8**, 015024 (2023).

[63] W. Larson and B. E. A. Saleh, Two-lens anisotropic image-inversion system for interferometric information processing, Opt. Express **29**, 15403 (2021).

[64] R. Aiello, F. Sgobba, P. Maddaloni, F. D. Lena, and L. S. Amato, Sub-rayleigh imaging enhanced by phase-locked inversion interferometry, Opt. Express **33**, 44834 (2025).

[65] C. S. Mitchell, D. Dhruva, Z. P. Burke, D. J. Durden,

A. I. Dingilian, and M. P. Backlund, Quantum-inspired super-resolution of fluorescent point-like sources (2024).
[66] A. L. Efros and D. J. Nesbitt, Origin and control of blinking in quantum dots, Nature Nanotechnology **11**, 661 (2016).
[67] R. M. Dickson, A. B. Cubitt, R. Y. Tsien, and W. E. Moerner, On/off blinking and switching behaviour of single molecules of green fluorescent protein, Nature **388**, 355 (1997).
[68] S. Kurdziałek and R. Demkowicz-Dobrzański, Super-resolution optical fluctuation imaging—fundamental estimation theory perspective, Journal of Optics **23**, 075701 (2021).
[69] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory* (Prentice-Hall, Inc., 1993).
[70] M. Rosenblatt, A central limit theorem and a strong mixing condition, Proceedings of the National Academy of Sciences **42**, 43 (1956), https://www.pnas.org/content/42/1/43.full.pdf.
[71] P. Dedecker, G. C. H. Mo, T. Dertinger, and J. Zhang, Widely accessible method for superresolution fluorescence imaging of living systems, Proceedings of the National Academy of Sciences **109**, 10909–10914 (2012).
[72] H. Deschout, T. Lukes, A. Sharipov, D. Szlag, L. Feletti, W. Vandenberg, P. Dedecker, J. Hofkens, M. Leutenegger, T. Lasser, and A. Radenovic, Complementarity of palm and sofi for super-resolution live-cell imaging of focal adhesions, Nature Communications **7**, 10.1038/ncomms13693 (2016).
[73] W. Vandenberg and P. Dedecker, Effect of probe diffusion on the sofi imaging accuracy, Scientific Reports **7**, 10.1038/srep44665 (2017).
[74] S. Z. Ang, R. Nair, and M. Tsang, Quantum limit for two-dimensional resolution of two incoherent optical point sources, Phys. Rev. A **95**, 063847 (2017).
[75] D. R. Brillinger, The calculation of cumulants via conditioning, Annals of the Institute of Statistical Mathematics **21**, 215–218 (1969).
[76] DLMF, *NIST Digital Library of Mathematical Functions*, `https://dlmf.nist.gov/`, Release 1.2.4 of 2025-03-15, f. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.
[77] K. M. Wolter, Taylor series methods, in *Introduction to Variance Estimation* (Springer New York, New York, NY, 2007) pp. 226–271.

Appendix A: Cumulants properties

We can get explicit formulas for single-variable and multi-variable cumulants by expanding cumulants generating functions, (13) and (15) respectively. This allows us to express cumulants in terms of moments, see (16). The expressions for single-variable cumulants up to order 4 are:

$$\kappa^{(1)}(X) = \langle X\rangle, \tag{A1a}$$
$$\kappa^{(2)}(X) = \langle X^2\rangle - \langle X\rangle^2, \tag{A1b}$$
$$\kappa^{(3)}(X) = \langle X^3\rangle - 3\langle X^2\rangle\langle X\rangle + 2\langle X\rangle^3, \tag{A1c}$$
$$\kappa^{(4)}(X) = \langle X^4\rangle - 4\langle X^3\rangle\langle X\rangle - 3\langle X^2\rangle^2 + 12\langle X^2\rangle\langle X\rangle^2 - 6\langle X\rangle^4. \tag{A1d}$$

Note that 2nd and 3rd cumulants are equivalent to central moments. The analogous expression for multivariable cumulants are

$$\kappa(X,Y) = \langle XY\rangle - \langle X\rangle\langle Y\rangle, \tag{A2a}$$
$$\kappa(X,Y,Z) = \langle XYZ\rangle - \langle XY\rangle\langle Z\rangle - \langle XZ\rangle\langle Y\rangle - \langle YZ\rangle\langle X\rangle + 2\langle X\rangle\langle Y\rangle\langle Z\rangle, \tag{A2b}$$
$$\begin{aligned}\kappa(X,Y,Z,W) = {}& \langle XYZW\rangle - \langle XYZ\rangle\langle W\rangle - \langle XYW\rangle\langle Z\rangle - \langle XZW\rangle\langle Y\rangle - \langle YZW\rangle\langle X\rangle \\ & - \langle XY\rangle\langle ZW\rangle - \langle XZ\rangle\langle YW\rangle - \langle XW\rangle\langle YZ\rangle \\ & + 2\big(\langle XY\rangle\langle Z\rangle\langle W\rangle + \langle XZ\rangle\langle Y\rangle\langle W\rangle + \langle XW\rangle\langle Y\rangle\langle Z\rangle \\ & \quad + \langle YZ\rangle\langle X\rangle\langle W\rangle + \langle YW\rangle\langle X\rangle\langle Z\rangle + \langle ZW\rangle\langle X\rangle\langle Y\rangle\big) \\ & - 6\langle X\rangle\langle Y\rangle\langle Z\rangle\langle W\rangle.\end{aligned} \tag{A2c}$$

One can also express moments in terms of cumulants; inverted relation (16) reads

$$\mu(X_1,...,X_r) = \langle X_1...X_r\rangle = \sum_{\pi}\prod_{B\in\pi}\kappa(X_i : i\in B), \tag{A3}$$

π is set of all partitions of $\{1,...,r\}$ into subsets. Importantly, multivariate cumulants are linear in each argument

$$\kappa(\alpha X + \beta Y, X_2, ..., X_n) = \alpha\kappa(X, X_2, ..., X_n) + \beta(Y, X_2, ..., X_n), \tag{A4}$$

α, β are numbers. For single-variable cumulant of order r this formula leads to

$$\kappa^{(r)}(\alpha X + \beta Y) = \sum_{j=0}^{r}\binom{r}{j}\alpha^j\beta^{r-j}\kappa^{(j,r-j)}(X,Y) \tag{A5}$$

The crucial property of cumulants , which distinguishes them from other statistics like moments and central moments, is

$$\kappa^{(r_1,...,r_n)}(X_1,...,X_n)=0 \text{ for independent } X_1,...,X_n \text{ if } \#\{i:r_i>0\}\geq 2, \tag{A6}$$

where $\#\{i:\bullet\}$ is the number of indices i satisfying condition $\bullet$. This property makes cumulants so useful in SOFI, SOFSPADE and SOFIII techniques as long as different light emitters are statistically independent. Using (A5) and (A6), we get

$$\kappa^{(r)}(\alpha X+\beta Y)=\alpha^r\kappa^{(r)}(X)+\beta^r\kappa^{(r)}(Y) \text{ for independent } X,Y, \tag{A7}$$

and, more generally

$$\kappa^{(r)}\left(\sum_i \alpha_i X_i\right)=\sum_i \alpha_i^r\ \kappa^{(r)}(X_i) \text{ for independent } X_1,...,X_n. \tag{A8}$$

This equation is a basic principle of SOFI technique, as it directly translates to (14) when X_i are replaced with q_i and α_i are replace with $U(x_j-x_i)$.

The other property, which will be used in appendix B is the law of total cumulance [75]: when $X_1,X_2,...,X_n,Y$ are random variables, then

$$\kappa(X_1,X_2,...,X_n)=\sum_\pi \kappa_Y(\kappa(X_i:i\in B|Y):B\in\pi), \tag{A9}$$

where summation is over all partitions π of set of indices$\{1,2,...,n\}$ into subsets, B denotes a subset of a given partition; for example, when $n=6$, then one exemplary partition of the set of indices is $\pi=\{\{1,2\},\{3,4\},\{5\},\{6\}\}$, elements of this partition, are $B=\{1,2\}$, $B=\{3,4\}$, $B=\{5\}$, $B=\{6\}$. Conditional cumulants $\kappa(X_i:i\in B|Y)$ depend on the value of Y, so they are random variables because Y is a random variable; we use notation κ_Y to indicate that the external cumulant is computed taking into account the randomness of Y only. To get some more intuition, let us see how (A9) looks like for $n=2$ and $n=3$

$$\kappa(X_1,X_2)=\kappa_Y(\kappa(X_1,X_2|Y))+\kappa_Y(\kappa(X_1|Y),\kappa(X_2|Y)) \tag{A10}$$

$$\begin{aligned}\kappa(X_1,X_2,X_3)=\kappa_Y(\kappa(X_1,X_2,X_3|Y))+\kappa_Y(\kappa(X_1|Y),\kappa(X_2,X_3|Y))+\kappa_Y(\kappa(X_2|Y),\kappa(X_1,X_3|Y))+\\ +\kappa_Y(\kappa(X_3|Y),\kappa(X_1,X_2|Y))+\kappa_Y(\kappa(X_1|Y),\kappa(X_2|Y),\kappa(X_3|Y))\end{aligned} \tag{A11}$$

Finally, another useful property is that all cumulants of a Poissonian random variable are equal to mean value of this variable,

$$X\sim\text{Poiss}(\mu)\Rightarrow\forall_r\ \ \kappa^{(r)}(X)=\mu, \tag{A12}$$

Appendix B: Relation between n-cumulants and I-cumulants

In this work we distinguish between photon counts numbers n and intensities I which are linked by relation $n\sim\text{Poiss}(I)$, see (1). It is much easier to use intensities in derivations because relation $I_j=\sum_i T(j|\boldsymbol{r}_i)q_i$ holds for I (expected values of Poissonian random variables), but not for n—this is related to the fact that coherent (Poissonian) state at beam-splitter transforms into independent coherent states, but Fock state transforms into non-trivially correlated states. In fact, for a fixed numbers of photons emitted by different sources n_i, photon counts at detectors n_j are correlated and distributed according to multinomial distribution. We can then include fluctuations by introducing a probability of different numbers of photons emitted $P(n_i)$, and write down the statistical model for n_j as a combination of multinomial distributions—this was done for direct imaging measurement in Ref. [31].

Here we take different approach, valid for super-poissonian sources only—we assume that the number of photons emitted by each source in a single frame is given by $\text{Poiss}(q_i)$, where q_i is a random variable describing brightness of a given source. Therefore, the number of emitted photons fluctuates both due to variability of q_i and due to shot noise, which leads to superpoissonian fluctuations. Then, we use (1) to transform brightnesses q_i into detectors intensities I_j, and, further on, into detectors photon counts n_j.

For this, we need (24), which links measurable cumulants of photon counts with easier-to-calculate cumulants of intensities. Let us now prove (24).

Firstly, let us write an equation for $\kappa^{(r_1,...,r_l)}(n_{j_1},...,n_{j_l})$ (when j_1, ..., j_l are non-repeating indices) using the law of total cumulance (A9) with $Y=(I_{j_1},...,I_{j_l})$, $X_i=n_{j_i}$; then conditional random variables $X_1,...,X_l|Y$ are independent Poissonian variables with mean values $\langle X_k|Y\rangle = I_{j_k}$. Using the fact that cumulants of independent random variables vanish (A6) and using (A12), we get

$$\kappa^{(r_1,...,r_l)}(X_1,...,X_l|Y)=\begin{cases} I_{j_k} & \text{if } r_k>0, r_1=...=r_{k-1}=r_{k+1}=...r_l=0 \\ 0 & \text{in all other cases}\end{cases} \tag{B1}$$

Therefore, many terms in the total cumulance expansion (A9) will vanish, and the only non-zero terms are those, in which partition π contains only subsets B consisting of one repeating variable. Let π be a partition in which r_1 variables X_1 were divided into i_1 subsets, r_2 variables X_2 were divided into i_2 subsets, etc. Then,

$$\kappa_Y(\kappa(X_i : i\in B|Y) : B\in\pi) = \kappa^{(i_1,...,i_l)}(I_{j_1},...,I_{j_l}) \tag{B2}$$

Moreover, there are $S(r_1,i_1)...S(r_l,i_l)$ such partitions π, where $S(r,i)$ denotest the number of partitions of r-element set into i subsets. Therefore,

$$\kappa^{(r_1,...,r_l)}(n_{j_1},...,n_{j_l}) = \sum_{i_1=0}^{r_1}\sum_{i_2=0}^{r_2}...\sum_{i_l=0}^{r_l} S(r_1,i_1)...S(r_l,i_l)\kappa^{(i_1,...,i_l)}(I_{j_1},...,I_{j_l}). \tag{B3}$$

Numbers $S(r,i)$ are called Stirling numbers of the 2nd type, and they have the following property [76]

$$\sum_{k=0}^{n} s(n,k)\,S(k,m) = \delta_{n,m} = \sum_{k=0}^{n} S(n,k)\,s(k,m) \tag{B4}$$

where $s(r,i)$ are Stirling numbers of the 1st type, defined as the coefficients in the following expansion

$$x(x+1)...(x+n-1) = \sum_{k=0}^{n} s(n,k)x^k \tag{B5}$$

Equation (B4) allows us to invert relation (B3), which leads to (24) from the main text.

Appendix C: Details of estimator construction

Let us remind that vectors $\vec{\mu}^{(n)}$ and $\vec{\kappa}^{(I)}$ collect all theoretical photon counts moments and intensity cumulants we want to use; analogously, vectors $\hat{\vec{\mu}}^{(n)}$ and $\hat{\vec{\kappa}}^{(I)}$ collect estimators of photon counts moments and estimators of intensity cumulants; elements of $\hat{\vec{\mu}}^{(n)}$ are given by (22), elements of $\hat{\vec{\kappa}}^{(I)}$ can be expressed in terms of $\hat{\vec{\mu}}^{(n)}$ using (23) and (24), and the same equations can be used to express $\vec{\kappa}^{(I)}$ in terms of $\vec{\mu}^{(n)}$. We can define a Jacobian matrix of transformation from $\vec{\mu}^{(n)}$ to $\vec{\kappa}^{(I)}$ as

$$\frac{\partial[\vec{\kappa}^{(I)}]_i}{\partial[\vec{\mu}^{(n)}]_j} = \boldsymbol{J}_{ij}(\vec{\mu}^{(n)}), \tag{C1}$$

elements of this matrix can be found using (23) and (24). The matrix $\boldsymbol{J}_{ij}(\hat{\vec{\mu}}^{(n)})$ is a Jacobian of transformation from $\hat{\vec{\mu}}^{(n)}$ to $\hat{\vec{\kappa}}^{(I)}$. Cumulants estimators calculated using (23) and (24) are asymptotically unbiased, which means that

$$\langle\hat{\vec{\kappa}}^{(I)}\rangle \overset{M\to\infty}{\longrightarrow} \vec{\kappa}^{(I)} = \boldsymbol{A}\vec{\theta}. \tag{C2}$$

Let us observe that (22) can be written as $\hat{\vec{\mu}}^{(n)} = \sum_{m=1}^{M}\vec{n}_m$, where each $\vec{n}_m$ is an independent random vector containing the products of photon counts $n_{j_1,m}\cdots n_{j_r,m}$ used to estimate the moments. By the central limit theorem,

$$\hat{\vec{\mu}}^{(n)} \overset{M\to\infty}{\longrightarrow} \mathcal{N}(\vec{\mu}^{(n)},\ \boldsymbol{\Sigma}^{(n)}/M), \tag{C3}$$

where $\mathcal{N}(\vec{\mu},\boldsymbol{\Sigma})$ denotes multivariate normal distribution with mean vector $\vec{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$. In our case, a covariance matrix can be estimated from collected data using

$$\hat{\boldsymbol{\Sigma}}^{(n)} = \frac{1}{M}\left(\sum_{m=1}^{M}\vec{n}_m\vec{n}_m^T\right) - \hat{\vec{\mu}}^{(n)}\hat{\vec{\mu}}^{(n)T} \tag{C4}$$

Cumulants estimators are functions of moments estimators, so from delta method [77], $\hat{\vec{\kappa}}$ is also normally distributed, and

$$\hat{\vec{\kappa}}^{(I)} \overset{M\to\infty}{\longrightarrow} \mathcal{N}\left(\boldsymbol{A}\vec{\theta}, \boldsymbol{\Sigma}^{(\kappa)}/M\right), \quad \boldsymbol{\Sigma}^{(\kappa)} = \boldsymbol{J}(\vec{\mu}^{(n)})\boldsymbol{\Sigma}^{(n)}\boldsymbol{J}^T(\vec{\mu}^{(n)}). \tag{C5}$$

We proved that cumulants estimators are described by linear Gaussian model (25). This allows to construct a simple maximum likelihood estimator [69] of object's spatial moments

$$\hat{\vec{\theta}} = \left(\boldsymbol{A}^T\left(\boldsymbol{\Sigma}^{(\kappa)}\right)^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\left(\boldsymbol{\Sigma}^{(\kappa)}\right)^{-1}\hat{\vec{\kappa}}^{(I)} \tag{C6}$$

We cannot use this formula directly because we do not know the covariance matrix $\boldsymbol{\Sigma}^{(\kappa)}$; however, we can replace $\boldsymbol{\Sigma}^{(\kappa)}$ with its estimator constructed as

$$\hat{\boldsymbol{\Sigma}}_0^{(\kappa)} = \boldsymbol{J}(\hat{\vec{\mu}}^{(n)})\hat{\boldsymbol{\Sigma}}^{(n)}\boldsymbol{J}^T(\hat{\vec{\mu}}^{(n)}), \tag{C7}$$

where $\hat{\boldsymbol{\Sigma}}^{(n)}$ is constructed using (C4), $\hat{\vec{\mu}}^{(n)}$ entering $\boldsymbol{J}(\hat{\vec{\mu}}^{(n)})$ are constructed using (22); this leads to initial moments estimate

$$\hat{\vec{\theta}}_0 = \left(\boldsymbol{A}^T\left(\hat{\boldsymbol{\Sigma}}_0^{(\kappa)}\right)^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\left(\hat{\boldsymbol{\Sigma}}_0^{(\kappa)}\right)^{-1}\hat{\vec{\kappa}}^{(I)} \tag{C8}$$

We can further increase the estimation accuracy and robustness by using estimated moments to create new, more accurate estimators $\hat{\boldsymbol{\Sigma}}^{(\kappa)}$, which are then used to reestimate θ; subsequent estimators can be constructed iteratively:

$$\hat{\vec{\theta}}_{l+1} = \left(\boldsymbol{A}^T\left(\hat{\boldsymbol{\Sigma}}_{l+1}^{(\kappa)}(\hat{\vec{\theta}}_l)\right)^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\left(\hat{\boldsymbol{\Sigma}}_{l+1}^{(\kappa)}(\hat{\vec{\theta}}_l)\right)^{-1}\hat{\vec{\kappa}}^{(I)}, \tag{C9}$$

where $\hat{\boldsymbol{\Sigma}}_{l+1}^{(\kappa)}(\hat{\vec{\theta}}_l)$ means that we calculate covariance matrix estimator assuming $\hat{\vec{\theta}}_l$ are real moments—firstly, we calculate cumulants as a function of $\hat{\vec{\theta}}_l$ using (18), then we transform intensity cumulants to photon counts moments using (B3) and (A3), and finally, covariance matrix estimator is calculated using (C4) and (C7).

Chapter 7

Conclusions & Outlook

The main goal of Chapters 2–4 was to advance the state of the art in QFI optimization for single-parameter estimation problems.

The emphasis was placed on the little-explored area of correlated quantum channels. The main challenge was to extend existing methods based on the quantum comb formalism to the large-N limit, where even storing the complete comb matrix in computer memory becomes impossible. To overcome this problem, tensor-network based (Chapter 2) or iterative (Chapters 3, 4) methods were applied.

The frequentist framework was a natural choice for exploring the uncharted waters of quantum metrology. Despite recent widespread interest in Bayesian [Rub21; Bav24; Zho24; Hay25; Alb25] and other non-local [Ges23; Mey25] approaches to quantum metrology , extending the results presented in this thesis beyond local estimation regime is still a challenge for the future.

Significant progress in Bayesian quantum metrology was recently made in Ref. [Zho24], which showed how to translate Bayesian problems into frequentist ones using the imaginary-time evolution technique. This enabled application of MOP-based methods to formulate Bayesian cost optimization as a single SDP. However, these results apply only to small N, and whether iterative techniques like those in Chapter 4 can extend them to large-N fundamental bounds remains unclear.

The channel discrimination problem [Bav21; Kat21a; Kat21b] can be viewed as a particular case of Bayesian problem with a discrete prior. For channel discrimination, the tensor-network based optimization from Chapter 2 has been recently successfully applied [Sie26]. Nevertheless, it is not obvious how to extend this result to continuous priors without drastically increasing the tensor network bond dimension.

It would be also interesting to extend the results from Chapters 2–4 to a multiparameter case. Nontrivial tradeoffs then appear because different input probes or estimation strategies are optimal for different estimated parameters, a phenomenon called probe incompatibility [Alb22; Alb25].

The methods proposed for correlated models are efficient for relatively small environmental space dimensions $\dim(\mathcal{E})$. However, this limitation can be sometimes

circumvented using the Markovian embedding technique [Ple20; Das25a], which enables treatment of even infinite-dimensional environments.

Regarding the results presented in Chapter 5, the introduced FI MeNoS quantity has been generalized to multiparameter scenarios in Ref. [Alb24], and the problem of noise vulnerability of quantum estimation protocols has been further examined in Ref. [For26].

The results of Chapter 6 are particularly interesting from an experimental perspective. I show that temporal signal correlations enable the image inversion interferometry (III) technique to extract information otherwise accessible only via higher H-G modes. The III technique is much simpler than measurement in the full H-G mode basis and has recently been demonstrated in realistic microscopic setups [Mit24]. Practical demonstration of fluctuation-enhanced III for imaging of complex sub-diffraction objects would be a valuable experimental direction.

Appendix A

Other publications

Published before PhD studies:

- **Super-resolution optical fluctuation imaging—fundamental estimation theory perspective**
  Stanisław Kurdziałek and Rafał Demkowicz-Dobrzański
  Journal of Optics **23**, 075701 (2021)
  DOI: 10.1088/2040-8986/ac059c

Published during PhD studies but not included in this thesis:

- **Back to sources – the role of losses and coherence in super-resolution imaging revisited**
  Stanisław Kurdziałek
  Quantum **6**, 697 (2022)
  DOI: 10.22331/q-2022-04-27-697

- **Noise-resistant phase imaging with intensity correlation**
  Jerzy Szuniewicz, Stanisław Kurdziałek, Sanjukta Kundu, Wojciech Zwolinski, Radosław Chrapkiewicz, Mayukh Lahiri, and Radek Lapkiewicz
  Science Advances **9**, adh5396 (2023)
  DOI: 10.1126/sciadv.adh5396

- **Certifying the quantum Fisher information from a given set of mean values: a semidefinite programming approach**
  Guillem Müller-Rigat, Anubhav Kumar Srivastava, Stanisław Kurdziałek, Grzegorz Rajchel-Mieldzioc, Maciej Lewenstein, and Irénée Frérot
  Quantum **7**, 1152 (2023)
  DOI: 10.22331/q-2023-10-24-1152

- **Optimal phase estimation in the presence of correlated dephasing**
  Srijon Ghosh, Arkadiusz Kobus, Stanisław Kurdziałek, and Rafał Demkowicz-Dobrzański
  arXiv preprint: 2511.07211 (2025)

- **QMetro++ – Python optimization package for large scale quantum metrology with customized strategy structures**
  Piotr Dulian, Stanisław Kurdziałek, and Rafał Demkowicz-Dobrzański
  Quantum **10**, 1991 (2026)
  DOI: 10.22331/q-2026-01-29-1991

Bibliography

[Aas13] J. Aasi et al. "Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light". In: *Nature Photonics* 7.8 (July 2013), pp. 613–619. ISSN: 1749-4893. DOI: 10.1038/nphoton.2013.177. URL: http://dx.doi.org/10.1038/nphoton.2013.177.

[Abb16] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116 (6 Feb. 2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

[Ací18] A. Acín et al. "The quantum technologies roadmap: a European community view". In: *New Journal of Physics* 20.8 (Aug. 2018), p. 080201. ISSN: 1367-2630. DOI: 10.1088/1367-2630/aad1ea. URL: http://dx.doi.org/10.1088/1367-2630/aad1ea.

[Aco09] V. M. Acosta et al. "Diamonds with a high density of nitrogen-vacancy centers for magnetometry applications". In: *Phys. Rev. B* 80 (11 Sept. 2009), p. 115202. DOI: 10.1103/PhysRevB.80.115202. URL: https://link.aps.org/doi/10.1103/PhysRevB.80.115202.

[Aff88] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki. "Valence bond ground states in isotropic quantum antiferromagnets". In: *Communications in Mathematical Physics* 115.3 (Sept. 1988), pp. 477–528. ISSN: 1432-0916. DOI: 10.1007/bf01218021. URL: http://dx.doi.org/10.1007/BF01218021.

[Alb22] F. Albarelli and R. Demkowicz-Dobrzański. "Probe Incompatibility in Multiparameter Noisy Quantum Metrology". In: *Phys. Rev. X* 12 (1 Mar. 2022), p. 011039. DOI: 10.1103/PhysRevX.12.011039. URL: https://link.aps.org/doi/10.1103/PhysRevX.12.011039.

[Alb24] F. Albarelli, I. Gianani, M. G. Genoni, and M. Barbieri. "Fisher-information susceptibility for multiparameter quantum estimation". In: *Phys. Rev. A* 110 (3 Sept. 2024), p. 032436. DOI: 10.1103/PhysRevA.110.032436. URL: https://link.aps.org/doi/10.1103/PhysRevA.110.032436.

[Alb25] F. Albarelli, D. Branford, and J. Rubio. *Measurement incompatibility in Bayesian multiparameter quantum estimation*. 2025. arXiv: 2511.16645 [quant-ph]. URL: https://arxiv.org/abs/2511.16645.

[Alm21] J. O. de Almeida, J. Kołodyński, C. Hirche, M. Lewenstein, and M. Skotiniotis. "Discrimination and estimation of incoherent sources under misalignment". In: *Phys. Rev. A* 103 (2 Feb. 2021), p. 022406. DOI: 10.

1103/PhysRevA.103.022406. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.022406.

[Alt21] A. Altherr and Y. Yang. "Quantum Metrology for Non-Markovian Processes". In: *Phys. Rev. Lett.* 127 (6 Aug. 2021), p. 060501. DOI: 10.1103/PhysRevLett.127.060501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.060501.

[Ara14] M. Araújo, F. Costa, and C. Č. Brukner. "Computational Advantage from Quantum-Controlled Ordering of Gates". In: *Phys. Rev. Lett.* 113 (25 Dec. 2014), p. 250402. DOI: 10.1103/PhysRevLett.113.250402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.250402.

[Bar20] J. F. Barry, J. M. Schloss, E. Bauch, M. J. Turner, C. A. Hart, L. M. Pham, and R. L. Walsworth. "Sensitivity optimization for NV-diamond magnetometry". In: *Rev. Mod. Phys.* 92 (1 Mar. 2020), p. 015004. DOI: 10.1103/RevModPhys.92.015004. URL: https://link.aps.org/doi/10.1103/RevModPhys.92.015004.

[Bar22] M. Barbieri. "Optical Quantum Metrology". In: *PRX Quantum* 3 (1 Jan. 2022), p. 010202. DOI: 10.1103/PRXQuantum.3.010202. URL: https://link.aps.org/doi/10.1103/PRXQuantum.3.010202.

[Bav21] J. Bavaresco, M. Murao, and M. T. Quintino. "Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination". In: *Phys. Rev. Lett.* 127 (20 Nov. 2021), p. 200504. DOI: 10.1103/PhysRevLett.127.200504. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.200504.

[Bav24] J. Bavaresco, P. Lipka-Bartosik, P. Sekatski, and M. Mehboudi. "Designing optimal protocols in Bayesian quantum parameter estimation with higher-order operations". In: *Phys. Rev. Res.* 6 (2 June 2024), p. 023305. DOI: 10.1103/PhysRevResearch.6.023305. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.6.023305.

[Bea18] F. Beaudoin, L. M. Norris, and L. Viola. "Ramsey interferometry in correlated quantum noise environments". In: *Phys. Rev. A* 98 (2 Aug. 2018), p. 020102. DOI: 10.1103/PhysRevA.98.020102. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.020102.

[Bel21] K. Beloy et al. "Frequency ratio measurements at 18-digit accuracy using an optical clock network". In: *Nature* 591.7851 (Mar. 2021), pp. 564–569. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03253-4. URL: http://dx.doi.org/10.1038/s41586-021-03253-4.

[Ber00] D. W. Berry and H. M. Wiseman. "Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry". In: *Phys. Rev. Lett.* 85 (24 Dec. 2000), pp. 5098–5101. DOI: 10.1103/PhysRevLett.85.5098. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.5098.

[Bet06] E. Betzig et al. "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313.5793 (2006), pp. 1642–1645. ISSN: 0036-8075. DOI: 10.1126/science.1127344. eprint: https://science.sciencemag.org/content/313/5793/1642.full.pdf. URL: https://science.sciencemag.org/content/313/5793/1642.

[Bol96] J. J. . Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen. "Optimal frequency measurements with maximally correlated states". In: *Phys. Rev. A* 54 (6 Dec. 1996), R4649–R4652. DOI: 10.1103/PhysRevA.54.R4649. URL: https://link.aps.org/doi/10.1103/PhysRevA.54.R4649.

[Bon84] R. S. Bondurant and J. H. Shapiro. "Squeezed states in phase-sensing interferometers". In: *Phys. Rev. D* 30 (12 Dec. 1984), pp. 2548–2556. DOI: 10.1103/PhysRevD.30.2548. URL: https://link.aps.org/doi/10.1103/PhysRevD.30.2548.

[Bor13] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[Bot00] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling. "Quantum Interferometric Optical Lithography: Exploiting Entanglement to Beat the Diffraction Limit". In: *Phys. Rev. Lett.* 85 (13 Sept. 2000), pp. 2733–2736. DOI: 10.1103/PhysRevLett.85.2733. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.2733.

[Boy04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[Bra15] J. B. Brask, R. Chaves, and J. Kołodyński. "Improved Quantum Magnetometry beyond the Standard Quantum Limit". In: *Phys. Rev. X* 5 (3 July 2015), p. 031010. DOI: 10.1103/PhysRevX.5.031010. URL: https://link.aps.org/doi/10.1103/PhysRevX.5.031010.

[Bra94] S. L. Braunstein and C. M. Caves. "Statistical distance and the geometry of quantum states". In: *Phys. Rev. Lett.* 72 (22 May 1994), pp. 3439–3443. DOI: 10.1103/PhysRevLett.72.3439. URL: https://link.aps.org/doi/10.1103/PhysRevLett.72.3439.

[Bre02] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. OUP Oxford, 2002.

[Bre16] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini. "Colloquium: Non-Markovian dynamics in open quantum systems". In: *Rev. Mod. Phys.* 88 (2 Apr. 2016), p. 021002. DOI: 10.1103/RevModPhys.88.021002. URL: https://link.aps.org/doi/10.1103/RevModPhys.88.021002.

[Bud07] D. Budker and M. Romalis. "Optical magnetometry". In: *Nature Physics* 3.4 (Apr. 2007), pp. 227–234. ISSN: 1745-2481. DOI: 10.1038/nphys566. URL: http://dx.doi.org/10.1038/nphys566.

[But24] E. P. Butler, G. E. Fux, C. Ortega-Taberner, B. W. Lovett, J. Keeling, and P. R. Eastham. "Optimizing Performance of Quantum Operations with Non-Markovian Decoherence: The Tortoise or the Hare?" In: *Phys. Rev. Lett.* 132 (6 Feb. 2024), p. 060401. DOI: 10.1103/PhysRevLett.132.060401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.132.060401.

[Cat15] C. Catana, L. Bouten, and M. Guţă. "Fisher informations and local asymptotic normality for continuous-time quantum Markov processes". In: *Journal of Physics A: Mathematical and Theoretical* 48.36 (Aug. 2015), p. 365301. ISSN: 1751-8121. DOI: 10.1088/1751-8113/48/36/365301. URL: http://dx.doi.org/10.1088/1751-8113/48/36/365301.

[Cav80] C. M. Caves. "Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer". In: *Phys. Rev. Lett.* 45 (2 July 1980), pp. 75–79. DOI: 10.1103/PhysRevLett.45.75. URL: https://link.aps.org/doi/10.1103/PhysRevLett.45.75.

[Cav81] C. M. Caves. "Quantum-mechanical noise in an interferometer". In: *Phys. Rev. D* 23 (8 Apr. 1981), pp. 1693–1708. DOI: 10.1103/PhysRevD.23.1693. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.1693.

[Cha20] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański. "Tensor-network approach for quantum metrology in many-body quantum systems". In: *Nature Communications* 11.1 (Jan. 2020). ISSN: 2041-1723. DOI: 10.1038/s41467-019-13735-9. URL: http://dx.doi.org/10.1038/s41467-019-13735-9.

[Cha21] F. ç. Chapeau-Blondeau. "Noisy quantum metrology with the assistance of indefinite causal order". In: *Phys. Rev. A* 103 (3 Mar. 2021), p. 032615. DOI: 10.1103/PhysRevA.103.032615. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.032615.

[Cha22] K. Chabuda and R. Demkowicz-Dobrzański. "TNQMetro: Tensor-network based package for efficient quantum metrology computations". In: *Computer Physics Communications* 274 (May 2022), p. 108282. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2021.108282. URL: http://dx.doi.org/10.1016/j.cpc.2021.108282.

[Chi08] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Quantum Circuit Architecture". In: *Phys. Rev. Lett.* 101 (6 Aug. 2008), p. 060401. DOI: 10.1103/PhysRevLett.101.060401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.101.060401.

[Chi09] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Theoretical framework for quantum networks". In: *Phys. Rev. A* 80 (2 Aug. 2009), p. 022339. DOI: 10.1103/PhysRevA.80.022339. URL: https://link.aps.org/doi/10.1103/PhysRevA.80.022339.

[Chi12] A. W. Chin, S. F. Huelga, and M. B. Plenio. "Quantum Metrology in Non-Markovian Environments". In: *Phys. Rev. Lett.* 109 (23 Dec. 2012), p. 233601. DOI: 10.1103/PhysRevLett.109.233601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.233601.

[Cic22] F. Ciccarello, S. Lorenzo, V. Giovannetti, and G. M. Palma. "Quantum collision models: Open system dynamics from repeated interactions". In: *Physics Reports* 954 (Apr. 2022), pp. 1–70. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2022.01.001. URL: http://dx.doi.org/10.1016/j.physrep.2022.01.001.

[Cza19] J. Czajkowski, K. Pawłowski, and R. Demkowicz-Dobrzański. "Many-body effects in quantum metrology". In: *New Journal of Physics* 21.5 (May 2019), p. 053031. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ab1fc2. URL: http://dx.doi.org/10.1088/1367-2630/ab1fc2.

[Das25a] A. Das and R. Demkowicz-Dobrzański. *Quantum metrology in presence of correlated noise via Markovian embedding*. 2025. arXiv: 2509.19685 [quant-ph]. URL: https://arxiv.org/abs/2509.19685.

[Das25b] A. Das, W. Górecki, and R. Demkowicz-Dobrzański. "Universal time scalings of sensitivity in Markovian quantum metrology". In: *Phys. Rev. A* 111 (2 Feb. 2025), p. L020403. DOI: 10.1103/PhysRevA.111.L020403. URL: https://link.aps.org/doi/10.1103/PhysRevA.111.L020403.

[Deg17] C. L. Degen, F. Reinhard, and P. Cappellaro. "Quantum sensing". In: *Rev. Mod. Phys.* 89.3 (2017), p. 035002. URL: https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.89.035002.

[Dem11] R. Demkowicz-Dobrzański. "Optimal phase estimation with arbitrary a priori knowledge". In: *Phys. Rev. A* 83 (6 June 2011), p. 061802. DOI: 10.1103/PhysRevA.83.061802. URL: https://link.aps.org/doi/10.1103/PhysRevA.83.061802.

[Dem12] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă. "The Elusive Heisenberg Limit in Quantum-Enhanced Metrology". In: *Nat. Commun.* 3 (Sept. 2012), p. 1063. DOI: 10.1038/ncomms2067. arXiv: 1201.3940.

[Dem13] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel. "Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector GEO 600". In: *Phys. Rev. A* 88 (4 Oct. 2013), p. 041802. DOI: 10.1103/PhysRevA.88.041802. URL: https://link.aps.org/doi/10.1103/PhysRevA.88.041802.

[Dem14] R. Demkowicz-Dobrzański and L. Maccone. "Using Entanglement Against Noise in Quantum Metrology". In: *Phys. Rev. Lett.* 113.25 (Dec. 2014), p. 250801. DOI: 10.1103/PhysRevLett.113.250801. arXiv: 1407.2934.

[Dem15] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński. "Quantum Limits in Optical Interferometry". In: *Progress in Optics*. Elsevier, 2015, pp. 345–435. ISBN: 9780128022849. DOI: 10.1016/bs.po.2015.02.003. URL: http://dx.doi.org/10.1016/bs.po.2015.02.003.

[Dem17] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski. "Adaptive Quantum Metrology under General Markovian Noise". In: *Phys. Rev. X* 7.4 (Oct. 2017), p. 041009. DOI: 10.1103/PhysRevX.7.041009. arXiv: 1704.06280.

[Dem20] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă. "Multi-parameter estimation beyond quantum Fisher information". In: *Journal of Physics A: Mathematical and Theoretical* 53.36 (Aug. 2020), p. 363001. ISSN: 1751-8121. DOI: 10.1088/1751-8121/ab8ef3. URL: http://dx.doi.org/10.1088/1751-8121/ab8ef3.

[Der09] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein. "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)". In: *Proceedings of the National Academy of Sciences* 106.52 (2009), pp. 22287–22292. ISSN: 0027-8424. DOI: 10.1073/pnas.0907866106. eprint: https://www.pnas.org/content/106/52/22287.full.pdf. URL: https://www.pnas.org/content/106/52/22287.

[Dow03] J. P. Dowling and G. J. Milburn. "Quantum technology: the second quantum revolution". In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (June 2003). Ed. by A. G. J. MacFarlane, pp. 1655–1674. ISSN:

1471-2962. DOI: 10.1098/rsta.2003.1227. URL: http://dx.doi.org/10.1098/rsta.2003.1227.

[Dow08] J. P. Dowling. "Quantum optical metrology – the lowdown on high-N00N states". In: *Contemporary Physics* 49.2 (Mar. 2008), pp. 125–143. ISSN: 1366-5812. DOI: 10.1080/00107510802091298. URL: http://dx.doi.org/10.1080/00107510802091298.

[Dow98] J. P. Dowling. "Correlated input-port, matter-wave interferometer: Quantum-noise limits to the atom-laser gyroscope". In: *Phys. Rev. A* 57 (6 June 1998), pp. 4736–4746. DOI: 10.1103/PhysRevA.57.4736. URL: https://link.aps.org/doi/10.1103/PhysRevA.57.4736.

[Du24] J. Du, F. Shi, X. Kong, F. Jelezko, and J. Wrachtrup. "Single-molecule scale magnetic resonance spectroscopy using quantum diamond sensors". In: *Rev. Mod. Phys.* 96 (2 May 2024), p. 025001. DOI: 10.1103/RevModPhys.96.025001. URL: https://link.aps.org/doi/10.1103/RevModPhys.96.025001.

[Dul26] P. Dulian, S. Kurdziałek, and R. Demkowicz-Dobrzański. "QMetro++ – Python optimization package for large scale quantum metrology with customized strategy structures". In: *Quantum* 10 (Jan. 2026), p. 1991. ISSN: 2521-327X. DOI: 10.22331/q-2026-01-29-1991. URL: https://doi.org/10.22331/q-2026-01-29-1991.

[Dup25] A. Duplinskiy, J. Frank, K. Bearne, and A. I. Lvovsky. "Tsang's resolution enhancement method for imaging with focused illumination". In: *Light: Science & Applications* 14.1 (Apr. 2025). ISSN: 2047-7538. DOI: 10.1038/s41377-025-01791-4. URL: https://dx.doi.org/10.1038/s41377-025-01791-4.

[Esc11] B. M. Escher, R. L. de Matos Filho, and L. Davidovich. "General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology". In: *Nature Physics* 7.5 (Mar. 2011), pp. 406–411. ISSN: 1745-2481. DOI: 10.1038/nphys1958. URL: http://dx.doi.org/10.1038/nphys1958.

[Fan92] M. Fannes, B. Nachtergaele, and R. F. Werner. "Finitely correlated states on quantum spin chains". In: *Communications in Mathematical Physics* 144.3 (Mar. 1992), pp. 443–490. ISSN: 1432-0916. DOI: 10.1007/bf02099178. URL: http://dx.doi.org/10.1007/BF02099178.

[For26] A. K. Forbes, M. A. Rodríguez-García, and I. H. Deutsch. *Fragility of Optimal Measurements due to Noise in Probe States for Quantum Sensing*. 2026. arXiv: 2601.08712 [quant-ph]. URL: https://arxiv.org/abs/2601.08712.

[Fra23] J. Frank, A. Duplinskiy, K. Bearne, and A. I. Lvovsky. "Passive super-resolution imaging of incoherent objects". In: *Optica* 10.9 (Sept. 2023), pp. 1147–1152. DOI: 10.1364/OPTICA.493718. URL: https://opg.optica.org/optica/abstract.cfm?URI=optica-10-9-1147.

[Fre19] M. Frey. "Indefinite causal order aids quantum depolarizing channel identification". In: *Quantum Information Processing* 18.4 (Feb. 2019). ISSN: 1573-1332. DOI: 10.1007/s11128-019-2186-9. URL: http://dx.doi.org/10.1007/s11128-019-2186-9.

[Fuj01] A. Fujiwara. "Quantum channel identification problem". In: *Phys. Rev. A* 63 (4 Mar. 2001), p. 042304. DOI: 10.1103/PhysRevA.63.042304. URL: https://link.aps.org/doi/10.1103/PhysRevA.63.042304.

[Fuj08] A. Fujiwara and H. Imai. "A Fibre Bundle over Manifolds of Quantum Channels and Its Application to Quantum Statistics". In: *J. Phys. A* 41.25 (June 2008), p. 255304. DOI: 10.1088/1751-8113/41/25/255304.

[Gat14] D. Gatto Monticone et al. "Beating the Abbe Diffraction Limit in Confocal Microscopy via Nonclassical Photon Statistics". In: *Phys. Rev. Lett.* 113 (14 Sept. 2014), p. 143602. DOI: 10.1103/PhysRevLett.113.143602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.143602.

[Ges20] M. Gessner, C. Fabre, and N. Treps. "Superresolution Limits from Measurement Crosstalk". In: *Phys. Rev. Lett.* 125 (10 Aug. 2020), p. 100501. DOI: 10.1103/PhysRevLett.125.100501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.100501.

[Ges23] M. Gessner and A. Smerzi. "Hierarchies of Frequentist Bounds for Quantum Metrology: From Cramér-Rao to Barankin". In: *Phys. Rev. Lett.* 130 (26 June 2023), p. 260801. DOI: 10.1103/PhysRevLett.130.260801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.260801.

[Gio04] V. Giovannetti, S. Lloyd, and L. Maccone. "Quantum-Enhanced Measurements: Beating the Standard Quantum Limit". In: *Science* 306.5700 (2004), pp. 1330–1336. DOI: 10.1126/science.1104149. eprint: https://www.science.org/doi/pdf/10.1126/science.1104149. URL: https://www.science.org/doi/abs/10.1126/science.1104149.

[Gio06] V. Giovannetti, S. Lloyd, and L. Maccone. "Quantum Metrology". In: *Phys. Rev. Lett.* 96 (1 Jan. 2006), p. 010401. DOI: 10.1103/PhysRevLett.96.010401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.010401.

[God23] A. Godley and M. Guta. "Adaptive measurement filter: efficient strategy for optimal estimation of quantum Markov chains". In: *Quantum* 7 (Apr. 2023), p. 973. ISSN: 2521-327X. DOI: 10.22331/q-2023-04-06-973. URL: https://doi.org/10.22331/q-2023-04-06-973.

[Gór20] W. Górecki, R. Demkowicz-Dobrzański, H. M. Wiseman, and D. W. Berry. "π-Corrected Heisenberg Limit". In: *Phys. Rev. Lett.* 124 (3 Jan. 2020), p. 030501. DOI: 10.1103/PhysRevLett.124.030501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.030501.

[Gór25] W. Górecki, X. Lu, C. Macchiavello, and L. Maccone. "Mutual information bounded by Fisher information". In: *Phys. Rev. Res.* 7 (2 Apr. 2025), p. L022013. DOI: 10.1103/PhysRevResearch.7.L022013. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.7.L022013.

[Gor76] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. "Completely positive dynamical semigroups of N-level systems". In: *Journal of Mathematical Physics* 17.5 (May 1976), pp. 821–825. ISSN: 1089-7658. DOI: 10.1063/1.522979. URL: http://dx.doi.org/10.1063/1.522979.

[Gos18] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White. "Indefinite Causal Order in a Quantum Switch". In: *Phys. Rev. Lett.* 121 (9 Aug. 2018), p. 090503. DOI: 10.

1103/PhysRevLett.121.090503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.090503.

[Guţ11] M. Guţă. "Fisher information and asymptotic normality in system identification for quantum Markov chains". In: *Phys. Rev. A* 83 (6 June 2011), p. 062324. DOI: 10.1103/PhysRevA.83.062324. URL: https://link.aps.org/doi/10.1103/PhysRevA.83.062324.

[Haa18] J. F. Haase, A. Smirne, J. Kołodyński, R. Demkowicz-Dobrzański, and S. F. Huelga. "Fundamental limits to frequency estimation: a comprehensive microscopic perspective". In: *New Journal of Physics* 20.5 (May 2018), p. 053009. ISSN: 1367-2630. DOI: 10.1088/1367-2630/aab67f. URL: http://dx.doi.org/10.1088/1367-2630/aab67f.

[Hay25] M. Hayashi. "Indefinite causal order strategy does not improve the estimation of group action". In: *Quantum* 9 (Oct. 2025), p. 1891. ISSN: 2521-327X. DOI: 10.22331/q-2025-10-22-1891. URL: http://dx.doi.org/10.22331/q-2025-10-22-1891.

[Hel69] C. W. Helstrom. "Quantum detection and estimation theory". In: *Journal of Statistical Physics* 1.2 (1969), pp. 231–252.

[Hin13] N. Hinkley et al. "An Atomic Clock with 10 –18 Instability". In: *Science* 341.6151 (Sept. 2013), pp. 1215–1218. ISSN: 1095-9203. DOI: 10.1126/science.1240420. URL: http://dx.doi.org/10.1126/science.1240420.

[Hol11] A. S. Holevo. *Probabilistic and statistical aspects of quantum theory*. Vol. 1. Springer Science & Business Media, 2011.

[Hol79] J. N. Hollenhorst. "Quantum limits on resonant-mass gravitational-radiation detectors". In: *Phys. Rev. D* 19 (6 Mar. 1979), pp. 1669–1679. DOI: 10.1103/PhysRevD.19.1669. URL: https://link.aps.org/doi/10.1103/PhysRevD.19.1669.

[Hol93] M. J. Holland and K. Burnett. "Interferometric detection of optical phase shifts at the Heisenberg limit". In: *Phys. Rev. Lett.* 71 (9 Aug. 1993), pp. 1355–1358. DOI: 10.1103/PhysRevLett.71.1355. URL: https://link.aps.org/doi/10.1103/PhysRevLett.71.1355.

[Hor09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. "Quantum entanglement". In: *Rev. Mod. Phys.* 81 (2 June 2009), pp. 865–942. DOI: 10.1103/RevModPhys.81.865. URL: https://link.aps.org/doi/10.1103/RevModPhys.81.865.

[Hor12] T. Horrom, R. Singh, J. P. Dowling, and E. E. Mikhailov. "Quantum-enhanced magnetometer with low-frequency squeezing". In: *Phys. Rev. A* 86 (2 Aug. 2012), p. 023803. DOI: 10.1103/PhysRevA.86.023803. URL: https://link.aps.org/doi/10.1103/PhysRevA.86.023803.

[Hor25] D. B. Horoshko, A. B. Mikhalychev, F. Jelezko, and P. P. Kuzhir. "Weak convexity of Fisher information matrix and superresolved localization of blinking sources of light". In: *Phys. Rev. Res.* 7 (3 July 2025), p. 033065. DOI: 10.1103/t48h-6hcf. URL: https://link.aps.org/doi/10.1103/t48h-6hcf.

[Hue97] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac. "Improvement of Frequency Standards with Quantum

Entanglement". In: *Phys. Rev. Lett.* 79 (20 Nov. 1997), pp. 3865–3868. DOI: 10.1103/PhysRevLett.79.3865. URL: https://link.aps.org/doi/10.1103/PhysRevLett.79.3865.

[Ita90] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. "Quantum Zeno effect". In: *Phys. Rev. A* 41 (5 Mar. 1990), pp. 2295–2300. DOI: 10.1103/PhysRevA.41.2295. URL: https://link.aps.org/doi/10.1103/PhysRevA.41.2295.

[Ita93] W. M. Itano et al. "Quantum projection noise: Population fluctuations in two-level systems". In: *Phys. Rev. A* 47 (5 May 1993), pp. 3554–3570. DOI: 10.1103/PhysRevA.47.3554. URL: https://link.aps.org/doi/10.1103/PhysRevA.47.3554.

[Jar13] M. Jarzyna and R. Demkowicz-Dobrzański. "Matrix Product States for Quantum Metrology". In: *Phys. Rev. Lett.* 110 (24 June 2013), p. 240405. DOI: 10.1103/PhysRevLett.110.240405. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.240405.

[Jar15] M. Jarzyna and R. Demkowicz-Dobrzański. "True precision limits in quantum metrology". In: *New Journal of Physics* 17.1 (Jan. 2015), p. 013010. DOI: 10.1088/1367-2630/17/1/013010. URL: https://doi.org/10.1088/1367-2630/17/1/013010.

[Jes14] J. Jeske, J. H. Cole, and S. F. Huelga. "Quantum metrology subject to spatially correlated Markovian noise: restoring the Heisenberg limit". In: *New Journal of Physics* 16.7 (July 2014), p. 073039. ISSN: 1367-2630. DOI: 10.1088/1367-2630/16/7/073039. URL: http://dx.doi.org/10.1088/1367-2630/16/7/073039.

[Kat11] H. Katori. "Optical lattice clocks and quantum metrology". In: *Nature Photonics* 5.4 (Mar. 2011), pp. 203–210. ISSN: 1749-4893. DOI: 10.1038/nphoton.2011.45. URL: http://dx.doi.org/10.1038/nphoton.2011.45.

[Kat21a] V. Katariya and M. M. Wilde. "Evaluating the advantage of adaptive strategies for quantum channel distinguishability". In: *Phys. Rev. A* 104 (5 Nov. 2021), p. 052406. DOI: 10.1103/PhysRevA.104.052406. URL: https://link.aps.org/doi/10.1103/PhysRevA.104.052406.

[Kat21b] V. Katariya and M. M. Wilde. "Geometric distinguishability measures limit quantum channel estimation and discrimination". In: *Quantum Information Processing* 20.2 (Feb. 2021). ISSN: 1573-1332. DOI: 10.1007/s11128-021-02992-7. URL: http://dx.doi.org/10.1007/s11128-021-02992-7.

[Kay93] S. M. Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[Kit93] M. Kitagawa and M. Ueda. "Squeezed spin states". In: *Phys. Rev. A* 47 (6 June 1993), pp. 5138–5143. DOI: 10.1103/PhysRevA.47.5138. URL: https://link.aps.org/doi/10.1103/PhysRevA.47.5138.

[Koł13] J. Kołodyński and R. Demkowicz-Dobrzański. "Efficient tools for quantum metrology with uncorrelated noise". In: *New Journal of Physics* 15.7 (July 2013), p. 073043. ISSN: 1367-2630. DOI: 10.1088/1367-2630/15/7/073043. URL: http://dx.doi.org/10.1088/1367-2630/15/7/073043.

[Kur21] S. Kurdziałek and R. Demkowicz-Dobrzański. "Super-resolution optical fluctuation imaging—fundamental estimation theory perspective". In: *Journal of Optics* 23.7 (2021), p. 075701.

[Kur23a] S. Kurdziałek and R. Demkowicz-Dobrzański. "Measurement Noise Susceptibility in Quantum Estimation". In: *Phys. Rev. Lett.* 130 (16 Apr. 2023), p. 160802. DOI: 10.1103/PhysRevLett.130.160802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.160802.

[Kur23b] S. Kurdziałek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański. "Using Adaptiveness and Causal Superpositions Against Noise in Quantum Metrology". In: *Phys. Rev. Lett.* 131 (9 Aug. 2023), p. 090801. DOI: 10.1103/PhysRevLett.131.090801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.090801.

[Kur25a] S. Kurdziałek. *Super-resolution microscopy via fluctuation-enhanced spatial mode demultiplexing*. 2025. arXiv: 2511.20790 [physics.optics]. URL: https://arxiv.org/abs/2511.20790.

[Kur25b] S. Kurdziałek, F. Albarelli, and R. Demkowicz-Dobrzański. "Universal Bounds for Quantum Metrology in the Presence of Correlated Noise". In: *Phys. Rev. Lett.* 135 (13 Sept. 2025), p. 130801. DOI: 10.1103/jy3v-wkcb. URL: https://link.aps.org/doi/10.1103/jy3v-wkcb.

[Kur25c] S. Kurdziałek, P. Dulian, J. Majsak, S. Chakraborty, and R. Demkowicz-Dobrzański. "Quantum metrology using quantum combs and tensor network formalism". In: *New Journal of Physics* 27.1 (Jan. 2025), p. 013019. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ada8d1. URL: http://dx.doi.org/10.1088/1367-2630/ada8d1.

[Lay18] D. Layden and P. Cappellaro. "Spatial noise filtering through error correction for quantum sensing". In: *npj Quantum Information* 4.1 (July 2018). ISSN: 2056-6387. DOI: 10.1038/s41534-018-0082-2. URL: http://dx.doi.org/10.1038/s41534-018-0082-2.

[Lay19] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang. "Ancilla-Free Quantum Error Correction Codes for Quantum Metrology". In: *Phys. Rev. Lett.* 122 (4 Jan. 2019), p. 040502. DOI: 10.1103/PhysRevLett.122.040502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.040502.

[Lee02] H. Lee, P. Kok, and J. P. Dowling. "A quantum Rosetta stone for interferometry". In: *Journal of Modern Optics* 49.14–15 (Nov. 2002), pp. 2325–2338. ISSN: 1362-3044. DOI: 10.1080/0950034021000011536. URL: http://dx.doi.org/10.1080/0950034021000011536.

[Len22] Y. L. Len, T. Gefen, A. Retzker, and J. Kołodyński. "Quantum metrology with imperfect measurements". In: *Nature Communications* 13.1 (Nov. 2022). ISSN: 2041-1723. DOI: 10.1038/s41467-022-33563-8. URL: http://dx.doi.org/10.1038/s41467-022-33563-8.

[Leu11] B. O. Leung and K. C. Chou. "Review of Super-Resolution Fluorescence Microscopy for Biology". In: *Applied Spectroscopy* 65.9 (2011). PMID: 21929850, pp. 967–980. DOI: 10.1366/11-06398. eprint: https://doi.org/10.1366/11-06398. URL: https://doi.org/10.1366/11-06398.

[Lin76] G. Lindblad. "On the generators of quantum dynamical semigroups". In: *Communications in Mathematical Physics* 48.2 (June 1976), pp. 119–130.

ISSN: 1432-0916. DOI: 10.1007/bf01608499. URL: http://dx.doi.org/10.1007/BF01608499.

[Liu23] Q. Liu, Z. Hu, H. Yuan, and Y. Yang. "Optimal Strategies of Quantum Metrology with a Strict Hierarchy". In: *Phys. Rev. Lett.* 130 (7 Feb. 2023), p. 070803. DOI: 10.1103/PhysRevLett.130.070803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.070803.

[Liu24] Q. Liu and Y. Yang. "Efficient tensor networks for control-enhanced quantum metrology". In: *Quantum* 8 (Dec. 2024), p. 1571. ISSN: 2521-327X. DOI: 10.22331/q-2024-12-18-1571. URL: https://doi.org/10.22331/q-2024-12-18-1571.

[Lud15] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt. "Optical atomic clocks". In: *Rev. Mod. Phys.* 87 (2 June 2015), pp. 637–701. DOI: 10.1103/RevModPhys.87.637. URL: https://link.aps.org/doi/10.1103/RevModPhys.87.637.

[Lup20] C. Lupo. "Subwavelength quantum imaging with noisy detectors". In: *Phys. Rev. A* 101 (2 Feb. 2020), p. 022323. DOI: 10.1103/PhysRevA.101.022323. URL: https://link.aps.org/doi/10.1103/PhysRevA.101.022323.

[Mac13] K. Macieszczak. "Quantum Fisher Information: Variational principle and simple iterative algorithm for its efficient computation". In: *arXiv e-prints*, arXiv:1312.1356 (Dec. 2013), arXiv:1312.1356. arXiv: 1312.1356 [quant-ph].

[Mac14] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański. "Bayesian quantum frequency estimation in presence of collective dephasing". In: *New Journal of Physics* 16.11 (2014), p. 113002. URL: http://stacks.iop.org/1367-2630/16/i=11/a=113002.

[Mac15] K. Macieszczak. "Zeno limit in frequency estimation with non-Markovian environments". In: *Phys. Rev. A* 92 (1 July 2015), p. 010102. DOI: 10.1103/PhysRevA.92.010102. URL: https://link.aps.org/doi/10.1103/PhysRevA.92.010102.

[Mat11] Y. Matsuzaki, S. C. Benjamin, and J. Fitzsimons. "Magnetic field sensing beyond the standard quantum limit under the effect of decoherence". In: *Phys. Rev. A* 84 (1 July 2011), p. 012103. DOI: 10.1103/PhysRevA.84.012103. URL: https://link.aps.org/doi/10.1103/PhysRevA.84.012103.

[Maz08] J. R. Maze et al. "Nanoscale magnetic sensing with an individual electronic spin in diamond". In: *Nature* 455.7213 (Oct. 2008), pp. 644–647. ISSN: 1476-4687. DOI: 10.1038/nature07279. URL: http://dx.doi.org/10.1038/nature07279.

[Men25] T. M. Mendonça, D. O. Soares-Pinto, and M. Paternostro. "Information flow-enhanced precision in collisional quantum thermometry". In: *Phys. Rev. Res.* 7 (4 Nov. 2025), p. 043203. DOI: 10.1103/ymjl-hpm4. URL: https://link.aps.org/doi/10.1103/ymjl-hpm4.

[Mey25] J. J. Meyer, S. Khatri, D. Stilck Fran ça, J. Eisert, and P. Faist. "Quantum Metrology in the Finite-Sample Regime". In: *PRX Quantum* 6 (3 Aug. 2025), p. 030336. DOI: 10.1103/qbn1-p6bq. URL: https://link.aps.org/doi/10.1103/qbn1-p6bq.

[Mit24] C. S. Mitchell, D. Dhruva, Z. P. Burke, D. J. Durden, A. I. Dingilian, and M. P. Backlund. *Quantum-inspired super-resolution of fluorescent point-like sources*. 2024. DOI: 10.48550/ARXIV.2412.16835. URL: https://arxiv.org/abs/2412.16835.

[Moe15] W. E. Moerner. "Nobel Lecture: Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy". In: *Rev. Mod. Phys.* 87 (4 Oct. 2015), pp. 1183–1212. DOI: 10.1103/RevModPhys.87.1183. URL: https://link.aps.org/doi/10.1103/RevModPhys.87.1183.

[Mor19] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett. "Imaging with quantum states of light". In: *Nature Reviews Physics* 1.6 (May 2019), pp. 367–380. ISSN: 2522-5820. DOI: 10.1038/s42254-019-0056-0. URL: http://dx.doi.org/10.1038/s42254-019-0056-0.

[Nag07] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi. "Beating the Standard Quantum Limit with Four-Entangled Photons". In: *Science* 316.5825 (May 2007), pp. 726–729. ISSN: 1095-9203. DOI: 10.1126/science.1138007. URL: http://dx.doi.org/10.1126/science.1138007.

[Nai16] R. Nair and M. Tsang. "Far-Field Superresolution of Thermal Electromagnetic Sources at the Quantum Limit". In: *Physical Review Letters* 117.19 (Nov. 2016). ISSN: 1079-7114. DOI: 10.1103/physrevlett.117.190801. URL: http://dx.doi.org/10.1103/PhysRevLett.117.190801.

[Nie10] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[Oh21] C. Oh, S. Zhou, Y. Wong, and L. Jiang. "Quantum Limits of Superresolution in a Noisy Environment". In: *Phys. Rev. Lett.* 126 (12 Mar. 2021), p. 120502. DOI: 10.1103/PhysRevLett.126.120502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.120502.

[Öst95] S. Östlund and S. Rommer. "Thermodynamic Limit of Density Matrix Renormalization". In: *Phys. Rev. Lett.* 75 (19 Nov. 1995), pp. 3537–3540. DOI: 10.1103/PhysRevLett.75.3537. URL: https://link.aps.org/doi/10.1103/PhysRevLett.75.3537.

[Par18] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański. "Beating the Rayleigh Limit Using Two-Photon Interference". In: *Phys. Rev. Lett.* 121 (25 Dec. 2018), p. 250503. DOI: 10.1103/PhysRevLett.121.250503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.250503.

[Per21] J. L. Pereira, L. Banchi, and S. Pirandola. "Bounding the Benefit of Adaptivity in Quantum Metrology Using the Relative Fidelity". In: *Phys. Rev. Lett.* 127 (15 Oct. 2021), p. 150501. DOI: 10.1103/PhysRevLett.127.150501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.150501.

[Pez18] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein. "Quantum metrology with nonclassical states of atomic ensembles". In: *Rev. Mod. Phys.* 90 (3 Sept. 2018), p. 035005. DOI: 10.1103/RevModPhys.

90.035005. URL: https://link.aps.org/doi/10.1103/RevModPhys.90.035005.

[Pir17] S. Pirandola and C. Lupo. "Ultimate Precision of Adaptive Noise Estimation". In: *Phys. Rev. Lett.* 118 (10 Mar. 2017), p. 100502. DOI: 10.1103/PhysRevLett.118.100502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.100502.

[Pir18] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd. "Advances in photonic quantum sensing". In: *Nature Photonics* 12.12 (Nov. 2018), pp. 724–733. ISSN: 1749-4893. DOI: 10.1038/s41566-018-0301-6. URL: http://dx.doi.org/10.1038/s41566-018-0301-6.

[Pla22] G. Planella, M. F. B. Cenni, A. Acín, and M. Mehboudi. "Bath-Induced Correlations Enhance Thermometry Precision at Low Temperatures". In: *Phys. Rev. Lett.* 128 (4 Jan. 2022), p. 040502. DOI: 10.1103/PhysRevLett.128.040502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.128.040502.

[Ple20] G. Pleasance, B. M. Garraway, and F. Petruccione. "Generalized theory of pseudomodes for exact descriptions of non-Markovian quantum processes". In: *Phys. Rev. Res.* 2 (4 Oct. 2020), p. 043058. DOI: 10.1103/PhysRevResearch.2.043058. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.043058.

[Pol20] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino. "Photonic quantum metrology". In: *AVS Quantum Science* 2.2 (June 2020). ISSN: 2639-0213. DOI: 10.1116/5.0007577. URL: http://dx.doi.org/10.1116/5.0007577.

[Pus21] A. A. Pushkina, G. Maltese, J. I. Costa-Filho, P. Patel, and A. I. Lvovsky. "Superresolution Linear Optical Imaging in the Far Field". In: *Phys. Rev. Lett.* 127 (25 Dec. 2021), p. 253602. DOI: 10.1103/PhysRevLett.127.253602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.253602.

[Ray79] Rayleigh. "XXXI. Investigations in optics, with special reference to the spectroscope". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.49 (Oct. 1879), pp. 261–274. ISSN: 1941-5990. DOI: 10.1080/14786447908639684. URL: http://dx.doi.org/10.1080/14786447908639684.

[Rib22] F. Riberi, L. M. Norris, F. Beaudoin, and L. Viola. "Frequency estimation under non-Markovian spatially correlated quantum noise". In: *New Journal of Physics* 24.10 (Oct. 2022), p. 103011. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ac92a2. URL: http://dx.doi.org/10.1088/1367-2630/ac92a2.

[Ron14] L. Rondin, J.-P. Tetienne, T. Hingant, J.-F. Roch, P. Maletinsky, and V. Jacques. "Magnetometry with nitrogen-vacancy defects in diamond". In: *Reports on Progress in Physics* 77.5 (May 2014), p. 056503. ISSN: 1361-6633. DOI: 10.1088/0034-4885/77/5/056503. URL: http://dx.doi.org/10.1088/0034-4885/77/5/056503.

[Roz14] L. A. Rozema, J. D. Bateman, D. H. Mahler, R. Okamoto, A. Feizpour, A. Hayat, and A. M. Steinberg. "Scalable Spatial Superresolution Using

Entangled Photons". In: *Phys. Rev. Lett.* 112.22 (June 2014), p. 223602. ISSN: 0031-9007.

[Roz24] L. A. Rozema, T. Strömberg, H. Cao, Y. Guo, B.-H. Liu, and P. Walther. "Experimental aspects of indefinite causal order in quantum mechanics". In: *Nature Reviews Physics* 6.8 (July 2024), pp. 483–499. ISSN: 2522-5820. DOI: 10.1038/s42254-024-00739-8. URL: http://dx.doi.org/10.1038/s42254-024-00739-8.

[Rub21] J. Rubio, J. Anders, and L. A. Correa. "Global Quantum Thermometry". In: *Phys. Rev. Lett.* 127 (19 Nov. 2021), p. 190402. DOI: 10.1103/PhysRevLett.127.190402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.190402.

[Rus06] M. J. Rust, M. Bates, and X. Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature Methods* 3.10 (Oct. 2006), pp. 793–796. ISSN: 1548-7105. DOI: 10.1038/nmeth929. URL: https://doi.org/10.1038/nmeth929.

[Sar06] M. Sarovar and G. J. Milburn. "Optimal estimation of one-parameter quantum channels". In: *Journal of Physics A: Mathematical and General* 39.26 (June 2006), p. 8487. DOI: 10.1088/0305-4470/39/26/015. URL: https://doi.org/10.1088/0305-4470/39/26/015.

[Sch12] O. Schwartz and D. Oron. "Improved resolution in fluorescence microscopy using quantum correlations". In: *Phys. Rev. A* 85 (3 Mar. 2012), p. 033812. DOI: 10.1103/PhysRevA.85.033812. URL: https://link.aps.org/doi/10.1103/PhysRevA.85.033812.

[Sch13] O. Schwartz, J. M. Levitt, R. Tenne, S. Itzhakov, Z. Deutsch, and D. Oron. "Superresolution microscopy with quantum emitters". In: *Nano letters* 13.12 (2013), pp. 5832–5836.

[Sea19] S. Seah, S. Nimmrichter, D. Grimmer, J. P. Santos, V. Scarani, and G. T. Landi. "Collisional Quantum Thermometry". In: *Phys. Rev. Lett.* 123 (18 Oct. 2019), p. 180602. DOI: 10.1103/PhysRevLett.123.180602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.180602.

[Sek17] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür. "Quantum metrology with full and fast quantum control". In: *Quantum* 1 (Sept. 2017), p. 27. ISSN: 2521-327X. DOI: 10.22331/q-2017-09-06-27. URL: http://dx.doi.org/10.22331/q-2017-09-06-27.

[Sew12] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell. "Magnetic Sensitivity Beyond the Projection Noise Limit by Spin Squeezing". In: *Phys. Rev. Lett.* 109 (25 Dec. 2012), p. 253605. DOI: 10.1103/PhysRevLett.109.253605. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.253605.

[Shu20] A. Shu, S. Seah, and V. Scarani. "Surpassing the thermal Cramér-Rao bound with collisional thermometry". In: *Phys. Rev. A* 102 (4 Oct. 2020), p. 042417. DOI: 10.1103/PhysRevA.102.042417. URL: https://link.aps.org/doi/10.1103/PhysRevA.102.042417.

[Sie26] S. Sieniawski and R. Demkowicz-Dobrzanski. "Adaptive quantum channel discrimination using methods of quantum metrology". In: *New Journal of*

Physics (Jan. 2026). ISSN: 1367-2630. DOI: 10.1088/1367-2630/ae3edd. URL: http://dx.doi.org/10.1088/1367-2630/ae3edd.

[Smi16] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański. "Ultimate Precision Limits for Noisy Frequency Estimation". In: *Phys. Rev. Lett.* 116 (12 Mar. 2016), p. 120801. DOI: 10.1103/PhysRevLett.116.120801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.120801.

[Str17] A. Streltsov, G. Adesso, and M. B. Plenio. "Colloquium: Quantum coherence as a resource". In: *Rev. Mod. Phys.* 89 (4 Oct. 2017), p. 041003. DOI: 10.1103/RevModPhys.89.041003. URL: https://link.aps.org/doi/10.1103/RevModPhys.89.041003.

[Sut16] D. Suter and G. A. Álvarez. "Colloquium: Protecting quantum information against environmental noise". In: *Rev. Mod. Phys.* 88 (4 Oct. 2016), p. 041001. DOI: 10.1103/RevModPhys.88.041001. URL: https://link.aps.org/doi/10.1103/RevModPhys.88.041001.

[Szi20] S. S. Szigeti, S. P. Nolan, J. D. Close, and S. A. Haine. "High-Precision Quantum-Enhanced Gravimetry with a Bose-Einstein Condensate". In: *Phys. Rev. Lett.* 125 (10 Sept. 2020), p. 100402. DOI: 10.1103/PhysRevLett.125.100402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.100402.

[Tho80] K. S. Thorne. "Gravitational-wave research: Current status and future prospects". In: *Rev. Mod. Phys.* 52 (2 Apr. 1980), pp. 285–297. DOI: 10.1103/RevModPhys.52.285. URL: https://link.aps.org/doi/10.1103/RevModPhys.52.285.

[Tsa11] M. Tsang, H. M. Wiseman, and C. M. Caves. "Fundamental Quantum Limit to Waveform Estimation". In: *Phys. Rev. Lett.* 106 (9 Mar. 2011), p. 090401. DOI: 10.1103/PhysRevLett.106.090401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.090401.

[Tsa16] M. Tsang, R. Nair, and X.-M. Lu. "Quantum Theory of Superresolution for Two Incoherent Optical Point Sources". In: *Phys. Rev. X* 6 (3 Aug. 2016), p. 031033. DOI: 10.1103/PhysRevX.6.031033. URL: https://link.aps.org/doi/10.1103/PhysRevX.6.031033.

[Tsa17] M. Tsang. "Subdiffraction incoherent optical imaging via spatial-mode demultiplexing". In: *New Journal of Physics* 19.2 (Feb. 2017), p. 023054. DOI: 10.1088/1367-2630/aa60ee. URL: https://dx.doi.org/10.1088/1367-2630/aa60ee.

[Tsa18] M. Tsang. "Subdiffraction incoherent optical imaging via spatial-mode demultiplexing: Semiclassical treatment". In: *Phys. Rev. A* 97 (2 Feb. 2018), p. 023830. DOI: 10.1103/PhysRevA.97.023830. URL: https://link.aps.org/doi/10.1103/PhysRevA.97.023830.

[Tse19] M. Tse et al. "Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy". In: *Phys. Rev. Lett.* 123 (23 Dec. 2019), p. 231107. DOI: 10.1103/PhysRevLett.123.231107. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.231107.

[Veg17] I. de Vega and D. Alonso. "Dynamics of non-Markovian open quantum systems". In: *Rev. Mod. Phys.* 89 (1 Jan. 2017), p. 015001. DOI: 10.1103/

RevModPhys.89.015001. URL: https://link.aps.org/doi/10.1103/RevModPhys.89.015001.

[Wan22] K. Wan and R. Lasenby. "Bounds on Adaptive Quantum Metrology under Markovian Noise". In: *Phys. Rev. Research* 4.3 (Aug. 2022), p. 033092. DOI: 10.1103/PhysRevResearch.4.033092. arXiv: 2108.11390.

[Wec21] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard. "Quantum Circuits with Classical Versus Quantum Control of Causal Order". In: *PRX Quantum* 2 (3 Aug. 2021), p. 030335. DOI: 10.1103/PRXQuantum.2.030335. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.030335.

[Whi22] G. White, F. Pollock, L. Hollenberg, K. Modi, and C. Hill. "Non-Markovian Quantum Process Tomography". In: *PRX Quantum* 3 (2 May 2022), p. 020344. DOI: 10.1103/PRXQuantum.3.020344. URL: https://link.aps.org/doi/10.1103/PRXQuantum.3.020344.

[Win92] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. "Spin squeezing and reduced quantum noise in spectroscopy". In: *Phys. Rev. A* 46 (11 Dec. 1992), R6797–R6800. DOI: 10.1103/PhysRevA.46.R6797. URL: https://link.aps.org/doi/10.1103/PhysRevA.46.R6797.

[Xia87] M. Xiao, L.-A. Wu, and H. J. Kimble. "Precision measurement beyond the shot-noise limit". In: *Phys. Rev. Lett.* 59 (3 July 1987), pp. 278–281. DOI: 10.1103/PhysRevLett.59.278. URL: https://link.aps.org/doi/10.1103/PhysRevLett.59.278.

[Yan19] Y. Yang. "Memory Effects in Quantum Metrology". In: *Phys. Rev. Lett.* 123 (11 Sept. 2019), p. 110501. DOI: 10.1103/PhysRevLett.123.110501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.110501.

[Yan24] X. Yang et al. "Control-enhanced non-Markovian quantum metrology". In: *Communications Physics* 7.1 (Aug. 2024). ISSN: 2399-3650. DOI: 10.1038/s42005-024-01758-8. URL: http://dx.doi.org/10.1038/s42005-024-01758-8.

[Yue76] H. P. Yuen. "Two-photon coherent states of the radiation field". In: *Phys. Rev. A* 13 (6 June 1976), pp. 2226–2243. DOI: 10.1103/PhysRevA.13.2226. URL: https://link.aps.org/doi/10.1103/PhysRevA.13.2226.

[Yur86a] B. Yurke. "Input States for Enhancement of Fermion Interferometer Sensitivity". In: *Phys. Rev. Lett.* 56 (15 Apr. 1986), pp. 1515–1517. DOI: 10.1103/PhysRevLett.56.1515. URL: https://link.aps.org/doi/10.1103/PhysRevLett.56.1515.

[Yur86b] B. Yurke, S. L. McCall, and J. R. Klauder. "SU(2) and SU(1,1) interferometers". In: *Phys. Rev. A* 33 (6 June 1986), pp. 4033–4054. DOI: 10.1103/PhysRevA.33.4033. URL: https://link.aps.org/doi/10.1103/PhysRevA.33.4033.

[Zha20] X. Zhao, Y. Yang, and G. Chiribella. "Quantum Metrology with Indefinite Causal Order". In: *Phys. Rev. Lett.* 124 (19 May 2020), p. 190503. DOI: 10.1103/PhysRevLett.124.190503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.190503.

[Zho17] Z. Zhou, Z. Lü, H. Zheng, and H.-S. Goan. "Quantum Zeno and anti-Zeno effects in open quantum systems". In: *Phys. Rev. A* 96 (3 Sept. 2017), p. 032101. DOI: 10.1103/PhysRevA.96.032101. URL: https://link.aps.org/doi/10.1103/PhysRevA.96.032101.

[Zho18] S. Zhou, M. Zhang, J. Preskill, and L. Jiang. "Achieving the Heisenberg Limit in Quantum Metrology Using Quantum Error Correction". In: *Nat. Commun.* 9.1 (Dec. 2018), p. 78. DOI: 10.1038/s41467-017-02510-3. arXiv: 1706.02445.

[Zho20] S. Zhou and L. Jiang. "Optimal approximate quantum error correction for quantum metrology". In: *Phys. Rev. Res.* 2 (1 Mar. 2020), p. 013235. DOI: 10.1103/PhysRevResearch.2.013235. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.013235.

[Zho21] S. Zhou and L. Jiang. "Asymptotic Theory of Quantum Channel Estimation". In: *PRX Quantum* 2 (1 Mar. 2021), p. 010343. DOI: 10.1103/PRXQuantum.2.010343. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.010343.

[Zho24] Z.-Y. Zhou, J.-T. Qiu, and D.-J. Zhang. "Strict hierarchy of optimal strategies for global estimations: Linking global estimations with local ones". In: *Phys. Rev. Res.* 6 (3 Aug. 2024), p. L032048. DOI: 10.1103/PhysRevResearch.6.L032048. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.6.L032048.

[Zur03] W. H. Zurek. "Decoherence, einselection, and the quantum origins of the classical". In: *Rev. Mod. Phys.* 75 (3 May 2003), pp. 715–775. DOI: 10.1103/RevModPhys.75.715. URL: https://link.aps.org/doi/10.1103/RevModPhys.75.715.