An outline of quantum interferometry

Jan Chwedeńczuk

Contents

Chapter 1

Basics of the estimation theory

1.1 Introduction

The goal of this manuscript is to introduce the basic notions of the estimation theory and discuss its applications to quantum interferometry. To begin, assume that some properties of a system (either classical or quantum) depend on a parameter which we will denote by θ. This parameter could stand, for instance, for a temperature T of a state at thermal equilibrium, a mass m or a position of a particle $\mathbf{r}$. For a quantum system, θ could also be a relative phase between two sub-systems, and the estimation of this parameter lies at the heart of quantum interferometry. Nevertheless, we now present some general results of the esitmation theory without actually specifying what θ could be.

The aim of the estimation procedure is to learn about the unknown parameter θ with the highest precission possible. Let us assume that the experimental sequence consists of m measurements of some quantity x. A central object is the *estimator*, which is some function which utilizes the results $x_1^{(1)} \ldots x_m^{(1)}$ to provide the value of the parameter, namely

$$\theta_{\text{est}}^{(1)} = \theta_{\text{est}}(x_1^{(1)} \ldots x_m^{(1)}). \tag{1.1}$$

The upper index $^{(1)}$ indicates that a first block of m outcomes was used to obtain $\theta_{\text{est}}^{(1)}$. Note that in order to tell what is the estimation error, one must repeat the sequence many times, to obtain many values of $\theta_{\text{est}}^{(i)}$. They differ from shot to shot because the outcomes of the measurements x_i inevitably fluctuate (either due to the experimental imperfections or – more fundamentally – quantum fluctuations), thus another set of $x_1^{(2)} \ldots x_m^{(2)}$ will provide a different value of $\theta_{\text{est}}^{(2)}$.

After M repetitions of the whole estimation sequence, one can plot a his-
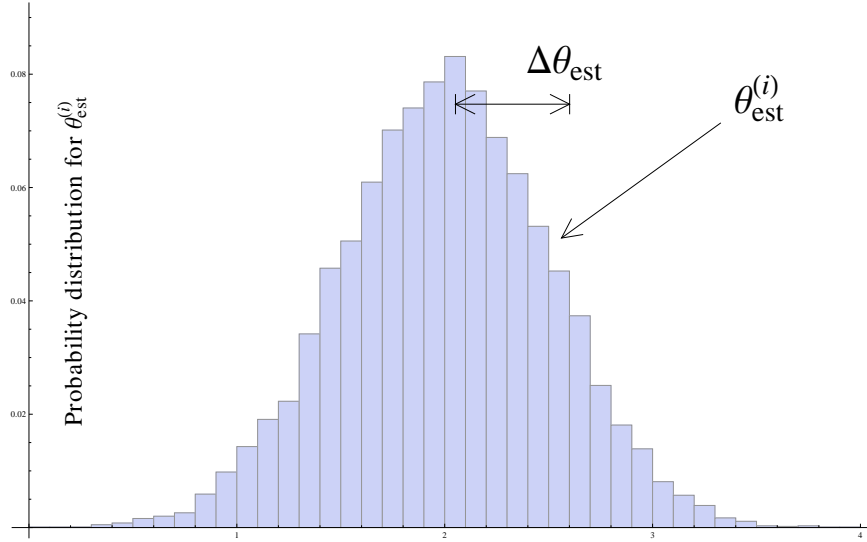
Figure 1.1: A histogram of $\theta_{\text{est}}^{(i)}$, which for large m tends to a Gaussian. The average of this histogram is our estimated value of the parameter, while its width is the "sensitivity" of the metrological devie.

togram of $\theta_{\text{est}}^{(i)}$, as shown in Fig. 1.1 [1] and say that the *estimated* parameter is the averege value of this distribution, namely

$$\langle \theta_{\text{est}} \rangle = \frac{1}{M} \sum_{k=1}^{M} \theta_{\text{est}}^{(k)}. \tag{1.2}$$

This averaging over the blocks can be formally replaced by the averaging over all the possible outcomes of the measured values of the parameters x_i. We introduce $P(x_1 \ldots x_m | \theta)$, which is a joinded probability distributions of these outcomes. It conditionally depends on the *true* value of the parameter θ. Clearly, if that would not be the case, no relation between $\langle \theta_{\text{est}} \rangle$ and θ could be established. Equation (1.2) can be thus written as

$$\langle \theta_{\text{est}} \rangle = \int dx_1 \ldots dx_m \ \theta_{\text{est}}(x_1 \ldots x_m) \, P(x_1 \ldots x_m | \theta). \tag{1.3}$$

Let us now mention two important properties of the estimators. One is called *consistency* which is defined by

$$\lim_{m \to \infty} \theta_{\text{est}}^{(i)}(x_1^{(i)} \ldots x_m^{(i)}) = \theta \quad \text{for all } i. \tag{1.4}$$

Another definition, which is crucial for us at the moment states that an estimator is called *unbiased* if its average value tends to the true value of the parameter,

---

[1] which by the way will tend to a Gaussian for large m – a consequence of the Central Limit Theorem discussed later

namely

$$\langle \theta_{\text{est}} \rangle = \theta. \tag{1.5}$$

From now on we will focus only on such unbiased estimators.

1.2 Cramer-Rao Lower Bound

We will now manipulate the Eq. (1.5) to derive the final expression for the sensitivity of the parameter estimation. First note, that it can be trivially rewritten as

$$\int dx_1 \ldots dx_m \; (\theta_{\text{est}} - \theta) \, P(x_1 \ldots x_m | \theta) = 0, \tag{1.6}$$

where we dropped the explicit dependence of θ_{est} on the measurement outcomes for simplicity. By taking the derivative of this equation with respect to θ we obtain that

$$\int d\vec{x} \; (\theta_{\text{est}} - \theta) \, \frac{\partial P(\vec{x}|\theta)}{\partial \theta} = 1, \tag{1.7}$$

with $\vec{x} = x_1 \ldots x_m$.

In the next step, a Cauchy-Schwartz inequality can be used, which says that if f and g are two functions of $\vec{x}$, then

$$\left| \int d\vec{x} \; f(\vec{x}) g(\vec{x}) \right|^2 \leqslant \int d\vec{x} \; \left| f(\vec{x}) \right|^2 \int d\vec{x} \; \left| g(\vec{x}) \right|^2. \tag{1.8}$$

By setting

$$f(\vec{x}) = \frac{1}{\sqrt{P(\vec{x}|\theta)}} \frac{\partial P(\vec{x}|\theta)}{\partial \theta} \quad \text{and} \quad g(x) = \sqrt{P(\vec{x}|\theta)}(\theta_{\text{est}} - \theta) \tag{1.9}$$

and by combining Eq. (5.43) and (1.8) we obtain that

$$\int d\vec{x} \; (\theta_{\text{est}} - \theta)^2 P(\vec{x}|\theta) \int d\vec{x} \; \frac{1}{P(\vec{x}|\theta)} \left(\frac{\partial P(\vec{x}|\theta)}{\partial \theta} \right)^2 \geqslant 1. \tag{1.10}$$

The first term in the left-hand-side is the variance of the estimator

$$\Delta^2 \theta_{\text{est}} \equiv \int d\vec{x} \; (\theta_{\text{est}} - \theta)^2 P(\vec{x}|\theta) \tag{1.11}$$

and is the central object in the estimation theory. In the remaining part of this manuscript we will focus solely on the properties of this variance and show how it can be minimalized. Note that $\Delta \theta_{\text{est}}$ is often called the estimation *precision* or the *sensitivity* of the estimation protocol. The second term in the left-hand-side is called the *Fisher information* which is denoted by F_m, namely

$$F_m \equiv \int d\vec{x} \; \frac{1}{P(\vec{x}|\theta)} \left(\frac{\partial P(\vec{x}|\theta)}{\partial \theta} \right)^2. \tag{1.12}$$

With the help of the above definitions, equation (1.10) can be written in a particularly simple form

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{F_m}}. \tag{1.13}$$

This inequality is called the Cramer-Rao Lower Bound [1, 2] and it says that the precision of *any* estimator which uses the results distributed with the probability $P(\vec{x}|\theta)$ is bounded by the square root of the inverse of the Fisher information. In a particular case, when $x_1 \ldots x_m$ are all the same single object measered in the m independent repetitions of the experiment (for instance this could the population imbalance between the two arms of an interferometer measured m times), the joint probability becomes a simple product

$$P(\vec{x}|\theta) = p(x_1|\theta) \ldots p(x_m|\theta). \tag{1.14}$$

By plugging this probability into Eq. (1.13) we obtain

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}} \frac{1}{\sqrt{F}} \tag{1.15}$$

where the Fisher information now reads

$$F = \int dx \ \frac{1}{p(x|\theta)} \left(\frac{\partial p(x|\theta)}{\partial \theta} \right)^2. \tag{1.16}$$

Note that according to the CRLB (5.10), the higher the value of the Fisher information, the less restrictive is the lower bound for the estimation precision.

Also, we can identify for which kind probability P, the bound is saturated. To this end we note that the Cauchy-Schwartz relation becomes an equality when $f(x) = \alpha g(x)$, where α is a constant. Using the definitions from Eq. (1.9), we obtain that

$$\frac{1}{\sqrt{P(\vec{x}|\theta)}} \frac{\partial P(\vec{x}|\theta)}{\partial \theta} = \alpha \sqrt{P(\vec{x}|\theta)} \big(\theta_{\text{est}} - \theta \big). \tag{1.17}$$

This is a simple differential equation, which has a solution

$$P(\vec{x}|\theta) \propto e^{-\frac{1}{2}|\alpha|(\theta - \theta_{\text{est}})^2}, \tag{1.18}$$

where the proportionality sign stands for the normalization. In other words, when the probability is gaussian in the estimator, it automatically saturates the CRLB. But what happens when P is not gaussian? Can the CRLB be nevertheless saturated?

1.3   Maximum-Likelihood Estimator

The answer is positive and it turns out that the correct choice is the Maximum-Likelihood Estimator (MLE). The construction of this estimator is following.

Let us assume again that a quantity x is measured m times and that the probability $p(x|\theta)$ is known [2]. With the measurement outcomes $x_1 \ldots x_m$ at hand, a likelihood function is defined in a following way

$$\mathcal{L}(\varphi) = \ln\left(\prod_{i=1}^{m} p(x_i|\varphi)\right) = \sum_{i=1}^{m} \ln\left(p(x_i|\varphi)\right). \tag{1.19}$$

It is a function of a single variable φ, with the other variables fixed by the measurement outcomes. The MLE is now defined as the value of φ, at which $\mathcal{L}$ reaches its maximal value,

$$\theta_{\text{MLE}}^{(m)}: \qquad \left.\frac{\partial \mathcal{L}(\varphi)}{\partial \varphi}\right|_{\varphi=\theta_{\text{MLE}}^{(m)}} = 0. \tag{1.20}$$

The uppper index m is to underline that the MLE estimator is obtained from a series of m experiments. We will now show that such chosen estimator is consistent, unbiased and its variance saturates the CRLB.

• **consistency of the MLE** To demonstate that the MLE is consistent, let us calculate the following difference

$$\mathcal{L}(\varphi) - \mathcal{L}(\theta) = \sum_{i=1}^{m} \ln\left(\frac{p(x_i|\varphi)}{p(x_i|\theta)}\right) \xrightarrow{m\to\infty} m \int dx\, p(x|\theta) \ln\left(\frac{p(x|\varphi)}{p(x|\theta)}\right). \tag{1.21}$$

This last expression can be easily bounded from above using $\ln(y) \leqslant y - 1$ to give

$$m \int dx\, p(x|\theta) \ln\left(\frac{p(x|\varphi)}{p(x|\theta)}\right) \leqslant m \int dx\, p(x|\theta) \left(\frac{p(x|\varphi)}{p(x|\theta)} - 1\right) = 0. \tag{1.22}$$

This inequality says that the maximal value of the left-hand-side is zero, and by definition it is reached when $\varphi = \theta_{\text{MLE}}^{(m)}$. But in this case $\mathcal{L}(\theta_{\text{MLE}}^{(m)}) = \mathcal{L}(\theta)$, which proves consistency, i.e.

$$\lim_{m\to\infty} \theta_{\text{MLE}}^{(m)} = \theta. \quad \square \tag{1.23}$$

• **MLE is unbiased** In the next step we demonstate, that the MLE is unbiased. To this end, we calculate the derivative of the likelihood function and expand it in φ around the true value of the parameter,

$$\mathcal{L}'(\varphi) = \mathcal{L}'(\theta) + \mathcal{L}''(\theta)(\varphi - \theta) + \ldots. \tag{1.24}$$

We now set $\varphi = \theta_{\text{MLE}}^{(m)}$, which gives the R.H.S. equal to zero (by definition), and assuming $m \gg 1$ we make use of the consistency of the MLE. Namely, for large

---

[2]In true experiments it is usually very difficult to find out what this probability actually looks like. One does it in the *calibration* stage, where for a given *known* value of θ the probability is measured, and then θ is changed to another value and the whole procedure is repeated. This way, a two-dimensional function is sampled $p(x|\theta)$. Clearly, to calibrate the device requires much experimental effort.

m, $\theta_{\mathrm{MLE}}^{(m)}$ is close to θ, so the above expansion can be safely truncated at the first order, giving

$$\theta_{\mathrm{MLE}}^{(m)} - \theta \simeq -\frac{\mathcal{L}'(\theta)}{\mathcal{L}''(\theta)}. \tag{1.25}$$

We now concentrate for a moment on the denominator of this equation and notice that according to the definition of the likelihood function from Eq. (1.19), its second derivative reads

$$\mathcal{L}''(\theta) = \sum_{i=1}^{m} \left[-\frac{1}{p^2(x_i|\theta)} \left(\frac{\partial p(x_i|\varphi)}{\partial \varphi} \Big|_{\varphi=\theta} \right)^2 + \frac{1}{p(x_i|\theta)} \frac{\partial^2 p(x_i|\varphi)}{\partial \varphi^2} \Big|_{\varphi=\theta} \right] \tag{1.26}$$

For large m, summation over individual events can be replaced by the averaging with the probability $p(x|\theta)$ and we obtain

$$\mathcal{L}''(\theta) \xrightarrow{m\to\infty} m \int dx \left[-\frac{1}{p(x|\theta)} \left(\frac{\partial p(x|\varphi)}{\partial \varphi} \Big|_{\varphi=\theta} \right)^2 + \frac{\partial^2 p(x|\varphi)}{\partial \varphi^2} \Big|_{\varphi=\theta} \right] = -m\,F, \tag{1.27}$$

since the second term vanishes due to the normalization of the probability. We now plug this expression into Eq. (1.25) and obtain

$$\theta_{\mathrm{MLE}}^{(m)} - \theta \simeq \frac{1}{m} \sum_{i=1}^{m} \left[\frac{1}{F} \frac{1}{p(x_i|\theta)} \frac{\partial p(x_i|\varphi)}{\partial \varphi} \Big|_{\varphi=\theta} \right]. \tag{1.28}$$

Note that the object in the square parenthesis, which we will denote by s_i is a random variable, since the outcomes of the measurements are random by itself. Therefore, the upper equation which can be written as

$$\theta_{\mathrm{MLE}}^{(m)} - \theta \simeq \frac{1}{m} \sum_{i=1}^{m} s_i \tag{1.29}$$

tells that the L.H.S. is an average value of m random variables. At this point, we can invoke the Cental Limit Theorem (CLT), which says that if an object is an average value of m random variables, then when $m \to \infty$, this object is governed by the Gaussian (normal) distribution, with a mean equal to the mean of the random variable, $\langle s \rangle$, and the fluctuations proportional to the fluctuations of the random variable and equal to $\frac{1}{m}\Delta s$. In other words, we can now easily calculate both the average value and the variance of the difference between the MLE and the true value of the estimator. According to the CLT, we thus have that

$$\left\langle \theta_{\mathrm{MLE}}^{(m)} - \theta \right\rangle \xrightarrow{m\to\infty} \int dx\, p(x|\theta) \left[\frac{1}{F} \frac{1}{p(x|\theta)} \frac{\partial p(x|\varphi)}{\partial \varphi} \Big|_{\varphi=\theta} \right] = \tag{1.30a}$$

$$= \frac{1}{F} \int dx \frac{\partial p(x|\varphi)}{\partial \varphi} \Big|_{\varphi=\theta} = 0. \quad \square \tag{1.30b}$$

Therefore, according to the definition from Eq. (1.5), the MLE estimator is unbiased.

• **Variance of the MLE saturates the CRLB.** According to the CRLB we obtain that

$$\Delta^2 \theta_{\text{MLE}}^{(m)} \xrightarrow{m \to \infty} \frac{1}{m} \int dx \, p(x|\theta) \left[\frac{1}{F} \frac{1}{p(x|\theta)} \frac{\partial p(x|\varphi)}{\partial \varphi}\bigg|_{\varphi=\theta} \right]^2 \tag{1.31a}$$

$$= \frac{1}{m} \frac{1}{F^2} \int dx \, \frac{1}{p(x|\theta)} \left(\frac{\partial p(x|\varphi)}{\partial \varphi}\bigg|_{\varphi=\theta} \right)^2 = \frac{1}{m} \frac{1}{F}. \quad \square \tag{1.31b}$$

This shows that indeed the MLE saturates the CRLB, according to Eq. (5.10).

Chapter 2

De revolutionibus

In this chapter we wish to introduce the algebra of the two-mode angular momentum operators. As will become evident in the following chapters, these operators are particularly useful in the description of two-mode interferometers. In order to gain some deeper insight into the physics behind these operators and their eigen-states, we will first concentrate on the rotations in the three-dimensional space. Further on, we will show the link between the generators of these transformations and the two-mode operators.

2.1 Rotations in 3D space

2.1.1 SO(3) group

Rotation in three-dimensional space transforms a vector $\vec{r}$ into $\vec{r}'$ as follows

$$r'_j = \sum_i R_{ij}(\vec{\varphi})r_j. \tag{2.1}$$

The vector $\vec{\varphi}$ determines the direction of the rotation, while $\hat{R}$ is the rotation matrix. One important property of the rotiations is that they conserve the scalar product. Formally, this can be written as

$$\vec{a}'\vec{b}' = \sum_i a'_i b'_i = \sum_{ijk} R_{ij}(\vec{\varphi})R_{ik}(\vec{\varphi})a_j b_k = \sum_j a_j b_j. \tag{2.2}$$

This is true, when $\hat{R}^T\hat{R} = \hat{R}\hat{R}^T = \hat{\mathbb{1}}$. Since $\det\hat{A}\hat{B} = \det B \det A$ and $\det\hat{A}^T = \det\hat{A}$, then we imediatelly obtain that $\det\hat{R} = \pm 1$. It can be shown that the minus sign is obtained for the transformations, which are not only rotations, but also inversions of the coordinate system. Pure rotations require $\det\hat{R} = 1$.

The above considerations allow us to define the set of rotation matrices in the 3D space, named as SO(3), and having the following properties

$$\text{SO(3)} = \{\text{real } 3 \times 3 \text{ matrices } \hat{R}; \ \hat{R}^T\hat{R} = \hat{R}\hat{R}^T = \hat{\mathbb{1}}; \ \det\hat{R} = 1\}. \tag{2.3}$$

It turns out that this set has the properties of the group $\mathcal{G}$, namely

1. For any $\hat{R}_1 \in \mathcal{G}$ and $\hat{R}_2 \in \mathcal{G}$, also $\hat{R}_1 \cdot \hat{R}_2 \in \mathcal{G}$.

2. There exists $\hat{R}_e \in \mathcal{G}$, for which $\hat{R} \cdot \hat{R}_e = \hat{R}_e \cdot \hat{R} = \hat{R}$ for all $\hat{R} \in G$

3. For all $\hat{R} \in \mathcal{G}$, there exists $\hat{R}^{-1}$, for which $\hat{R} \cdot \hat{R}^{-1} = \hat{R}^{-1} \cdot \hat{R} = \hat{R}_e$

4. An associative law holds for a product of three elements, namely $\hat{R}_1 \cdot (\hat{R}_2 \cdot \hat{R}_3) = (\hat{R}_1 \cdot \hat{R}_2) \cdot \hat{R}_3$

It can be easily demonstrated that all these properties hold for the rotation matrices in 3D space.

2.1.2   Generators of rotations

We now wish to express the rotation matrices $\hat{R}(\vec{\varphi})$ in the most convenient and compact way. Our claim is that these matrices can be represented in a following way

$$\hat{R}(\vec{\varphi}) = e^{-\frac{i}{\hbar}\vec{\varphi}\vec{J}}. \tag{2.4}$$

Let us assume for a moment, that indeed this representation is valid. We will now demomnstrate an property of the operators

$$\hat{L}_k = -\frac{i}{\hbar}\hat{J}_k. \tag{2.5}$$

Namely, we will show that these operator must be antisymmetric. To this end, note that intuitively it should hold that

$$\hat{R}^{-1}(\varphi_k) = \hat{R}(-\varphi_k) = e^{-\varphi_k \hat{L}_k}. \tag{2.6}$$

Using a simple relation for operator functions $f(\hat{O})^T = f(\hat{O}^T)$, and since $\hat{R}^{-1} = \hat{R}^T$, we obtain that

$$\hat{R}^{-1}(\varphi_k) = \hat{R}^T(\varphi_k) = e^{\varphi_k \hat{L}_k^T}. \tag{2.7}$$

Combining this result with Eq. (2.6), we arrive at the antisymmetry condition

$$\hat{L}_k^T = -\hat{L}_k. \tag{2.8}$$

Therefore, these matrices can be written in a quite general form

$$\hat{L}_k = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}. \tag{2.9}$$

Since there are only three independent numbers, which deterime the matrices $\hat{L}_k$, one can expect that there should be three matrices generating the rotations.

For instance, since from geometric arguments we know that the rotation matrix around the z axis has a form

$$\hat{R}(\varphi_z) = \begin{pmatrix} \cos\varphi_z & -\sin\varphi_z & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.10}$$

then using Eq. (2.4) we obtain that

$$\hat{L}_z = \lim_{\varphi_z \to 0} \frac{1}{\varphi_z} \left(\hat{R}(\varphi_z) - \hat{\mathbb{1}} \right) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.11}$$

In a similar way, one can show that

$$\hat{L}_x = \lim_{\varphi_x \to 0} \frac{1}{\varphi_x} \left(\hat{R}(\varphi_x) - \hat{\mathbb{1}} \right) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.12}$$

and

$$\hat{L}_y = \lim_{\varphi_y \to 0} \frac{1}{\varphi_y} \left(\hat{R}(\varphi_y) - \hat{\mathbb{1}} \right) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{2.13}$$

Now that the explicit form the rotation generators is known, one can easily demonstrate that

$$\left[\hat{L}_j, \hat{L}_k \right] = \varepsilon_{jkl} \hat{L}_l. \tag{2.14}$$

Groups, which elemnts can be written is a form (2.4) and their generators form a closed algebra (2.14) are called Lie groups, while their elements form the Lie algebra. Note that one can equally well use the $\hat{J}_k$ generators to arrive at

$$\left[\hat{J}_j, \hat{J}_k \right] = i\hbar\varepsilon_{jkl} \hat{J}_l. \tag{2.15}$$

2.1.3 Spatial form of the generators

We would like now to see how the rotations appear in quantum mechanics. Consider a single-particle wave-function $\psi(\mathbf{r})$. We can, rather then considering a rotation matrix $\hat{R}(\vec{\varphi})$, which maps $\mathbb{R}^3 \to \mathbb{R}^3$, introduce a rotation operator $\hat{\mathcal{R}}(\vec{\varphi})$, which maps the set of complex valued functions over $\mathbb{R}^3$ onto itself, i.e. $\mathbb{C}_\infty(3) \to \mathbb{C}_\infty(3)$. This rotation operators should be defined as follows

$$\hat{\mathcal{R}}(\vec{\varphi})\psi(\mathbf{r}) \equiv \psi(\hat{R}^{-1}(\vec{\varphi})\mathbf{r}). \tag{2.16}$$

It can be expected, that in analogy to Eq. (2.4) also the operators $\hat{\mathcal{R}}(\vec{\varphi})$ should form a Lie group, namely they should be representable in a form

$$\hat{\mathcal{R}}(\vec{\varphi}) = e^{-\frac{i}{\hbar}\vec{\varphi}\hat{\vec{\mathcal{J}}}}. \tag{2.17}$$

The generators can be determined in a similar way as in the case of generators $\hat{J}_k$. For instance

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_z = \lim_{\varphi_z \to 0} \frac{1}{\varphi_z}\left(\hat{\mathcal{R}}(\varphi_z) - \hat{\mathbb{1}}\right). \tag{2.18}$$

We can now act with the above equation onto the wave-function $\psi(\mathbf{r})$ to obtain

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_z\psi(\mathbf{r}) = \lim_{\varphi_z \to 0} \frac{1}{\varphi_z}\left(\hat{\mathcal{R}}(\varphi_z) - \hat{\mathbb{1}}\right)\psi(\mathbf{r}) = \lim_{\varphi_z \to 0} \frac{1}{\varphi_z}\left(\psi(\hat{R}(-\varphi_z)\mathbf{r}) - \psi(\mathbf{r})\right). \tag{2.19}$$

Using Eq. (2.11) we obtain that for small angles

$$\hat{R}(-\varphi_z)\mathbf{r} = \begin{pmatrix} 1 & \varphi_z & 0 \\ -\varphi_z & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + \varphi_z y \\ y - \varphi_z x \\ z \end{pmatrix}. \tag{2.20}$$

Therefore, using the Taylor expansion we end up with an expression

$$\psi(\hat{R}(-\varphi_z)\mathbf{r}) \simeq \psi(\mathbf{r}) + y\partial_x\psi(\mathbf{r}) - x\partial_y\psi(\mathbf{r}). \tag{2.21}$$

Plugging this expression into Eq. (2.19) we get

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_z = y\partial_x - x\partial_y. \tag{2.22}$$

Similarly for x and y we have that

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_x = z\partial_y - y\partial_z \tag{2.23a}$$

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_y = x\partial_z - z\partial_x. \tag{2.23b}$$

Cleraly, the generators of rotations are the angular momentum operators, or in other words

$$\vec{\hat{\mathcal{J}}} = \mathbf{r} \times \mathbf{p} \equiv \mathbf{r} \times \frac{\hbar}{i}\vec{\nabla}. \tag{2.24}$$

It can be easily checked, that the elements of the angular momentum vector form the Lie algebra

$$\left[\hat{\mathcal{J}}_j, \hat{\mathcal{J}}_k\right] = i\hbar\varepsilon_{jkl}\hat{\mathcal{J}}_l. \tag{2.25}$$

2.1.4   Eigen-states of $\hat{\mathcal{J}}_z$

We now briefly recall the basic properties of the most commonly used basis for the angular momentum operators. Note, that since the three components do not commute, as underlined by the relation (2.25), then there is no signle eigen-basis common for all of them. Nevertheless, one can construct the total angular momentum operator as follows

$$\hat{J}^2 = \hat{\mathcal{J}}_x^2 + \hat{\mathcal{J}}_y^2 + \hat{\mathcal{J}}_z^2. \tag{2.26}$$

It can be trivially demonstrated, that this operator commutes with all three components, i.e. $[\hat{\mathcal{J}}^2, \hat{\mathcal{J}}_i] = 0$ for $i = x, y, z$. Therefore, one can seek for the eigen-states of $\hat{\mathcal{J}}^2$ and one of the three components – traditionally the z one.

To introduce this basis it is convenient to change the basis and express the angular momentum operators in the spherical coordinates. A simple change of variables gives

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_x = \sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi} \tag{2.27a}$$

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_y = -\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi} \tag{2.27b}$$

$$-\frac{i}{\hbar}\hat{\mathcal{J}}_z = -\frac{\partial}{\partial\phi}. \tag{2.27c}$$

Using this representation and the commutation relations (2.25), a set of eigen-vectors of $\hat{\mathcal{J}}_z$ and $\hat{\mathcal{J}}^2$ can be found. These functions of ϕ and θ are labelled by two quantum numbers l and m, where $m \in [-l, -l+1 \ldots l-1, l]$ and have the following properties

$$\hat{\mathcal{J}}^2 Y_{lm}(\phi, \theta) = \hbar^2 l(l+1) Y_{lm}(\phi, \theta) \tag{2.28a}$$
$$\hat{\mathcal{J}}_z Y_{lm}(\phi, \theta) = \hbar m Y_{lm}(\phi, \theta). \tag{2.28b}$$

They are orthonormal, i.e.

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \, \sin\theta \, Y_{lm}^*(\phi, \theta) Y_{l'm'}(\phi, \theta) = \delta_{ll'}\delta_{mm'}. \tag{2.29}$$

They form a complete basis, thus any function of θ and ϕ can be represented as a combination of spherical harmonics. Another important property is the recursion relation, which reads

$$\hat{\mathcal{J}}_+ Y_{lm}(\phi, \theta) = \hbar\sqrt{(l+m+1)(l-m)} Y_{lm+1}(\phi, \theta) \tag{2.30a}$$
$$\hat{\mathcal{J}}_- Y_{lm}(\phi, \theta) = \hbar\sqrt{(l+m+1)(l+m)} Y_{lm-1}(\phi, \theta), \tag{2.30b}$$

where the rising/lowering operators are defined as $\hat{\mathcal{J}}_\pm = \hat{\mathcal{J}}_x \pm i\hat{\mathcal{J}}_y$. Finally, the functional form of the spherical harmonics is

$$Y_{lm}(\phi, \theta) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\frac{(-1)^m}{2^l l!}\sin^m(\theta)\left(\frac{\partial}{\partial\theta}\right)^m P_l(\cos\theta)e^{im\phi} \tag{2.31}$$

for $m \geq 0$ and $Y_{l-m}(\phi, \theta) = (-1)^m Y_{lm}^*(\phi, \theta)$. Here $P_l(x)$ are the Legendre polynomials defined by the recursion relation

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_{l+1}(x) = \frac{1}{l+1}\left((2l+1)xP_l(x) - lP_{l-1}(x)\right). \tag{2.32}$$

2.1.5   Irreducible representations

Let us now stop for a moment and write down how the rotation operators act on
any function $f(\phi, \theta) \in \mathbb{C}_\infty(S_2)$, i.e. complex functions over a two-dimensional
unit sphere. We have introduce the image of the rotation operator

$$\tilde{f}(\phi, \theta) = e^{-\frac{i}{\hbar}\vec{\varphi}\vec{\mathcal{J}}} f(\phi, \theta). \tag{2.33}$$

We now use the basis of the spherical harmonics to expand the image function
as follows

$$\tilde{f}(\phi, \theta) = \sum_{lm} C_{lm} Y_{lm}(\phi, \theta), \tag{2.34}$$

where the expansion coefficient simply reads

$$C_{lm} = \int d\Omega' Y_{lm}^*(\phi', \theta') e^{-\frac{i}{\hbar}\vec{\varphi}\vec{\mathcal{J}}} f(\phi', \theta'). \tag{2.35}$$

Clearly, the rotations can be fully characterized by the following rotation matrix

$$\mathcal{D}_{l'm';lm} = \int d\Omega Y_{l'm'}^*(\phi, \theta) e^{-\frac{i}{\hbar}\vec{\varphi}\vec{\mathcal{J}}} Y_{lm}(\phi, \theta) \tag{2.36}$$

or in other words, since the spherical harmonics span the whole space $\mathbb{C}_\infty(S_2)$

$$\mathbb{C}_\infty(S_2) = \{Y_{lm}, l = 0, 1 \ldots \infty, \ m = -l \ldots l\}, \tag{2.37}$$

then it is sufficient to tell how these functions transform under rotations. If we
now define a set $\mathbb{X}_l$ as

$$\mathbb{X}_l = \{Y_{lm}, \ m = -l \ldots l\} \quad \text{with fixed } l, \tag{2.38}$$

then the whole space can be represented as a sum of these sets

$$\mathbb{C}_\infty(S_2) = \bigcup_{l=1}^{\infty} \mathbb{X}_l. \tag{2.39}$$

This choice of $\mathbb{X}_l$ is not random. Since rotations do not change the l quantum
number of spherical harmonics, then the set $\mathbb{X}_l$ with some fixed l is invariant
under rotations. In other words, $\mathbb{X}_l$'s with different l's do not communicate
under rotations. This representation of the rotation operations is called the
irreducible representation, because the sets $\mathbb{X}_l$ form the smallest possible sub-
spaces of $\mathbb{C}_\infty(S_2)$.

2.1.6   Wigner rotation matrix

We now make a stup further, and determine the matrix elemnts of $\mathcal{D}$. To this
end, we introduce the Euler parametrization of rotations. What we have used
so far is the rotation operator in a form

$$\hat{\mathcal{R}}(\vec{\varphi}) = e^{-\frac{i}{\hbar}\vec{\varphi}\vec{\mathcal{J}}}, \tag{2.40}$$

so – in general – a simultaneous rotation around all three carthesian axes. It turns out however, that any such rotation can be expressed in another form, namely

$$\hat{\mathcal{R}}(\vec{\varphi}) = e^{-\frac{i}{\hbar}\gamma\hat{\mathcal{J}}_z''} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'} e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z}. \tag{2.41}$$

Here, the angles α, β and γ are uniquely determined by $\vec{\varphi}$. So, the first rotation is around the z-axis of the initial unprimed axis. The second one, is around the primed y-axis, which is rotated with respect to the initial one by the first z-rotation. Finally, the last rotation is around the z-axis of the coordinate system rotated twice by the previous rotations. The representation (2.41) migh seem an unnecessary complication with respect to (2.40), however we will now demonstrate that it actually allows to write down the $\mathcal{D}$ matrix in a particularly simple form. We will now demonstrate, that Eq. (2.41) can be re-written in terms of three rotations around the axes of the initial *unprimed* system. To this end, we note that

$$e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'} = e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} e^{\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z}. \tag{2.42}$$

Thus the first two rotations actually can be written as

$$e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'} e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} = e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} e^{\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} = e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y}. \tag{2.43}$$

Now, the last rotation is similarly

$$e^{-\frac{i}{\hbar}\gamma\hat{\mathcal{J}}_z''} = e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'} e^{-\frac{i}{\hbar}\gamma\hat{\mathcal{J}}_z} e^{\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'} e^{\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z}. \tag{2.44}$$

We can now plugg in the expression for $e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y'}$ from Eq. (2.42) and finally we obtain that

$$\hat{\mathcal{R}}(\vec{\varphi}) = e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} e^{-\frac{i}{\hbar}\gamma\hat{\mathcal{J}}_z}. \tag{2.45}$$

So clearly to express the rotation in the unprimed coordinate system, it is just necessary to invert the rotiation order and remove all the primes. Now, we insert the above result into the definition of the $\mathcal{D}$ matrix from Eq. (2.36) and obtain that

$$\mathcal{D}_{l'm';lm} = \int d\Omega Y_{l'm'}^*(\phi,\theta) e^{-\frac{i}{\hbar}\alpha\hat{\mathcal{J}}_z} e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} e^{-\frac{i}{\hbar}\gamma\hat{\mathcal{J}}_z} Y_{lm}(\phi,\theta). \tag{2.46}$$

We now know how the rotations around the z axis act on the spherical harmonics – they produce the m-dependent phase, so finally we obtain that

$$\mathcal{D}_{l'm';lm} = e^{-i\alpha m'} \int d\Omega Y_{l'm'}^*(\phi,\theta) e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} Y_{lm}(\phi,\theta) e^{-i\gamma m} \delta_{ll'}. \tag{2.47}$$

So to determine the full rotation properties, it is sufficient to find the matrix elemnt of the rotation around y-axis calculated in the basis of the spherical harmonics. We will provide the final expression for that in the following section. At this stage, we only introduce the *Wigner matrix*, which is defined as

$$d_{mm'}^{(l)}(\beta) = \int d\Omega Y_{l'm'}^*(\phi,\theta) e^{-\frac{i}{\hbar}\beta\hat{\mathcal{J}}_y} Y_{lm}(\phi,\theta). \tag{2.48}$$

2.2   Two-mode angular momentum operators

So what does this whole story have to do with quantum interferometry? We have been discussing here the general properties of the angular momentum operators and rotations in quantum mechanics. Now it is the right moment to demonstrate the link between the rotation operators and two-mode interferometry.

In general, any two mode pure quantum state consisting N indistinguishable bosons can be written as

$$|\psi\rangle = \sum_{n=0}^{N} C_n |n, N - n\rangle. \tag{2.49}$$

Here, the ket $|n, N - n\rangle$ denotes a state, where n particles occupy the first mode a and $N - n$ occupy the mode b. The amplitudes C_n are normalized, i.e. $\sum_{n=0}^{N} |C_n|^2 = 1$. Note that any such ket can be generated from the vacuum, by acting n times with the operator $\hat{a}^\dagger$, which creates a particle in the mode a and $N - n$ times with $\hat{b}^\dagger$, which creates a particle in the mode b, namely

$$|n, N - n\rangle = \frac{1}{\sqrt{n!(N-n)!}} \left(\hat{a}^\dagger\right)^n \left(\hat{b}^\dagger\right)^{N-n} |0, 0\rangle. \tag{2.50}$$

To demonstrate a close relation between the two-mode algebra and the theory of angular momentum, let us construct a family of "angular momentum" operators from the creation/anihilation operators of the two modes. The definition is following

$$\hat{J}_x = \frac{1}{2} \left(\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}\right) \tag{2.51a}$$

$$\hat{J}_y = \frac{1}{2i} \left(\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a}\right) \tag{2.51b}$$

$$\hat{J}_z = \frac{1}{2} \left(\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}\right). \tag{2.51c}$$

These three hemritian operators form a closed Lie algebra, because their commutation relation is

$$\left[\hat{J}_j, \hat{J}_k\right] = i\varepsilon_{jkl} \hat{L}_l. \tag{2.52}$$

Now, due to the aforementioned algebraic properties, one can again construct a common eigen-basis for the $\hat{J}_z$ and the total angular momentum operator

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2. \tag{2.53}$$

So let us take a ket $|n, N - n\rangle$ and act on it with the operators $\hat{J}_z$ and $\hat{J}^2$. What do we get is following

$$\hat{J}_z |n, N - n\rangle = \left(n - \frac{N}{2}\right) |n, N - n\rangle \ \text{ and } \ \hat{J}^2 |n, N - n\rangle = \frac{N}{2} \left(\frac{N}{2} + 1\right) |n, N - n\rangle. \tag{2.54}$$

So these kets form indeed the eigen-basis of the angular momentum operators. From the eigen-value of the $\hat{J}^2$ operator we can deduce the value of l, while from the eigen-value of $\hat{J}_z$ – about the m. So, the ket with n particles in one mode and $N - n$ in the other can be alternatively written as

$$|n, N - n\rangle \rightarrow \left|l = \frac{N}{2}, m = n - \frac{N}{2}\right\rangle. \tag{2.55}$$

Obviously, $m \in [-\frac{N}{2}, -\frac{N}{2}+1 \ldots \frac{N}{2}-1, \frac{N}{2}]$. Therefore, it is reasonable to say that N indistinguishable bosons in two well can be regarded as a single quasi-particle with the angular momentum equal to $\frac{N}{2}$.

To summarize, the angular momentum operators (3.31) form the same algebra as the operators $\hat{\mathcal{J}}_k$. The analogy also concerns the eigen-basis. The kets $|l, m\rangle$ are equivalent to the spherical harmonics $Y_{lm}(\phi, \theta)$. Consequently, the Wigner rotation matrix calculated can now be written as

$$d^{(l)}_{mm'}(\beta) = \langle l, m|e^{-i\beta \hat{J}_y}|l, m'\rangle. \tag{2.56}$$

It can be now shown (we skip the proof here), that the elemnts of this matrix have a following analytical form

$$
\begin{aligned}
d^{(l)}_{mm'}(\beta) &= \sqrt{\frac{m!(2l-m)!}{m'!(2l-m')!}} \left[\sin\left(\frac{\theta}{2}\right)\right]^{m-m'} \left[\cos\left(\frac{\theta}{2}\right)\right]^{m+m'-N} \times \\
&\quad \times P^{m-m',m+m'-2l}_{2l-m}(\cos\theta),
\end{aligned} \tag{2.57}
$$

where $P^{\alpha,\beta}_\mu(x)$ is the Jacobi polynomial.

Chapter 3

Quantum Fisher Information

3.1 Derivation of the Quantum Fisher information

Our discussion so far was quite general. We have derived the Cramer-Rao Lower Bound, which provides the lower limit for the precision of the estimation of the parameter θ from a series of m measurements of a physical quantity x performed on some system. The only important object, which entered the final expression for the Fisher information, is the conditional probability for obtaining the measurement outcome x given the value of the parameter θ.

From now on we will focus on such case, when the information about the parameter θ is *imprinted* on a *quantum* system through some specific unitary transformation. Although our choice might seem very restrictive, we will argue later on that most interferometric transformations can be described in this manner. Consider an "initial" quantum state $\hat{\varrho}_{\text{in}}$, which undergoes a following unitary transformation

$$\hat{\varrho}_{\text{in}} \rightarrow \hat{\varrho}(\theta) = e^{-i\theta\hat{h}}\hat{\varrho}_{\text{in}}e^{i\theta\hat{h}}. \tag{3.1}$$

Here, $\hat{h}$ is a hermitian generator of the evolution.

The goal, as in the general scheme outlined in the previous chapter is to perform a series of measurements on the *output* state $\hat{\varrho}(\theta)$ in order to deduce the value of the parameter. According to the CRLB from Eq. (5.10), the precision of the parameter estimation is bounded by

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}}\frac{1}{\sqrt{F}}, \tag{3.2}$$

where F is the Fisher information. As we already know, what enters the Fisher information is the conditional probability of obtaining the outcome x given θ.

Quantum mechanics provides the most general expression for this probability, in terms of the so-called Positive-Operator Valued Measures (POVMs) – a set of non-negative operators $\hat{E}_x$ which add up to unity, i.e.

$$\int dx\, \hat{E}_x = \hat{\mathbb{1}}. \tag{3.3}$$

These operators form the widest possible family of measurement operations allowed by quantum mechanics – a step behind the commonly used projection operators. With help of these operators, the conditional probability necessary for the evaluation of the Fisher information can be written as

$$p(x|\theta) = \mathrm{Tr}\left[\hat{E}_x \hat{\varrho}(\theta)\right] = \mathrm{Tr}\left[\hat{E}_x e^{-i\theta\hat{h}} \hat{\varrho}_{\mathrm{in}} e^{i\theta\hat{h}}\right]. \tag{3.4}$$

It is particularly instructive to give a closer look to the conditional probability expressed in this last form. It is now clear that there are three key ingredients which determine the conditional probability and in consequence the value of the Fisher information. These are the input state $\hat{\varrho}_{\mathrm{in}}$, the type of measurement chosen to infer the value of the parameter, here represented by the operator $\hat{E}_x$ and finally $\hat{h}$ – the generator of the interferometric transformation.

Usually, the interferometric sequence is determined by the experimental setup, which means that the form of $\hat{h}$ is fixed. Nevertheless, it is interesting to ask what is the maximal value of the Fisher information – optimized over all possible input states and measurements performed.

In the first step we calculate what is known as the Quantum Fisher Information (QFI) – the value of F for a given input state maximized over all possible measurements. In other words, our aim is to calculate the maximum of the expression

$$F = \int dx\, \frac{1}{\mathrm{Tr}\left[\hat{E}_x \hat{\varrho}(\theta)\right]} \left(\partial_\theta \mathrm{Tr}\left[\hat{E}_x \hat{\varrho}(\theta)\right]\right)^2 \tag{3.5}$$

with respect to all $\hat{E}_x$. The proof begins with an introduction of a superoperator

$$\hat{\mathcal{R}}_{\hat{\varrho}}(\hat{O}) \equiv \frac{1}{2}\left(\hat{\varrho}\hat{O} + \hat{O}\hat{\varrho}\right) = \sum_{j,k} \frac{1}{2}\left(p_j + p_k\right) O_{jk}|j\rangle\langle k|. \tag{3.6}$$

The last form is obtained using the eigen-basis of the density matrix operator, i.e. $\hat{\varrho} = \sum_j p_j |j\rangle\langle j|$. Naturally, as can be quickly verified, the inverse of this operator is

$$\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{O}) = \sum_{j,k} 2\frac{O_{jk}}{p_j + p_k}|j\rangle\langle k|. \tag{3.7}$$

This last expression is especially useful, because it allows to write down a trace of a product of two hermitian operators $\hat{A}\cdot\hat{B}$ in a form, which will be particularly useful to us in the next few steps, i.e.

$$\mathrm{Tr}\left[\hat{A}\cdot\hat{B}\right] = \mathrm{ReTr}\left[\varrho \cdot \hat{A} \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{B})\right]. \tag{3.8}$$

To check this above equality, first, note that

$$\varrho \cdot \hat{A} \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{B}) = \sum_{jklmn} p_j |j\rangle\langle j|A_{kl}|k\rangle\langle l|2\frac{B_{mn}}{p_m + p_n}|m\rangle\langle n| = \sum_{jmn} p_j \frac{2A_{jm}B_{mn}}{p_m + p_n}|j\rangle\langle n|.$$
(3.9)

The trace of this product of three operators is simply equal to

$$\mathrm{Tr}\left[\varrho \cdot \hat{A} \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{B})\right] = \sum_{jm} p_j \frac{2A_{jm}B_{mj}}{p_m + p_j}.$$
(3.10)

Now the realis is given by one-half of the sum of this expression plus the complex conjugate, i.e.

$$\mathrm{ReTr}\left[\varrho \cdot \hat{A} \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{B})\right] = \sum_{jm}\left[p_j\frac{A_{jm}B_{mj}}{p_m + p_j} + p_j\frac{A_{jm}^* B_{mj}^*}{p_m + p_j}\right].$$
(3.11)

But since the operator $\hat{A}$ and $\hat{B}$ are hermitian, then obviously $A_{jm}^* = A_{mj}$. So we have

$$\mathrm{ReTr}\left[\varrho \cdot \hat{A} \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{B})\right] = \sum_{jm}\left[p_j\frac{A_{jm}B_{mj}}{p_m + p_j} + p_j\frac{A_{mj}B_{jm}}{p_m + p_j}\right] =$$
(3.12)

$$= \sum_{jm}\left[p_j\frac{A_{jm}B_{mj}}{p_m + p_j} + p_m\frac{A_{jm}B_{mj}}{p_m + p_j}\right] = \sum_{jm} A_{jm}B_{mj} = \mathrm{Tr}\left[\hat{A} \cdot \hat{B}\right].$$

In particular, we can make use of this identity, to express the nominator of the Fisher information in a more useful way, namely

$$\mathrm{Tr}\left[\hat{E}_x \varrho'\right] = \mathrm{ReTr}\left[\varrho \cdot \hat{E}_x \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right].$$
(3.13)

We now insert this expression into (3.5) and obtain that

$$F = \int dx \frac{\left(\mathrm{ReTr}\left[\varrho \cdot \hat{E}_x \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right]\right)^2}{\mathrm{Tr}\left[\hat{E}_x \hat{\varrho}(\theta)\right]} \leqslant \int dx \frac{\left|\mathrm{Tr}\left[\varrho \cdot \hat{E}_x \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right]\right|^2}{\mathrm{Tr}\left[\hat{E}_x \hat{\varrho}(\theta)\right]}.$$
(3.14)

Now, we introduce

$$\hat{O}_1 \equiv \frac{\hat{\varrho}^{1/2}\hat{E}_x^{1/2}}{\sqrt{\hat{E}_x \varrho}} \quad \text{and} \quad \hat{O}_2 = \hat{E}_x^{1/2}\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}^{1/2}$$
(3.15)

and make use of the Cauchy-Schwartz inequality

$$\left|\mathrm{Tr}\left[\hat{O}_1^\dagger \hat{O}_2\right]\right|^2 \leq \mathrm{Tr}\left[\hat{O}_1^\dagger \hat{O}_1\right]\mathrm{Tr}\left[\hat{O}_2^\dagger \hat{O}_2\right]$$
(3.16)

and finally arrive at an expression

$$F \leqslant \mathrm{Tr}\left[\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right].$$
(3.17)

Let us now comment, under which conditions inequalities (3.14) and (3.17) are saturated. The first one becomes an equality when

$$\text{Im}\,\text{Tr}\left[\varrho \cdot \hat{E}_x \cdot \hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right] = 0 \quad \text{for all } x. \tag{3.18}$$

On the other hand, the Cauchy-Schwartz inequality (3.17) is satisfied, when the two operators are proportional to each other, and this condition can be written down in a compact form

$$\hat{E}_x^{1/2}\hat{\varrho}^{1/2} = \lambda_x \hat{E}_x^{1/2}\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}^{1/2}. \tag{3.19}$$

The coefficient λ_X must be found and can be for instance evaluated by tracing the above condition by sides. This way we obtain that

$$\lambda_x = \frac{\text{Tr}\left[\hat{E}_x\hat{\varrho}\right]}{\text{Tr}\left[\hat{E}_x\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}\right]}. \tag{3.20}$$

Note that since the nominator of this expression is real (because it is the probability for measuring x), then Eq. (5.17a) boils down to requirement that λ_x is real. Since the Fisher informaltion on the R.H.S. of inequality (3.17) does not depend on $\hat{E}_x$ anymore, conditions (5.17a) and (5.17b) can be regarded as maximalization of F with respect to all possible measurements. Therefore, the resulting object, which we call the Quantum Fisher Information (QFI) and denote by F_Q is a result of optimalization of F with respect to all $\hat{E}_x$.

To make the last step, and provide a simple expression for the QFI, we still need to calculate the derivative of the density matrix with respect to the parameter θ. To this end, we assume that the dependence of the system on the parameter is a result of some (not necessarily unitary) evolution in the Hilbert space, governed by the operator $\hat{h}$. Therefore, we obtain that

$$\hat{\varrho}+d\theta\hat{\varrho}' = \sum_j (p_j+dp_j)e^{id\theta\hat{h}}|j\rangle\langle j|e^{-id\theta\hat{h}} \simeq \sum_j dp_j|j\rangle\langle j|+\sum_j p_j e^{id\theta\hat{h}}|j\rangle\langle j|e^{-id\theta\hat{h}}. \tag{3.21}$$

By expanding the exponents to the lowest order we get that

$$d\theta\varrho' = \sum_j dp_j|j\rangle\langle j| + id\theta\sum_{jk}(p_j - p_k)h_{kj}|k\rangle\langle j|, \tag{3.22}$$

where $h_{kj} = \langle k|\hat{h}|j\rangle$. Now we manipulate the expression for the QFI, i.e.

$$F_Q = \text{Tr}\left[\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right] = \text{Tr}\left[\left(\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\right)^2\hat{\varrho}\right] = \text{Tr}\left[\hat{\mathcal{R}}_{\hat{\varrho}}^{-1}(\hat{\varrho}')\hat{\varrho}'\right]. \tag{3.23}$$

Using the expression for $\hat{\mathcal{R}}^{-1}$ from equation (3.7), we obtain that

$$F_Q = \sum_{jk} \varrho'_{jk}\varrho'_{kj}\frac{2}{p_j + p_k} \tag{3.24}$$

and finally plugging (3.22) we end up with

$$F_Q = \sum_j \frac{1}{p_j} \left(\frac{\partial p_j}{\partial \theta} \right)^2 + 2 \sum_{j \neq k} \frac{(p_j - p_k)^2}{p_j + p_k} |h_{jk}|^2. \tag{3.25}$$

If the θ-dependence is introduced via some unitary transformation as for instance in Eq. (3.4), then the probabilities p_j do not change and the first term of the QFI can be dropped, giving

$$F_Q = 2 \sum_{j \neq k} \frac{(p_j - p_k)^2}{p_j + p_k} |h_{jk}|^2. \tag{3.26}$$

Although it is difficult to provide a simple physical interpretation of the above expression, one can easily provide the higher bound for F_Q. To this end, note that

$$F_Q = 2 \sum_{j \neq k} \frac{(p_j - p_k)^2}{p_j + p_k} |h_{jk}|^2 \leqslant 2 \sum_{j \neq k} \frac{(p_j + p_k)^2}{p_j + p_k} |h_{jk}|^2 = \tag{3.27a}$$

$$= 2 \sum_{j \neq k} (p_j + p_k) |h_{jk}|^2 = 4 \sum_{j \neq k} p_j \langle j|\hat{h}|k \rangle \langle k|\hat{h}|j \rangle = 4 \langle (\Delta \hat{h})^2 \rangle. \tag{3.27b}$$

Here, $\Delta \hat{h} = \hat{h} - \langle \hat{h} \rangle$. This inequality becomes an equality only if $p_j p_k = 0$ for all $j \neq k$, which is satisfied only for pure states, so let us underline this very important result, that when the input state is pure, then

$$F_Q^{\text{pure}} = 4 \langle (\Delta \hat{h})^2 \rangle. \tag{3.28}$$

3.2 Entanglement and the sub shot-noise sensitivity

We will now show, that the QFI does not solely provie a mathematical expression for the Fisher information optimized over all possible measurements. From our point of view, the most important gain from the inequality

$$F \leqslant F_Q \leqslant 4 \langle (\Delta \hat{h})^2 \rangle \tag{3.29}$$

is that it provides a powerful tool which discrimanates more from less *useful* states from the parameter estimation point of view. To demonstrate this property of the QFI, we will follow the arguments of Smerzi and Pezzé, see [5]. Before we present their line of reasoning however, we need to finally focus our attention on the particular example of parameter estimation, which is two-mode quantum interferometry.

3.2.1   Two-mode interferometers

As underlined from the very beggining, our main point of interest is the op-
timalization of the performance of a device, which we should refer two as a
two-mode quantum interferometer. It is an instrument, which imprints a rela-
tive phase θ between two *input* ports and subsequently recombines the signal to
provide some kind of interference pattern. At the *output*, some measurement is
performed in order to estimate what the value of the phase was. The word *quan-
tum* underlines, that the inteferometer might be fed with a non-classical state.
In this chapter we will explain what the non-classicality refers to and argue that
in some cases it might be useful for improving the sensitivity of the device. Fi-
nally, before we continue, we should mention that probably the most well-known
such two-mode instrument is the Mach-Zehdner Interferometer, which we will
introduce and discuss its properties in detail in the next chapter.

As stated above, the quantum interferometer operates on two modes. In
order to describe such a device, we introduce two operators $\hat{a}_{\rm in}$ and $\hat{b}_{\rm in}$, which
determine the incoming field. An interferometer can be now modelled as an
instrument, which transforms these input fields into some output ones. Here,
we will focus on *linear* interferometers, where the output field is a linear com-
bination of the input, i.e.

$$\left(\begin{array}{c} \hat{a}_{\rm out} \\ \hat{b}_{\rm out} \end{array} \right) = \hat{U}(\theta) \left(\begin{array}{c} \hat{a}_{\rm in} \\ \hat{b}_{\rm in} \end{array} \right) = \left(\begin{array}{c} U_{11}\hat{a}_{\rm in} + U_{12}\hat{b}_{\rm in} \\ U_{21}\hat{a}_{\rm in} + U_{22}\hat{b}_{\rm in} \end{array} \right). \tag{3.30}$$

The operator $\hat{U}(\theta)$ fully determines the properties of the interferometer. The
above transformation is unitary for as long as it preserves the commutation
relations, $\left[\hat{a}_{\rm out}^\dagger, \hat{a}_{\rm out} \right] = \left[\hat{b}_{\rm out}^\dagger, \hat{b}_{\rm out} \right] = 1$ and other equal to zero. This means
that, for instance $|U_{11}|^2 + |U_{12}|^2 = 1$ and the same for the bottom row of the
evolution matrix. Clearly, to satisfy this relation, the elements of the evolution
matrix can be written in a form $U_{kl} = e^{i\varphi}f(\alpha)$, where f is either a sine or a
cosine function.

It will be argued here in more detail, that such transformations can be
generated by means of three objects, called *angular momentum operators*

$$\hat{J}_x = \frac{1}{2} \left(\hat{a}_{\rm in}^\dagger \hat{b}_{\rm in} + \hat{a}_{\rm in}\hat{b}_{\rm in}^\dagger \right) \tag{3.31a}$$

$$\hat{J}_y = \frac{1}{2i} \left(\hat{a}_{\rm in}^\dagger \hat{b}_{\rm in} - \hat{a}_{\rm in}\hat{b}_{\rm in}^\dagger \right) \tag{3.31b}$$

$$\hat{J}_y = \frac{1}{2} \left(\hat{a}_{\rm in}^\dagger \hat{a}_{\rm in} - \hat{b}_{\rm in}^\dagger \hat{b}_{\rm in} \right). \tag{3.31c}$$

Any linear interferometric transformation can now be represented by

$$\hat{U}(\theta) = e^{-i\vec{v}(\theta)\hat{\vec{J}}}. \tag{3.32}$$

Here, $\vec{v}(\theta)$ is a three-dimensional vector, which components depend on θ, while
$\hat{\vec{J}}$ stands for a vector of operators (3.31). Note by the way, that these hermi-
tian operators form a closed algebra, i.e. their commutation relation is cyclic

$\left[\hat{J}_k, \hat{J}_l\right] = i\varepsilon_{klm}\hat{J}_m$. This expression closely resembles the relation between the angular momentum operators generating rotations in the three-dimensional space $\mathbb{R}^3$. In this latter case, the rotations form a $SO(3)$ group. In analogy, the transformations (3.32), which act on a two-mode space, form an $SU(2)$ group. Therefore, the linear interferometers defined in Eq. (3.30) are sometimes referred to as $SU(2)$ interferometers.

A particularly important group of inerferometers, which includes the Mach-Zehnder and the time-of-flight interferometers is a sub-set of (3.32) and can be written in a form

$$\hat{U}(\theta) = e^{-i\theta\vec{n}\hat{\vec{J}}}, \tag{3.33}$$

where $\hat{n}$ is a unit, θ-independent vector. Obviously, some devices, like the Rabi interferometer which will be discussed in chapter ..., cannot be written in the form (3.33). Nevertheless, to prove the relation between the particle entanglement and the sub shot-noise interferometry, we focus on interferometers which can be represented in the form (3.33).

3.2.2 Role of entanglement in quantum interferometry

We will now argue, following the reference [5], that particle entanglement is a necessary resource for the better-then-classical precision of the parameter estimation in quantum interferometry. To this end, we define a non-entangled or *separable* state of N particles as such which can be written as a convex combination of product states, i.e.

$$\hat{\varrho}_{\text{sep}} = \sum_k p_k \hat{\varrho}_k^{(1)} \otimes \ldots \otimes \hat{\varrho}_k^{(N)}. \tag{3.34}$$

It is crucial to note, that a linear interferometer (3.30) does not introduce particle entanglement by itself, i.e.

$$\hat{\varrho}_{\text{sep}}^{\text{in}} \xrightarrow{\hat{U}(\theta)} \hat{\varrho}_{\text{sep}}^{\text{out}}(\theta). \tag{3.35}$$

To see this, we note that the angular momentum operators are sums of the Pauli matrices acting on each particle separately, namely

$$\hat{J}_k = \frac{1}{2}\sum_{i=1}^{N} \hat{\sigma}_k^{(i)}. \tag{3.36}$$

Therefore, the generic interferometric transformations (3.32) and consequently also (3.33) can be written as

$$\hat{U}(\theta) = \prod_{i=1}^{N} \hat{U}^{(i)}(\theta). \tag{3.37}$$

Such an operator, which acts on each particle independently, cannot correlate them, which justifies Eq. (3.35).

Note now that according to the CRLB, the ultimate lower bound for the phase sensitivity is provided by the inverse of the QFI, i.e.

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}} \frac{1}{\sqrt{F_Q}}. \tag{3.38}$$

We now combine Equations (3.1) and (3.29) with (3.33), we obtain that

$$F_Q \leqslant 4\langle(\Delta\hat{J}_{\vec{n}})^2\rangle, \tag{3.39}$$

where $\hat{J}_{\vec{n}} = \vec{n} \cdot \hat{\vec{J}}$. Consider first a separable state in a simple form

$$\hat{\varrho}_{\text{sep}} = \hat{\varrho}_k^{(1)} \otimes \ldots \otimes \hat{\varrho}_k^{(N)}. \tag{3.40}$$

For this state, and using Eq. (3.36) we have that

$$F_Q^{(\text{sep})} \leqslant 4\langle(\Delta\hat{J}_{\vec{n}})^2\rangle = \left\langle\left(\sum_{i=1}^N \hat{\sigma}_{\vec{n}}^{(i)}\right)^2\right\rangle - \left\langle\sum_{i=1}^N \hat{\sigma}_{\vec{n}}^{(i)}\right\rangle^2 \tag{3.41a}$$

$$= N - \left[\left\langle\sum_{i=1}^N \hat{\sigma}_{\vec{n}}^{(i)}\right\rangle^2 - \sum_{i\neq j}\langle\hat{\sigma}_{\vec{n}}^{(i)}\rangle\langle\hat{\sigma}_{\vec{n}}^{(j)}\rangle\right] = N - \sum_i \langle\hat{\sigma}_{\vec{n}}^{(i)}\rangle^2 \leqslant N \tag{3.41b}$$

Since the QFI is convex, then for an arbitrary separable state of the form (3.34) we obtain that

$$F_Q^{(\text{sep})} \leqslant N. \tag{3.42}$$

This is a very important result – for separable states the sensitivy of an interferometer is bounded by the *shot-noise limit*

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}} \frac{1}{\sqrt{N}}. \tag{3.43}$$

On the other hand, quite generally we can tell that

$$F_Q \leqslant 4\langle(\Delta\hat{J}_{\vec{n}})^2\rangle \leqslant \left\langle\left(\sum_{i=1}^N \hat{\sigma}_{\vec{n}}^{(i)}\right)^2\right\rangle \leqslant N^2. \tag{3.44}$$

This limiting value

$$F_Q = N^2 \tag{3.45}$$

is called the *Heisenber limit*. According to the above considerations, the precision of *any* linear interferometer like in Eq. (3.33) is therefore bounded by

$$\Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}} \frac{1}{N}. \tag{3.46}$$

Any values from the range

$$\frac{1}{\sqrt{m}} \frac{1}{\sqrt{N}} \geqslant \Delta\theta_{\text{est}} \geqslant \frac{1}{\sqrt{m}} \frac{1}{N} \tag{3.47}$$

are atteinable only with particle-entangled input states. The conclusion of this chapter is that the particle entanglement is the necessary resource for the sub shot-noise interferometry.

Chapter 4

Example: Mach-Zehdner Interferometer

In the previous chapters, we have introduced the basic theoretical tools of quantum interferometry. We are now prepared to analyze in detail the performance of the Mach-Zehnder Interferometer (MZI) – an archetype of all two-mode interferometers.

The MZI is realized in three separate stages, schematically shown in Fig. 4.1. At the beggining, two-mode state enters the interferometer and the signal from the modes is mixed on a 50/50 beam-splitter, denoted here by BS_1. Then, a relative phase θ is imprinted between the arms and finally another beam-splitter BS_2 mixes the modes again to produce the θ-dependent output signal[1]. We will first verify, how this whole sequence can be represented in terms of evolution operator like in Eq. (3.33), the identify the usefully entangled states and finally discuss the performance of different estimation strategies.

4.1 Evolution operator

To write down the action of the MZI in a concise form like in Eq. (3.33), we need to represent each of its three forming steps in terms of the two-mode language. The first beam-splitter takes the input operators $\hat{a}_{\text{in}}$ and $\hat{b}_{\text{in}}$ and produces a pair of mixed modes with equal weights, which is why it is actually refered to as a 50/50 device. It can be easily checked that one good choice for such a beam-splitter is

$$\hat{a}_1 = \frac{1}{\sqrt{2}} \left(\hat{a}_{\text{in}} - i\,\hat{b}_{\text{in}} \right) \tag{4.1a}$$

$$\hat{b}_1 = \frac{1}{\sqrt{2}} \left(\hat{b}_{\text{in}} - i\,\hat{a}_{\text{in}} \right). \tag{4.1b}$$

---

[1] An optical MZI requires also two mirrors M_1 and M_2 in order to focus the light beams on the second beam-splitter.
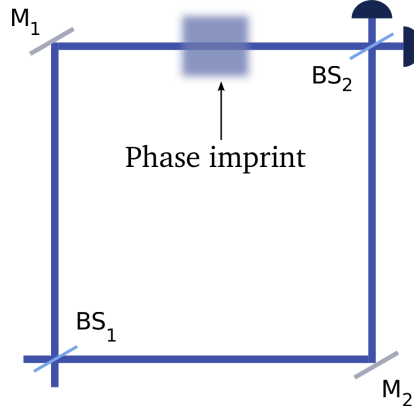
Figure 4.1: Schematic representation of the MZI. The two incoming modes are mixed on the first beam-splitter BS_1, and later a relative phase is imprinted beetween the arms. Finally, BS_2 produces an interferometric signal, which is measured by two detectors, represented here by two dark blue half-circles

It is "good" in a sence that it preserves the commutation relations, including the vanishing cross-commutator $[\hat{a}_1, \hat{b}_1^\dagger] = 0$. It is "one" of good choices, because the global phases can be taken arbitrary, according to the convention taken. In the next stage, a relative phase is imprinted between the arms, i.e.

$$\hat{a}_2 = e^{-i\frac{\theta}{2}}\hat{a}_1 \tag{4.2a}$$

$$\hat{b}_2 = e^{i\frac{\theta}{2}}\hat{b}_1. \tag{4.2b}$$

Finally, another beam-splitter recombines the signal again, giving

$$\hat{a}_{\text{out}} = \frac{1}{\sqrt{2}}\left(\hat{b}_2 - i\,\hat{a}_2\right) \tag{4.3a}$$

$$\hat{b}_{\text{out}} = \frac{1}{\sqrt{2}}\left(\hat{a}_2 - i\,\hat{b}_2\right). \tag{4.3b}$$

We can now express the output operators in terms of the input ones and arrive at[2]

$$\begin{pmatrix} \hat{a}_{\text{out}} \\ \hat{b}_{\text{out}} \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} = \begin{pmatrix} \hat{a}_{\text{in}} \\ \hat{b}_{\text{in}} \end{pmatrix}. \tag{4.4}$$

The MZI is therefore a rotation of the initial basis by the angle $\frac{\theta}{2}$. But around which axis is the rotation performed? To answer this question we make use of the Baker-Campbell-Hausdorff formula, which says that if $\hat{X}$ and $\hat{Y}$ do not commute, than

$$e^{\hat{X}}Ye^{-\hat{X}} = \hat{Y} + [\hat{X}, \hat{Y}] + \frac{1}{2!}[\hat{X}, [\hat{X}, \hat{Y}]] + \frac{1}{3!}[\hat{X}, [\hat{X}, [\hat{X}, \hat{Y}]]] + \dots \tag{4.5}$$

---

[2]up to an irrelevant phase-factor

With help of this equality, we finally obtain that

$$\left(\begin{array}{c} \hat{a}_{\mathrm{out}} \\ \hat{b}_{\mathrm{out}} \end{array} \right) = e^{-i\theta \hat{J}_y} \left(\begin{array}{c} \hat{a}_{\mathrm{in}} \\ \hat{b}_{\mathrm{in}} \end{array} \right) e^{i\theta \hat{J}_y}. \tag{4.6}$$

Therefore, the evolution operator of the MZI is simply

$$\hat{U}_{\mathrm{mzi}}(\theta) = e^{-i\theta \hat{J}_y}. \tag{4.7}$$

This expression allows for further analysis of the properties of the MZI.

4.2 Usefully entangled states

We have found the evolution operator for the MZI, thus we can proceed and check which types of states are usefully entangled for this interferometer. First of all, note that since the generator $\hat{h}$, which is necessary for the evaluation of the QFI, is $\hat{h} = \hat{J}_y$, then according to the inequality (3.29), the QFI itself can be bounded by

$$F_Q^{(\mathrm{mzi})} \leqslant 4 \left\langle \Delta^2 \hat{J}_y \right\rangle. \tag{4.8}$$

In order to provide the upper bound for the preformance of the MZI, we will now focus on pure states, for which the above inequality is saturated.

First, let us identify the ultimately entangled state, which provides the Heisenberg limit sensitivity. It is such a state, which must provide the maximal variance of the $\hat{J}_y$ operator. To find this state, first note that the NOON state

$$|\psi_z\rangle = \frac{1}{\sqrt{2}} \left(|N, 0\rangle + |0, N\rangle \right) \tag{4.9}$$

is a state, where all the particles are either in the a or in the b mode. Therefore, such a superposition leads to the maximal fluctuations of the population imbalance between the two arms of the interferometer. This can be formally stated by noting that the variance of the angular momentum operator, which is related to the population imbalance, i.e. $\hat{J}_z$ is maximal in the NOON state

$$4 \left\langle \Delta^2 \hat{J}_z \right\rangle_{\mathrm{NOON}} = N^2. \tag{4.10}$$

However, when we calculate the variance of the operator $\hat{J}_y$ in this state, we obtain the QFI to be limited to the shot-noise level, i.e.

$$F_Q^{(\mathrm{mzi})} = 4 \left\langle \Delta^2 \hat{J}_y \right\rangle = N. \tag{4.11}$$

However, we can now utilize the algebraic properties of the angular momentum operators. Namely, note that when we rotate the $\hat{J}_y$ operator around the x axis by the angle equal to $\frac{\pi}{2}$ we obtain the $\hat{J}_z$ operator, or formally

$$e^{-i\frac{\pi}{2}\hat{J}_x} \hat{J}_y e^{i\frac{\pi}{2}\hat{J}_x} = \hat{J}_z. \tag{4.12}$$

This is a crucial observation, since we can now write down Eq. (4.10) as

$$N^2 = 4 \left\langle \Delta^2 \hat{J}_z \right\rangle_{\mathrm{NOON}} = 4 \left\langle \Delta^2 \left(e^{-i\frac{\pi}{2}\hat{J}_x} \hat{J}_y e^{i\frac{\pi}{2}\hat{J}_x} \right) \right\rangle_{\mathrm{NOON}} \tag{4.13}$$

Now, the x-rotation operators can be pulled out to act on the state itself. Therefore, if we define the "y-NOON state" as

$$|\psi_y\rangle = e^{i\frac{\pi}{2}\hat{J}_x} \frac{1}{\sqrt{2}} \left(|N, 0\rangle + |0, N\rangle \right), \tag{4.14}$$

we find that

$$4 \left\langle \Delta^2 \hat{J}_z \right\rangle_{y-\mathrm{NOON}} = N^2. \tag{4.15}$$

We conclude that the state (4.14) is an ultimately entangled state which improves the performance of the MZI up to the Heisenberg level.

The remaining question is, how one could find a broader family of usefully entangled states for the MZI. To identify these states, we repeat that the NOON state has extreme site-to-site population imbalance fluctuations. The rotation around the x axis, as in Eq. (4.14) maps these fluctuations onto maximal phase fluctuations between the two modes. However, the relative phase and the population imbalance between the modes are two conjugate variables. Therefore, in order to produce an entangled state, which has large phase fluctuations (useful for the MZI), one can consider a family of *number squeezed* states. Now we need to formalize this statement.

First, consider what is called a Coherent Spin State (CSS), which has a following form

$$|\psi\rangle = \frac{1}{\sqrt{N!}} \left(\frac{\hat{a}^\dagger + \hat{b}^\dagger}{\sqrt{2}} \right)^N |0, 0\rangle. \tag{4.16}$$

It is a state, where each particle is in a coherent superposition of being in the left and in the right arm. Since it is a product state of N particles, naturally it is not entangled, and indeed we imediatelly obtain that

$$4 \left\langle \Delta^2 \hat{J}_z \right\rangle_{\mathrm{CSS}} = N. \tag{4.17}$$

Note that the CSS can be written in the basis of mode occupations, and then the coefficients of the expansion have binomial form

$$|\psi\rangle = \sum_{n=0}^{N} C_n |n, N - n\rangle = \frac{1}{2^{\frac{N}{2}}} \sum_{n=0}^{N} \sqrt{\binom{N}{n}} |n, N - n\rangle. \tag{4.18}$$

For large N, such binomial distribution can be approximated with a Gaussian, i.e.

$$\frac{1}{2^{\frac{N}{2}}} \sqrt{\binom{N}{n}} \simeq \frac{1}{\left(\frac{N\pi}{2} \right)^{\frac{1}{4}}} e^{-\frac{\left(n - \frac{N}{2} \right)^2}{N}}. \tag{4.19}$$

The width of this Gaussian is equal to $\sigma = \sqrt{N}$. This width is what governs the fluctuations between the two modes. And so, if $\sigma \to 0$, these fluctuations will we supressed down to zero. This is because the Gaussian will tend to the delta function, and the whole sum over all possible occupations labeled by n will boil down to a single term with $n = \frac{N}{2}$. Such state is called a twin-Fock state and reads

$$|\psi\rangle = \left| \frac{N}{2}, \frac{N}{2} \right\rangle. \tag{4.20}$$

Indeed, for this ultimately numbers-squeezed state, where there are no population imbalance fluctuations between the arms of the interferometer, we obtain that

$$4 \left\langle \Delta^2 \hat{J}_z \right\rangle_{\text{TF}} = \frac{N}{2} (N+1), \tag{4.21}$$

which is almost at the Heisenberg level. Also, all the states with the coefficients represented in a Gaussian form

$$C_n = \frac{1}{\left(\frac{\sigma^2 \pi}{2} \right)^{\frac{1}{4}}} e^{- \frac{\left(n - \frac{N}{2} \right)^2}{\sigma^2}} \tag{4.22}$$

with the fluctuations below the CSS level, i.e. $\sigma < N$, the sensitivity of the MZI is sub shot-noise, i.e.

$$F_Q^{(\text{mzi})} > N. \tag{4.23}$$

This way, we have identified the usefully entangled states for the MZI – these are the number-squeezed states.

4.3 Estimation from the population imbalance

We are now prepared to inestigate the performance of different estimation strategies at the output of the MZI. We will focus on the estimation from the measurements of the population imbalance between two arms of the MZI. We will demonstrate that the estimation from the full population imbalance probability is optimal (i.e. it saturates the QFI). Also, we will introduce the commonly used notion of spin-squeezing and show how it naturally appears in the estimation theory in the context of the MZI.

As stated above, typically the estimation of the parameter θ at the output of the MZI is based on the measurement of the population imbalance between the two arms of an interferometer. We will now provide some general results for this measurement and then focus on a particular estimator – the average population imbalance. To provide the upper bounds for the parameter precision, we will focus on pure input states.

We argued in Chapter 1 that the central object for the evaluation of the precision of the parameter estimation from a measurement of a quantity x is the conditional probability for obtaining x at the output of an interferometer given the value of the parameter θ. When the phase is estimated from the population imbalance, this probability is equal to the squared modulus of the

projection of the output state onto the Fock state with n particles in one and $N - n$ particles in the other arm, which formally reads

$$p(n|\theta) = \left| \left\langle n, N - n \middle| e^{-i\theta \hat{J}_y} \middle| \psi_0 \right\rangle \right|^2. \tag{4.24}$$

Decomposing the two-mode state $|\psi_0\rangle$ in the basis of mode occupations, we obtain that

$$p(n|\theta) = \left| \sum_{n'=0}^{N} C_{n'}^{(0)} \left\langle n, N - n \middle| e^{-i\theta \hat{J}_y} \middle| n', N - n' \right\rangle \right|^2. \tag{4.25}$$

The expression $\left\langle n, N - n \middle| e^{-i\theta \hat{J}_y} \middle| n', N - n' \right\rangle$ is real and is an element of what is called the Wigner rotation matrix [8]. As we argue in Chapter ..., these elements can be expressed analytically as follows

$$d_{jk}(\theta) = \sqrt{\frac{k!(N-k)!}{j!(N-j)!}} \left[\sin\left(\frac{\theta}{2}\right) \right]^{j-k} \left[\cos\left(\frac{\theta}{2}\right) \right]^{j+k-N} P_{N-j}^{j-k,j+k-N}(\cos\theta), \tag{4.26}$$

where $P_\mu^{\alpha,\beta}(x)$ is the Jacobi polynomial and $j = n - \frac{N}{2}$, $k = n' - \frac{N}{2}$. In order to calculate the CRLB, which is the lower limit for the parameter uncertainty, one should calculate the Fisher information related to the population imbalance measurement

$$F_{\text{imb}} = \sum_{n=0}^{N} \frac{1}{p(n|\theta)} \left(\frac{\partial}{\partial\theta} p(n|\theta) \right)^2. \tag{4.27}$$

The question now is, how this Fisher information is related to the QFI. In general, we know that

$$F_{\text{imb}} \leqslant 4 \left\langle \Delta^2 \hat{J}_y \right\rangle. \tag{4.28}$$

Note that once the interferometer (MZI) and the measurement (population imbalance) are fixed, the value of the Fisher information depends only on the phase θ and the input state $|\psi\rangle$. Our aim now is to verify if the above inequality is saturated for any input states $|\psi\rangle$, and if so, can we identify the family of such defined *optimal states*. Unfortunatelly, the Fisher information in the form of Eq. (4.27) cannot be calculated analytically, because the expression for the elements of the Wigner matrix (4.26) are complicated. Nevertheless, one can approach the whole problem of the CRLB for the population imbalance measurement from a different site. To this end, however, we must go back to the definition of the QFI and express it in a different manner than in Eq. (5.13).

4.3.1   QFI and the statistical distance

In the paper, where the QFI is introduced for the first time [6], Braunstein and Caves show how the QFI from Eq. (5.13) is related to the statistical distance between two neighboring states. Although this general result is valid whenever
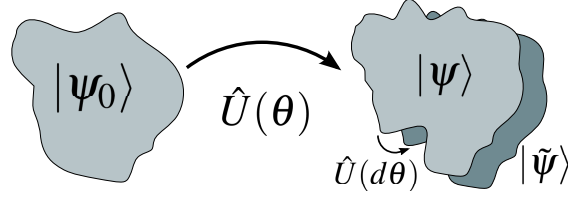
Figure 4.2: Schematic representation of the two steps performed to calculate the QFI. The input state $|\psi_0\rangle$, which enters the interferometer, is transformed by a unitary evolution operator $\hat{U}(\theta)$, giving the output state $|\psi\rangle$. To calculate the speed at which the state changes, and thus the statistical distance, we make a further infinitesimal rotation $\hat{U}(\theta)$ to obtain $\left|\tilde{\psi}\right\rangle$.

a parameter is estimated from measurements performed on a θ-dependent state $|\psi\rangle$, below we present this formalism in context of quantum interferometry.

To this end, we follow the scheme pictured in Fig. 5.2. First, consider a pure two-mode *input* state $|\psi_0\rangle$, which is transformed by a unitary evolution operator of the MZI, i.e. $\hat{U}(\theta) = e^{-i\theta\hat{J}_y}$, As a result, we obtain a θ-dependent *output* state, which when expanded in the basis of mode occupations reads

$$|\psi\rangle = e^{-i\theta\hat{J}_y}|\psi_0\rangle = \sum_{j=0}^{N}\sqrt{p_j}e^{i\varphi_j}|j\rangle \equiv \sum_{j=0}^{N}C_j|j\rangle. \tag{4.29}$$

Here for the simplicity of notation we replaced the $|j, N-j\rangle$ with $|j\rangle$. The neighboring state is found by applying a further infinitesimal transformation $e^{-id\theta\hat{J}_y}$, which gives

$$\left|\tilde{\psi}\right\rangle = e^{-id\theta\hat{J}_y}|\psi\rangle \simeq \sum_{j=0}^{N}(1 - id\theta\hat{J}_y)C_j|j\rangle = \sum_{j=0}^{N}(1 - id\theta\eta_j)C_j|j\rangle \equiv \sum_{j=0}^{N}\tilde{C}_j|j\rangle. \tag{4.30}$$

The coefficient η_j is a result of acting with $\hat{J}_y$ on a ket $|j\rangle$ and is equal to

$$\eta_j = \frac{1}{2i}\frac{\alpha_j C_{j+1} - \alpha_{j-1}C_{j-1}}{C_j}. \tag{4.31}$$

where $\alpha_j = \sqrt{(j+1)(N-j)}$. Note that the state The state (5.34) can be alternatively written as

$$\left|\tilde{\psi}\right\rangle = \sum_{j=0}^{N}\sqrt{p_j + dp_j}e^{i(\varphi_j + d\varphi_j)}|j\rangle. \tag{4.32}$$

where the probability- and phase-increments are

$$dp_j = |\tilde{C}_j|^2 - |C_j|^2 = 2\operatorname{Im}\eta_j|C_j|^2 d\theta \tag{4.33a}$$

$$e^{id\varphi_j} = \frac{\tilde{C}_j}{|\tilde{C}_j|}\frac{|C_j|}{C_j} = e^{-i\operatorname{Re}\eta_j d\theta}. \tag{4.33b}$$

We are now ready to calculate the crucial object, necessary for the QFI, which is the distance between two neighboring states and which is equal to

$$
\begin{aligned}
ds_{\mathrm{ps}}^2 &= 1 - |\langle\psi|\tilde{\psi}\rangle|^2 = \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4\left[\sum_{j=0}^{N} p_j d\varphi_j^2 - \left(\sum_{j=0}^{N} p_j d\varphi_j\right)^2\right] \\
&\equiv \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4\Delta^2 d\varphi.
\end{aligned}
\tag{4.34a}
$$

According to [6], the QFI can be interpreted as the speed at which the state changes upon the infinitesimal increment of the parameter θ and reads

$$
F_Q = \frac{ds_{\mathrm{ps}}^2}{d\theta^2}.
\tag{4.35}
$$

Substituting the expressions for the phase- and probability increments from Eq. (4.33) into (5.38), we obtain that

$$
F_Q = 4\sum_{j=0}^{N} |C_j|^2 \left(\mathrm{Im}\,\eta_j\right)^2 +
\tag{4.36a}
$$

$$
4\sum_{j=0}^{N} |C_j|^2 \left(\mathrm{Re}\,\eta_j\right)^2 - 4\left(\sum_{j=0}^{N} |C_j|^2 \mathrm{Re}\,\eta_j\right)^2.
\tag{4.36b}
$$

It can be easily demonstrated that when the phase is acquired through a unitary transformation like in Eq. (5.33), the above expression simplifies to the well known result $F_Q = 4\left\langle\Delta^2\hat{J}_y\right\rangle$. However, as we argue below, for the purpose of finding the optimal measurements, it is more convenient to keep the QFI in the form of Equations (5.39), i.e. as a sum of two non-negative parts – the change of the probability p_j and the variance of the phase increment $d\varphi_j$. Namely, if the latter term is zero, then the CFI calculated using the probability $p_j = |\langle j|\psi\rangle|^2$ of finding the system in state $|j\rangle$ is equal to QFI. This means that no information about θ is carried by the phases φ_j, which by definition are not witnessed by the projection measurement $|j\rangle\langle j|$ and thus do not contribute to probabilities p_j.

4.3.2   When CRLB saturates the QFI

With help of Equations (5.39) we are now ready to identify the family of states, for which the estimation from the population imbalance is optimal for the MZI. The conditional probability from Eq. (4.24) can be alternatively to Eq. (4.25) expressed using the formalism from the previous paragraph as

$$
p(j|\theta) = |\langle j|e^{-i\theta\hat{J}_y}|\psi_0\rangle|^2 = |\langle j|\psi\rangle|^2 = |C_j|^2.
\tag{4.37}
$$

The derivative of this expression, which is needed for the calculation of the CRLB, reads

$$\partial_\theta p(j|\theta) = 2|C_j|^2 \text{Im}\,\eta_j. \tag{4.38}$$

Therefore, the CFI calculated with (5.47) is equal to

$$F_{\text{imb}} = \sum_{j=0}^{N} \frac{1}{p(j|\theta)} \left(\frac{\partial\,p(j|\theta)}{\partial\theta}\right)^2 = 4\sum_{j=0}^{N} |C_j|^2\,(\text{Im}\,\eta_j)^2 \tag{4.39}$$

By comparing this expression to the QFI from Eq. (5.39), we conclude that the estimation from the population imbalance is optimal only if the other terms in line (5.39b) vanish, which requires $\text{Re}\,\eta_j \equiv 0$ for all j. This is equivalent to the condition $C_j = e^{i\phi}a_j$, where $a_j \in \mathbb{R}$ and ϕ is a common phase. In particular for $\phi = 0$, the measurement is optimal, when all C_j's are real. Since the elements of the Wigner rotation matrix (4.26) are all real, we conclude that if the input state of the MZI has real coefficients, the estimation from the population imbalance is optimal.

4.3.3 Estimation from the average population imbalance

What we have considered so far, is the CRLB for the estimation from the population imbalance. We know from Chapter 1.3, that the precision $\Delta\theta_{\text{est}} = \frac{1}{\sqrt{m}}\frac{1}{\sqrt{F_{\text{imb}}}}$ can be acheived when one uses the MLE constructed on the probability $p(j|\theta)$. However, as we already pointed out, usually it is very difficult to calibrate the interferometer and acquire precise information about this probability. Therefore, it is reasonable to construct a much simpler estimator based on the population imbalance measurement.

Here we will consider a case, when the phase is estimated from the *average* population imbalance, rather than from the full probability. In order to decide, how such estimator should be constructed, first we calculate the average population imbalance at the output of the MZI. We know that it is governed by the $\hat{J}_z$ operator. And so, we have that

$$\langle n \rangle = \langle \psi(\theta)|\hat{J}_z|\psi(\theta)\rangle = \langle\psi_0|e^{i\theta\hat{J}_y}\hat{J}_z e^{-i\theta\hat{J}_y}|\psi_0\rangle = \langle\hat{J}_z\rangle\cos\theta + \langle\hat{J}_x\rangle\sin\theta, \tag{4.40}$$

where the averages in the last expression are calculated in the input state. For simplicity, let us consider *path symmetric* states, for which the expansion coeffients posses the symmetry $C_n^{(0)} = C_{N-n}^{(0)}$. For these states, $\langle\hat{J}_z\rangle = 0$, so we obtain a simple expression

$$\langle n \rangle = \langle\hat{J}_x\rangle\sin\theta. \tag{4.41}$$

Now, imagine this function is used to estimate the relative phase θ. Let's assume, that in a first experiment, some population imbalance is measured, say n_1. We can now put this result in the R.H.S. of Eq. (4.41) and invert the function to obtain the first estimated value of θ, i.e.

$$\theta_{\text{est}}^{(1)} = \arcsin(n_1). \tag{4.42}$$

Now, the measurement is repeated m times and the resulting estimated θ's are averaged, to give

$$\langle\theta_{\text{est}}\rangle = \frac{1}{m}\sum_{i=1}^{m}\arcsin(n_i) \neq \arcsin(\langle n\rangle) = \theta. \tag{4.43}$$

Clearly, such naively constructed estimator is biased! As we will now argue, a proper construction of the estimator based on the average population imbalance, goes through – not surprisingly – the C.L.T.

Assume that instead we average the measured outcomes over first m repetitions and get a single number

$$n_1^{(m)} = \frac{1}{m}\sum_{i=1}^{m}n_i \tag{4.44}$$

before utilizing this data to obtain the value of the parameter in a manner analogue to Eq. (4.43), which is

$$\theta_{\text{est},1}^{(m)} = \arcsin(n_1^{(m)}). \tag{4.45}$$

Now, imagine that this value of the parameter is averaged over many (say M) mean results $n_k^{(m)}$. In this way we obtain

$$\langle\theta_{\text{est}}^{(m)}\rangle = \frac{1}{M}\sum_{k=1}^{M}\arcsin(n_k^{(m)}) \xrightarrow{M\to\infty} \int dn^{(m)}\; \mathcal{P}(n^{(m)}|\theta)\;\arcsin(n^{(m)}). \tag{4.46}$$

Now, according to the CLT, since $n^{(m)}$ is an average of a random variable over $m \gg 1$ repetitions, the probability $\mathcal{P}$ will tend to a Gaussian in a form

$$\mathcal{P}(n^{(m)}|\theta) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{m}}}\exp\left(-\frac{(n^{(m)} - \langle n\rangle)^2}{2\frac{\sigma^2}{m}}\right). \tag{4.47}$$

Here, $\langle n\rangle$ is the *true* average value defined in Eq. (4.41), while the widht is the *true* variance of the population imbalance operator, which for path-symmetric states is

$$\sigma^2 \equiv \langle\psi(\theta)|\hat{J}_z^2|\psi(\theta)\rangle - \langle\psi(\theta)|\hat{J}_z|\psi(\theta)\rangle^2 = \left\langle\Delta^2\hat{J}_x\right\rangle\sin^2(\theta) + \left\langle\hat{J}_z^2\right\rangle\cos^2(\theta). \tag{4.48}$$

Now, when $m \to \infty$ in Eq. (4.49), then the probability $\mathcal{P}$ tends to the Dirac delta peaked around the true average value and the integral gives

$$\langle\theta_{\text{est}}^{(m)}\rangle \xrightarrow{m\to\infty} \int dn^{(m)}\;\delta\big(n^{(m)} - \langle n\rangle\big)\;\arcsin(n^{(m)}) = \arcsin(\langle n\rangle) = \theta. \tag{4.49}$$

So in this case, careful choice provided an unbiased estimator. If it is so, we can derive the CRLB for the probability $\mathcal{P}$, i.e. calculate the Fisher information

$$F = \int dn^{(m)}\frac{1}{\mathcal{P}(n^{(m)}|\theta)}\left(\frac{\partial}{\partial\theta}\mathcal{P}(n^{(m)}|\theta)\right)^2. \tag{4.50}$$

Since the probability is Gaussian, then this integral can be evaluated analytically. Conistently using the fact that $m \gg 1$ we obtain

$$\Delta\theta_{\text{est}}^{(m)} = \frac{1}{\sqrt{m}} \frac{\sigma}{\left| \frac{\partial \langle n \rangle}{\partial \theta} \right|^2}. \tag{4.51}$$

This the well-known error propagation formula, often quoted with no explanation as the bound for the parameter estimation. We now see, that it appears, when the parameter θ is estimated from some average, and when the CLT is satisfied. We now plug in the above equation the expressions for the average (4.41) and the variance (4.48) and obtain

$$\Delta\theta_{\text{est}}^{(m)} = \frac{1}{\sqrt{m}} \frac{\sqrt{\left\langle \Delta^2 \hat{J}_x \right\rangle \sin^2(\theta) + \left\langle \hat{J}_z^2 \right\rangle \cos^2(\theta)}}{\left| \left\langle \hat{J}_x \right\rangle \cos(\theta) \right|}. \tag{4.52}$$

In particular, for $\theta \simeq 0$ we obtain the final equation for the squared phase sensitivity

$$\Delta^2 \theta_{\text{est}}^{(m)} = \frac{1}{m} \frac{\left\langle \hat{J}_z^2 \right\rangle}{\left\langle \hat{J}_x \right\rangle^2} \equiv \frac{1}{m} \frac{1}{N} \xi_N^2. \tag{4.53}$$

The object, which appears on the R.H.S. is called the *spin squeezing parameter* and is defined as

$$\xi_N^2 \equiv N \frac{\left\langle \hat{J}_z^2 \right\rangle}{\left\langle \hat{J}_x \right\rangle^2}. \tag{4.54}$$

Clearly, when $\xi_N^2 < 1$, then

$$\Delta\theta_{\text{est}}^{(m)} < \frac{1}{\sqrt{m}} \frac{1}{\sqrt{N}}, \tag{4.55}$$

so we obtain the sub shot-noise sensitivity. On the other hand, from Chapter 3.2 we know that it means, that the input state of the MZI was entanglement. To summarize, if one can demonstrate $\xi_N^2 < 1$, it is equivalent to showing that the system is particle-entangled. This motivates the current effort to generate and detect the spin-squeezing in two-mode systems created from Bose-Einstein Condensates.

Chapter 5

Optimal measurements

In this chapter, we show how optimal measurements can be determined in various two-mode interferometric systems. Our aim is to show, that optimal measurements in estimation theory are, in general, "subtle objects". To this end, we start with the simplest possible two-mode object, which is a single qubit. We find the optimal measurements, when the interferometric transformation is a rotation of the state on the Bloch sphere by an unknown angle θ. While for pure states $\hat{\rho}_0$, there is a whole continuity of optimal estimation strategies, they boil down to only two possibilities when $\hat{\rho}_0$ is mixed. For two qubits such general considerations are not possible anymore, so we will use a particular example, namely the symmetrized Werner state. We find the optimal measurement and discuss the precision which can be reached when the phase is estimated from the population imbalance between the two modes. We will then show under which conditions the estimation from the population imbalance or the N-th order correlation function is optimal, using N-qubit pure state at the input.

Finally, in order to show how mixing the input state makes things less trivial, we introduce some phase noise into the evolution of the state inside the interferometer. We show what is the impact of the noise on the precision of different, commonly employed estimation protocols which use the measurement of the population imbalance at the output ports.

5.1   Single qubit

We begin our analysis of the optimal estimation strategies with a most basic two-mode object, namely a single qubit, which is rotated on the Bloch sphere by an unknown angle θ. As will be shown below, in this case a full family of optimal measurements can be identified, both for pure and mixed states.

5.1.1   General formulation

The density operator of a single qubit is a 2x2 matrix. Typically, it is convenient to represent the general density matrix of a qubit using the set of three Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (5.1)$$

These three trace-less matrices, together with a unit matrix form a basis, so not surprisingly the density matrix of a qubit can be represented as a combination of a scalar and three Pauli matrices

$$\hat{\rho}_{\text{in}} = \frac{1}{2}(\hat{\mathbb{1}} + \vec{s}_{\text{in}} \cdot \vec{\hat{\sigma}}) \qquad (5.2)$$

Depending on the length s_{in} of the vector $\vec{s}_{\text{in}}$, the state is either pure $(s_{\text{in}} = 1)$ or mixed $(s_{\text{in}} < 1)$.

In the scenario considered here, this generic initial state $\hat{\rho}_{\text{in}}$ undergoes a unitary phase-dependent interferometric transformation $\hat{U}(\theta)$. To establish the analogy between the N-qubit Mach-Zehner Interferometer considered in detail in Chapter 4 and a single-qubit operation, we notice that the former case is represented by a unitary transformation

$$\hat{U}_{\text{mzi}}(\theta) = e^{-i\theta\hat{J}_y}. \qquad (5.3)$$

Here $\hat{J}_y = \sum_{i=1}^{N} \frac{\hat{\sigma}_y^{(i)}}{2}$ is a sum of y-component Pauli matrices acting on the i-th particle. Clearly, a single-qubit analogy of the MZI is

$$\hat{U}(\theta) = \exp\left[-i\theta\frac{\hat{\sigma}_y}{2}\right] \qquad (5.4)$$

which transforms the initial density matrix into

$$\hat{\rho} = \hat{U}(\theta)\,\hat{\rho}_{\text{in}}\,\hat{U}^\dagger(\theta) = \frac{1}{2}(\hat{\mathbb{1}} + \vec{s} \cdot \vec{\hat{\sigma}}), \qquad (5.5)$$

where the three components of the rotated vector are

$$s_x = s_{\text{in},x}\cos\theta + s_{\text{in},z}\sin\theta \qquad (5.6\text{a})$$
$$s_y = s_{\text{in},y} \qquad (5.6\text{b})$$
$$s_z = s_{\text{in},z}\cos\theta - s_{\text{in},x}\sin\theta. \qquad (5.6\text{c})$$

Note that we have omitted the explicit dependence of $\hat{\rho}$ on the phase θ – in order to clarify the notation.

5.1.2   Classical and quantum Fisher information

To establish a direct correspondence with the formalism of the QFI introduced in Chapter 3 we will look for the optimal phase estimation strategies by considering the broadest set of measurements allowed by quantum mechanics. The

elements of this set are described in terms of the POVMs (Positive-Operator Valued Measures) [1], which are self-adjoint and have non-negative eigenvalues. However, since for a single qubit any 2x2 matrix can be represented using the unity and the three Pauli matrices, then naturally any POVM can be represented by the following operator

$$\hat{E}_{\vec{q}} = \gamma_{\vec{q}}(\hat{1} + \vec{q} \cdot \vec{\hat{\sigma}}), \tag{5.7}$$

where $|\vec{q}| \leqslant 1$ and furthermore the normalization of the set of POVM's requires that

$$\int d\vec{q}\,\gamma_{\vec{q}} = 1 \quad \text{and} \quad \int d\vec{q}\,\gamma_{\vec{q}}\,\vec{q} = 0. \tag{5.8}$$

The trace of the product of the POVM (5.7) and the density matrix gives the probability for finding the qubit alligned along $\vec{q}$ on the Bloch sphere,

$$p(\vec{q}|\theta) = \text{Tr}\left[\hat{\rho}\,\hat{E}_{\vec{q}}\right]. \tag{5.9}$$

Recall again, that the precision of the phase estimation from a series of m measurements is limited by the Cramer-Rao Lower Bound [1, 2]

$$\Delta\theta \geqslant \frac{1}{\sqrt{m}}\frac{1}{\sqrt{F}}, \tag{5.10}$$

where F is called the Classical Fisher Information and is equal to

$$F = \int d\vec{q}\,\frac{1}{p(\vec{q}|\theta)}\left(\frac{\partial\,p(\vec{q}|\theta)}{\partial\theta}\right)^2. \tag{5.11}$$

The CFI, which is a measure of information about θ contained in $\hat{\rho}$, depends on the particular choice of measurement – some methods of estimating θ are better than other. The *optimal* measurements are those which give the maximal value of F, called the Quantum Fisher Information (QFI) [6] – and through Eq. (5.10) maximal precision of phase estimation.

The maximization procedure of (5.11) gives the QFI, which for unitary transformations is

$$F_Q = 2\sum_{j,k}\frac{(p_j - p_k)^2}{p_j + p_k}\left|\langle j|\hat{h}|k\rangle\right|^2. \tag{5.12}$$

Here $\hat{h}$ is a generator of the phase transformation, which according to Eq. (5.4) is $\hat{h} = \frac{\hat{\sigma}_y}{2}$. The ket $|k\rangle$ denotes the k-th eigen-vector of the density matrix $\hat{\rho}$ from Eq. (5.6) with a corresponding eigen-value p_k. For this 2×2 matrix, the eigen-problem is easily solved and we obtain

$$F_Q = s_x^2 + s_z^2. \tag{5.13}$$

Using the expressions from Eq. (5.6), we note that the QFI does not depend on θ and the maximum $F_Q = 1$ is achieved with pure states with $s_y = 0$. This is a clear result – the y-component of the density matrix is not affected by the $\hat{\sigma}_y$ rotation, and so there is no information encoded in it. It is thus better to have a qubit state with no y component, in order to maximize the information input.

5.1.3   Optimal measurements

Next, we determine which are the optimal measurements, which give (5.13). Remember, that they are given by the expression

$$\hat{E}_{\vec{q}}\,\hat{\rho} = \lambda_{\vec{q}}\,\hat{E}_{\vec{q}}\hat{\mathcal{L}}_{\hat{\rho}}\,\hat{\rho} \tag{5.14}$$

with $\lambda_{\vec{q}} \in \mathbb{R}$. The super-operator $\hat{\mathcal{L}}_{\hat{\rho}}$ is called the symmetric logarithmic derivative and is defined by a relation

$$\frac{1}{2}\left(\hat{\rho}\hat{\mathcal{L}}_{\hat{\rho}} + \hat{\mathcal{L}}_{\hat{\rho}}\hat{\rho}\right) = \partial_{\theta}\hat{\rho}. \tag{5.15}$$

The optimal measuerements are found in two steps. First, by declaring the general form of the symmetric logarithmic derivative

$$\hat{\mathcal{L}}_{\hat{\rho}} = \alpha\hat{\mathbb{1}} + \vec{s}_{\perp}\vec{\hat{\sigma}} \tag{5.16}$$

and using the definition (5.15) we get that $\alpha = 0$ and $\vec{s}_{\perp} = s_z\vec{e}_x - s_x\vec{e}_z$. This result justifies the notation, since $\vec{s}\cdot\vec{s}_{\perp} = 0$.

In the second step we insert $\hat{\mathcal{L}}_{\hat{\rho}}$ into (5.14), and use the general parametrization of the POVM (5.7). By comparing the scalar and vector parts and then the real and imaginary parts we obtain a set of equations

$$\vec{q}\cdot(\vec{s}_{\perp}\times\vec{s}) = 0 \tag{5.17a}$$

$$\lambda_{\vec{q}} = \frac{1 + \vec{q}\cdot\vec{s}}{\vec{q}\cdot\vec{s}_{\perp}} \tag{5.17b}$$

$$\vec{q} + \vec{s} = \lambda_{\vec{q}}\left[\vec{s}_{\perp} - \vec{q}\times(\vec{s}_{\perp}\times\vec{s})\right] \tag{5.17c}$$

$$\vec{q}\times\vec{s} = \lambda_{\vec{q}}\left[\vec{s}_{\perp}\times\vec{s} + \vec{q}\times\vec{s}_{\perp}\right]. \tag{5.17d}$$

From (5.17a) we deduce that $\vec{q}$ lies in the plane spanned by vectors $\vec{s}$ and $\vec{s}_{\perp}$, so it can be written as $\vec{q} = q_1\vec{e}_s + q_2\vec{e}_{s_{\perp}}$. Here, $\vec{e}_s$ and $\vec{e}_{s_{\perp}}$ are unit vectors pointing into directions $\vec{s}$ and $\vec{s}_{\perp}$. This observation reduces the set of eight equations (5.17) to

$$\lambda_{\vec{q}} = \frac{1 + s\,q_1}{s_{\perp}q_2} \tag{5.18a}$$

$$q_1 + s = \lambda_{\vec{q}}s\,s_{\perp}\,q_2 \tag{5.18b}$$

$$q_2 s = \lambda_{\vec{q}}s_{\perp}(s - q_1) \tag{5.18c}$$

$$q_2 = \lambda_{\vec{q}}s_{\perp}(1 - sq_1). \tag{5.18d}$$

This set of four equations for three variables q_1, q_2 and $\lambda_{\vec{q}}$ is non-condradictory when two of these equations are linearly dependent. When the state is pure ($s = 1$), the last two equations (5.18c) and (5.18d) are equivalent and the solution is $q_1^2 + q_2^2 = 1$. Thus for pure states there is a continuous set of optimal POVMs (5.7) parametrized by a vector $\vec{q}$ which lies on a circle of unit radius. If the state is mixed ($s < 1$), the last two equations are non-contradictory only if

$q_1 = 0$ and the other two equations give $q_2 = \pm 1$. This is a dramatic difference when compared to the pure state case – a continuous set of POVMs reduces to just two possibile projection operators.

Finally, we point that when the logarythmic derivative is known, some subset of optimal measurements is given by the projection operators onto the eigenstates of $\hat{\mathcal{L}}_{\hat{\rho}}$. In the case of a single qubit, this procedure gives only the $q_1 = 0$, $q_2 = \pm 1$ projectors, even for pure states.

5.1.4 Estimation from the population imbalance

In an N-qubit MZI, usually the phase is estimated from the population imbalance between the two arms of the interferometer. Here we show under which conditions this estimation strategy is optimal in the simplest one-qubit case.

The population imbalance POVMs are two operators

$$\hat{E}_+ = |+\rangle\langle+| \quad \text{and} \quad \hat{E}_- = |-\rangle\langle-|. \tag{5.19}$$

which project the state $\hat{\rho}$ on either of the two modes. The corresponding probabilities of detecting the qubit in one of the arms are

$$p_\pm = \text{Tr}\left[\hat{\rho}\,\hat{E}_\pm\right] = \frac{1}{2}(1 \pm s_z) \tag{5.20}$$

According to Eq. (5.11) and using the θ-dependence of the vector $\vec{s}$ from Eq. (5.6), we obtain that the CFI for the estimation from the population imbalance is

$$F_{\text{imb}} = \frac{1}{p_+}\left(\frac{\partial p_+}{\partial \theta}\right)^2 + \frac{1}{p_-}\left(\frac{\partial p_-}{\partial \theta}\right)^2 = \frac{s_x^2}{1 - s_z^2}. \tag{5.21}$$

This CFI saturates the bound of the QFI from Eq. (5.13) when either (a) $s_x^2 + s_z^2 = 1$ or (b) $s_z = 0$. When (a) is true, the state is pure, and the CFI does not depend on θ. In other words, for any initial pure state $\hat{\rho}_{\text{in}}$ lying in the $x - z$ plane, the population imbalance measurement is optimal for the phase estimation. When the state is mixed only (b) can be true and then for every θ there is only one orientation of the initial state $\hat{\rho}_{\text{in}}$, which gives $s_z = 0$.

This once again shows how the optimal estimation strategies change abruptly as soon as the state is mixed. While there is a continuum of pure states, which when used for the phase estimation from the population imbalance give maximal value of the CFI, there is only one such mixed state for each θ.

5.2 Two qubits

In this section, we extend the analysis of the optimal estimation strategies to two spin-$\frac{1}{2}$ bosons. A general density matrix of such system with its eight independent parameters, it too difficult to investigate. However there is a set

of "Werner" states, which are described with a single real coefficient α. These states have a particularly simple form

$$\hat{\rho}_{\mathrm{w}} = \frac{1-\alpha}{3}\hat{\mathbb{1}} + \alpha\,\hat{\Pi}_{\mathrm{TF}}, \tag{5.22}$$

where $\hat{\Pi}_{\mathrm{TF}} = \left|1,1\right\rangle\!\left\langle 1,1\right|$ is a projection over the Twin-Fock state. When α varies from 0 to 1, $\hat{\rho}$ changes from a complete mixture, which is useless for the parameter estimation, to a strongly entangled pure Twin-Fock state. Werner states are a narrow subset of all possible two spin-$\frac{1}{2}$ bosonic states, nevertheless – as we show below – they provide valuable insight into the optimal estimation strategies in quantum metrology.

In analogy to the previous section, we use a generic linear interferometric transformation

$$\hat{U}(\theta) = \exp\left[-i\theta\,\vec{n}\cdot\vec{\hat{J}}\right]. \tag{5.23}$$

The "composite" angular momentum operators are a sum of corresponding single-particle Pauli matrices, i.e. $\hat{J}_i = \frac{1}{2}\hat{\sigma}_i^{(1)} + \frac{1}{2}\hat{\sigma}_i^{(2)}$, where $i = x, y, z$ and the upper index labels the particles.

First, we calculate the QFI, using the expression from Eq. (5.12). The Werner state written in the mode occupation basis is already diagonal, i.e.

$$\hat{\rho}_{\mathrm{w}} = \begin{pmatrix} \frac{1-\alpha}{3} & 0 & 0 \\ 0 & \frac{1+2\alpha}{3} & 0 \\ 0 & 0 & \frac{1-\alpha}{3} \end{pmatrix} \tag{5.24}$$

Since the generator of the phase transformation from Eq. (5.23) is $\hat{h} = \vec{n}\cdot\vec{\hat{J}}$, evaluation of the QFI is straightforward and we obtain

$$F_Q(\alpha) = \frac{12\alpha^2}{2+\alpha}\left(n_x^2 + n_y^2\right) \tag{5.25}$$

Note that the z-component of the generator does not contribute to the QFI, because $\hat{\rho}_{\mathrm{w}}$ is invariant upon rotation around the z axis. Therefore, it is reasonable to consider only such transformations, which lie in the x-y plane. In this case, the value of the Fisher information depends only on the length of the vector $\vec{n}$. Thus, without any loss of generality in the remaining of this Section we can restrict to the MZI transformation $\hat{U}(\theta) = \exp\left[-i\theta\hat{J}_y\right]$. If so, the QFI is simply equal to

$$F_Q(\alpha) = \frac{12\alpha^2}{2+\alpha}. \tag{5.26}$$

As anticipated at the beginning of this section, $F_Q(0) = 0$. The Werner state is usefully entangled, when $F_Q(\alpha) \geqslant 2$, which gives $\alpha \geqslant \frac{2}{3}$. For a Twin-Fock state we obtain the Heisenberg scaling, i.e. $F_Q(1) = 4$.

5.2.1 Optimal measurements

In the next step, we find the optimal measurements, which saturate the bound set by the QFI (5.26). To this end, we determine the logarithmic derivative using Eq. (5.15). A simple calculation gives that

$$\hat{\mathcal{L}}_{\hat{\rho}} = -\frac{6i}{2+\alpha}\left[\hat{J}_y, \hat{\rho}_{\mathrm{w}}(\theta)\right], \tag{5.27}$$

where $\hat{\rho}_{\mathrm{w}}(\theta) = e^{-i\theta\hat{J}_y}\hat{\rho}_{\mathrm{w}}(\theta)e^{i\theta\hat{J}_y}$. In analogy to the single-qubit case, we should now parametrize the POVMs similarly as in Eq. (5.7) and find the parameters from the condition Eq. (5.14). However, this procedure gives equations, which cannot be solved in a simple way. Therefore, we restrict to those optimal POVMs, which can be found by the diagonalization of the logarithmic derivative. We write down $\hat{\mathcal{L}}_{\hat{\rho}}$ in the matrix form

$$\hat{\mathcal{L}}_{\hat{\rho}} = \frac{6\alpha}{\sqrt{2}(2+\alpha)}\begin{pmatrix} \frac{1}{\sqrt{2}}\sin 2\theta & -\cos 2\theta & -\frac{1}{\sqrt{2}}\sin 2\theta \\ -\cos 2\theta & -\sqrt{2}\sin 2\theta & \cos 2\theta \\ -\frac{1}{\sqrt{2}}\sin 2\theta & \cos 2\theta & \frac{1}{\sqrt{2}}\sin 2\theta \end{pmatrix}$$

and obtain a set of three eigen-states

$$|\Psi_1\rangle = \frac{(\cos\theta - \sin\theta)}{\sqrt{2}}|\psi_-\rangle - \frac{(\cos\theta + \sin\theta)}{\sqrt{2}}|1,1\rangle \tag{5.28a}$$

$$|\Psi_2\rangle = \frac{(\cos\theta + \sin\theta)}{\sqrt{2}}|\psi_-\rangle + \frac{(\cos\theta - \sin\theta)}{\sqrt{2}}|1,1\rangle \tag{5.28b}$$

$$|\Psi_3\rangle = |\psi_+\rangle, \tag{5.28c}$$

where $|\psi_\pm\rangle = \frac{|2,0\rangle \pm |0,2\rangle}{\sqrt{2}}$. The optimal measurements depend on θ, have a complicated form and it is difficult to tell how to realize them in the laboratory. Note however, that when $\theta = 0$ a following transformation

$$\hat{V} = \exp\left(i\frac{\pi}{2}\frac{\hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x}{2}\right)\exp\left(i\frac{\pi}{4}\hat{J}_y\right). \tag{5.29}$$

is applied to Equations (5.28), it results in $\hat{V}|\Psi_1\rangle = |0,2\rangle$, $\hat{V}|\Psi_2\rangle = |1,1\rangle$ and $\hat{V}|\Psi_3\rangle = |2,0\rangle$. In this way, we obtain the eigen-states of the $\hat{J}_z$ operator, and the optimal measurement is based on the simple determination of the population imbalance. Nevertheless, to accomplish this we needed an additional operation (5.29) on the state. This transformation is non-local – it correlates the particles, since the product of two angular-momentum operators cannot be written as a sum of operators acting on each qubit independently.

5.2.2 Estimation from the population imbalance

We now consider a common estimation strategy based on the measurement of the imbalance of the population of the two modes. The CFI defined in Eq. (5.11)

is a sum of three terms,

$$F_{\text{imb}} = \sum_{n=0}^{2} \frac{1}{p(n|\theta)} \left(\frac{\partial p(n|\theta)}{\partial \theta} \right)^2, \tag{5.30}$$

where $p(n|\theta)$ is a probability for finding n particles in one of the modes and $2 - n$ in the other and is given by

$$p(n|\theta) = \text{Tr}\left[|n, N - n\rangle\langle n, N - n| \hat{\rho}_{\text{w}}(\theta) \right]. \tag{5.31}$$

Using the density matrix of the Werner states from Eq. (5.22) we obtain that $p(0|\theta) = p(2|\theta) = \frac{1-\alpha}{3} + \frac{\alpha}{2} \sin^2 \theta$ and $p(1|\theta) = \frac{1-\alpha}{3} + \alpha \cos^2 \theta$, which when put into (5.30) gives

$$F_{\text{imb}} = \frac{36\,\alpha^2\,\sin^2(2\theta)}{[4 - \alpha(1 + 3\cos(2\theta))][2 + \alpha(1 + 3\cos(2\theta))]}. \tag{5.32}$$

Only for $\alpha = 1$, when the Werner state is pure, $F_{\text{imb}} = 4$, so it does not depend on θ and saturates the bound of the QFI. As shown in Fig. 5.1, for other values of α the estimation from the population imbalance is non-optimal for all values of θ.
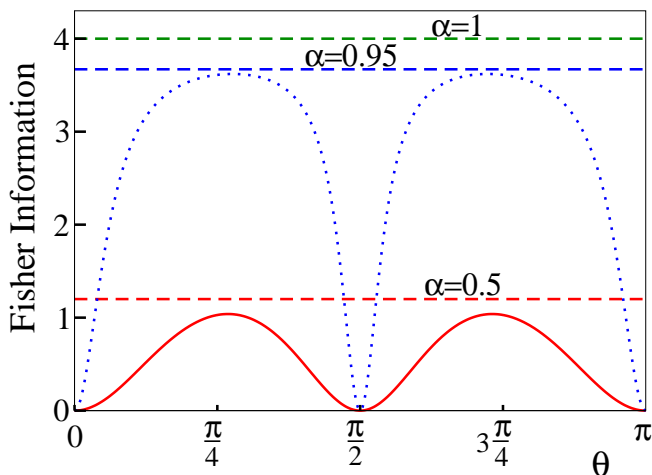

Figure 5.1: The CFI given by Eq. (5.32) for $\alpha = 0.5$ (solid red), $\alpha = 0.95$ (dotted blue) and $\alpha = 1$ (dashed green) as a function of θ. The dashed lines denote the corresponding values of the QFI (Eq. (5.26)), which for $\alpha = 1$ coincides with the CFI.

This is another example, after the single qubit case, of how the estimation strategy, which is optimal for a pure state, immediately deteriorates as soon as the state becomes mixed.

5.3 N qubits – pure states

So far, we have identified the optimal measurements for one and two qubits. It is natural to generalize these results and ask which are the optimal measurements for N qubits undergoing a linear interferometric transformation. However, the methods used in the previous two sections, which employed the logarithmic derivative $\hat{\mathcal{L}}_{\hat{\rho}}$, cannot be extended to higher N. This is because the mere analytical determination of $\hat{\mathcal{L}}_{\hat{\rho}}$ becomes impossible. Nevertheless, as we show below, for pure N-qubit states, some optimal estimation strategies can be found using the notion of the statistical distance.

5.3.1 QFI and the statistical distance

In [6] it is shown how the QFI from Eq. (5.12) is related to the statistical distance between two neighboring states. Although this general result is valid whenever a parameter is estimated from measurements performed on a θ-dependent state $|\psi\rangle$, below we present this formalism in context of quantum interferometry.
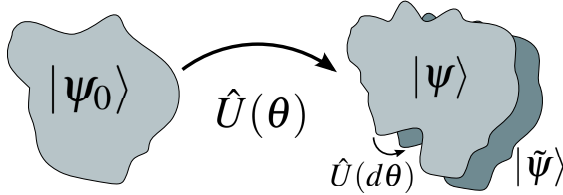

Figure 5.2: Schematic representation of the two steps performed to calculate the QFI. The input state $|\psi_0\rangle$, which enters the interferometer, is transformed by a unitary evolution operator $\hat{U}(\theta)$, giving the output state $|\psi\rangle$. To calculate the speed at which the state changes, and thus the statistical distance, we make a further infinitesimal rotation $\hat{U}(\theta)$ to obtain $\left|\tilde{\psi}\right\rangle$.

To this end, we follow the scheme pictured in Fig. 5.2. First, consider a pure two-mode *input* state $|\psi_0\rangle$, which is transformed by a unitary evolution operator $\hat{U}(\theta) = e^{-i\theta \hat{J}_k}$, where $k = x, y, z$. As a result, we obtain a θ-dependent *output* state, which when expanded in the basis of mode occupations reads

$$|\psi\rangle = e^{-i\theta \hat{J}_k} |\psi_0\rangle = \sum_{j=0}^{N} \sqrt{p_j} e^{i\varphi_j} |j\rangle \equiv \sum_{j=0}^{N} C_j |j\rangle. \tag{5.33}$$

Here $|j\rangle$ is a Fock state with j particles in the left and $N - j$ in the right arm. The neighboring state is found by applying a further infinitesimal transformation

$e^{-id\theta \hat{J}_k}$, which gives

$$\left| \tilde{\psi} \right\rangle \;\; = \;\; e^{-id\theta \hat{J}_k} \left| \psi \right\rangle \simeq \sum_{j=0}^{N} (1 - id\theta \hat{J}_k) C_j \left| j \right\rangle \tag{5.34a}$$

$$= \;\; \sum_{j=0}^{N} (1 - id\theta \eta_j^{(k)}) C_j \left| j \right\rangle \equiv \sum_{j=0}^{N} \tilde{C}_j \left| j \right\rangle , \tag{5.34b}$$

The coefficient $\eta_j^{(k)}$ is a result of acting with $\hat{J}_k$ on a ket $\left| j \right\rangle$ and is equal to

$$\eta_j^{(x)} \;\; = \;\; \frac{1}{2} \frac{\alpha_j C_{j+1} + \alpha_{j-1} C_{j-1}}{C_j} \tag{5.35a}$$

$$\eta_j^{(y)} \;\; = \;\; \frac{1}{2i} \frac{\alpha_j C_{j+1} - \alpha_{j-1} C_{j-1}}{C_j} \tag{5.35b}$$

$$\eta_j^{(z)} \;\; = \;\; j - \frac{N}{2}, \tag{5.35c}$$

where $\alpha_j = \sqrt{(j+1)(N-j)}$. The state (5.34b) can be alternatively written as

$$\left| \tilde{\psi} \right\rangle = \sum_{j=0}^{N} \sqrt{p_j + dp_j} \, e^{i(\varphi_j + d\varphi_j)} \left| j \right\rangle . \tag{5.36}$$

where the probability- and phase-increments are

$$dp_j = |\tilde{C}_j|^2 - |C_j|^2 = 2 \operatorname{Im} \eta_j^{(k)} |C_j|^2 d\theta \tag{5.37a}$$

$$e^{id\varphi_j} = \frac{\tilde{C}_j}{|\tilde{C}_j|} \frac{|C_j|}{C_j} = e^{-i \operatorname{Re} \eta_j^{(k)} d\theta} . \tag{5.37b}$$

The distance between two neighboring states is equal to

$$ds_{\text{ps}}^2 = 1 - |\langle \psi | \tilde{\psi} \rangle|^2 \tag{5.38a}$$

$$= \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4 \left[\sum_{j=0}^{N} p_j d\varphi_j^2 - \left(\sum_{j=0}^{N} p_j d\varphi_j \right)^2 \right] \tag{5.38b}$$

$$\equiv \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4\Delta^2 d\varphi. \tag{5.38c}$$

Finally, the QFI can be interpreted as the speed at which the state changes upon the infinitesimal increment of the parameter θ and therefore it reads

$$F_Q = \frac{ds_{\text{ps}}^2}{d\theta^2} = 4 \sum_{j=0}^{N} |C_j|^2 \left(\operatorname{Im} \eta_j^{(k)} \right)^2 + \tag{5.39a}$$

$$4 \sum_{j=0}^{N} |C_j|^2 \left(\operatorname{Re} \eta_j^{(k)} \right)^2 - 4 \left(\sum_{j=0}^{N} |C_j|^2 \operatorname{Re} \eta_j^{(k)} \right)^2 . \tag{5.39b}$$

When the phase is acquired through a unitary transformation like in Eq. (5.33), the above expression simplifies to the well known result $F_Q = 4\Delta^2 \hat{J}_k$. However, as we argue below, for the purpose of finding the optimal measurements, it is more convenient to keep the QFI in the form of Equations (5.39), i.e. as a sum of two non-negative parts – the change of the probability p_j and the variance of the phase increment $d\varphi_j$. Namely, if the latter term is zero, then the CFI calculated using the probability $p_j = |\langle j|\psi\rangle|^2$ of finding the system in state $|j\rangle$ is equal to QFI. This means that no information about θ is carried by the phases φ_j, which by definition are not witnessed by the projection measurement $|j\rangle\langle j|$ and thus do not contribute to probabilities p_j.

5.3.2 "In-situ" measurements – localized modes

In this section, by referring to Eq. (5.39), we identify two optimal measurements performed "in-situ", when the particles remain trapped in the two arms of the interferometer and their mode functions do not overlap.

Estimation from the full correlation

As a first example, we consider the phase estimation from the full N-body probability

$$
\begin{aligned}
p_N(\mathbf{r}|\theta) &= \frac{1}{N!}\langle\psi|\hat{\Psi}^\dagger(x_1)\ldots\hat{\Psi}^\dagger(x_N)\hat{\Psi}(x_N)\ldots\hat{\Psi}(x_1)|\psi\rangle \\
&\equiv \langle\psi|\hat{G}(\mathbf{r})|\psi\rangle .
\end{aligned}
\tag{5.40}
$$

of finding particles at positions $\mathbf{r} = (x_1\ldots x_N)$. The two-mode field operator is $\hat{\Psi}(x) = \psi_a(x)\hat{a} + \psi_b(x)\hat{b}$ and the wave-packets are separated in two arms of the interferometer, for instance by imposing $\psi_a(x) = 0$ for $x < 0$ and $\psi_b(x) = 0$ for $x > 0$. The θ-dependence of the probability $p_N(\mathbf{r}|\theta)$ comes from the state $|\psi\rangle$ from Eq. (5.33), which is used to calculate the average value of the operator $\hat{G}(\mathbf{r})$.

The estimation sequence relies upon detecting positions of N atoms in $m \gg 1$ experiments. If the phase is then deduced using the maximum likelihood estimator, then according to the Fisher theorem, its sensitivity is given by

$$
\Delta^2\theta = \frac{1}{m}\frac{1}{F_N},
\tag{5.41}
$$

where F_N is the CFI which is equal to

$$
F_N = \int d\mathbf{r}\,\frac{1}{p_N(\mathbf{r}|\theta)}\left(\frac{\partial p_N(\mathbf{r}|\theta)}{\partial\theta}\right)^2 .
\tag{5.42}
$$

In order to calculate F_N we first evaluate the derivative of the probability (5.40),

$$
\partial_\theta p_N(\mathbf{r}|\theta) = i\langle\psi|\hat{J}_k\hat{G}(\mathbf{r})|\psi\rangle - i\langle\psi|\hat{G}(\mathbf{r})\hat{J}_k|\psi\rangle =
\tag{5.43}
$$

$$
2\,\mathrm{Im}\,\langle\psi|\hat{G}(\mathbf{r})\hat{J}_k|\psi\rangle = 2\,\mathrm{Im}\sum_{j,j'=0}^{N} C_j^* C_{j'}\eta_{j'}^{(k)}\langle j|\hat{G}(\mathbf{r})|j'\rangle .
$$

The CFI is therefore equal to

$$F_N = 4 \int d\mathbf{r} \ \frac{\left[\mathrm{Im} \sum\limits_{j,j'=0}^{N} C_j^* C_{j'} \eta_{j'}^{(k)} \langle j | \hat{G}(\mathbf{r}) | j' \rangle \right]^2}{\mathrm{Im} \sum\limits_{j,j'=0}^{N} C_j^* C_{j'} \langle j | \hat{G}(\mathbf{r}) | j' \rangle}. \tag{5.44}$$

We now define Ω_μ by saying that $\mathbf{r} \in \Omega_\mu$ when $x_1 \ldots x_\mu < 0$ and $x_{\mu+1} \ldots x_N > 0$. Using this definition we obtain

$$F_N = 4 \sum_{\mu=0}^{N} \binom{N}{\mu} \int_{\mathbf{r} \in \Omega_\mu} d\mathbf{r} \ \frac{\left[\mathrm{Im} \sum\limits_{j,j'=0}^{N} C_j^* C_{j'} \eta_{j'}^{(k)} \langle j | \hat{G}(\mathbf{r}) | j' \rangle \right]^2}{\mathrm{Im} \sum\limits_{j,j'=0}^{N} C_j^* C_{j'} \langle j | \hat{G}(\mathbf{r}) | j' \rangle} \tag{5.45}$$

where the combinatory factor is due to indistiguishability of particles and stands for all possible choices of μ particles out of a set of N. When $\mathbf{r} \in \Omega_\mu$, then for separated wave-packets $\hat{G}(\mathbf{r}) | n \rangle \propto \frac{\mu!(N-\mu)!}{N!} | n \rangle \delta_{n\mu}$ and the above integral gives

$$F_N = 4 \sum_{j=0}^{N} |C_j|^2 \left(\mathrm{Im}\, \eta_j^{(k)} \right)^2. \tag{5.46}$$

We notice that this expression is equal to the first line of the QFI, see (5.39a).

Therefore, estimation from the N-body probability of trapped particles is optimal only if the other terms in line (5.39b) vanish, which requires $\mathrm{Re}\, \eta_j^{(k)} \equiv 0$ for all j. According to Equations (5.35), this condition can be satisfied only for the rotations around x and y axes. In the other case, $\mathrm{Im}\, \eta_j^{(z)} = 0$ and thus $F_N = 0$, because the simple phase imprint $e^{-i\theta J_z}$ requires further mode mixing to provide information about θ.

For the rotation around x-axis, $\mathrm{Re}\, \eta_j^{(x)} = 0$ if $C_j = i^j a_j$, while for the y-axis the condition is $C_j = e^{i\phi} a_j$, where $a_j \in \mathbb{R}$ and ϕ is a common phase. In particular for $\phi = 0$, the measurement is optimal, when all C_j's are real. Since the elements of the Wigner rotation matrix – which transforms the input state $|\psi_0\rangle$ into the output state $|\psi\rangle$ – are all real [8], we conclude that if the input state of the MZI has real coefficients, the estimation from p_N is optimal.

Estimation from the population imbalance

Although phase estimation from the N-body probability is optimal for a wide class of states and rotations around x and y, it has one major flaw – it is unpractical, since it requires sampling of a vast configurational space. We now show, that the same value of the CFI as in Eq. (5.46) can be obtained, when the phase is estimated from a simple population imbalance measurement.

The probability of having j atoms in the mode a and $N - j$ in b is

$$p(j|\theta) = |\langle j|\psi\rangle|^2 = |C_j|^2. \tag{5.47}$$

Similarly as in Eq. (5.43), its derivative reads

$$\partial_\theta p(j|\theta) = 2|C_j|^2 \mathrm{Im}\, \eta_j^{(k)}. \tag{5.48}$$

Therefore, the CFI calculated with (5.47)

$$F_{\mathrm{imb}} = \sum_{j=0}^N \frac{1}{p(j|\theta)} \left(\frac{\partial\, p(j|\theta)}{\partial\theta} \right)^2 = 4 \sum_{j=0}^N |C_j|^2 \left(\mathrm{Im}\, \eta_j^{(k)} \right)^2 \tag{5.49}$$

is equal to (5.46). In consequence, the QFI from Eq. (5.39) is saturated with the same family of states for the x and y rotations as in the case of the estimation from $p_N(\mathbf{r}|\theta)$. This is a step forward with respect to the work of Hofmann [9], where the saturation of the QFI bound with the population imbalance measurement was reported for the MZI and symmetric states with $C_j = C_{N-j}$.

5.3.3   Measurement after expansion

As argued above, when the interferometer rotates the state around the z-axis, giving a sole phase-imprint, further manipulation is necessary to exchange the information about the phase between the two modes. Here we assume, that this operation is realized by letting the two mode functions $\psi_a(x)$ and $\psi_b(x)$ expand and form an interference pattern. In such situation, the two modes cannot be distinguished anymore, and it is not possible to define a proper population imbalance operator. Instead, one must estimate θ in some different way. For instance, estimation from the least-squares fit of the one-body probability $p_1(x|\theta) = \frac{1}{N}\langle\psi|\hat{\Psi}^\dagger(x)\hat{\Psi}(x)|\psi\rangle$ to the interference pattern, although gives sub shot-noise sensitivity when the input state $|\psi_0\rangle$ is phase-squeezed [17], is never optimal [18].

Nevertheless, the optimal measurement can be identified and it is the N-body CFI from Eq. (5.42) which saturates the bound of the QFI under following additional assumptions [19]. First, the information between the two modes must be fully exchanged. This means, that the envelopes of $\psi_a(x)$ and $\psi_b(x)$ fully overlap and the functions only differ by the phase. This is true if initially $\psi_a(x)$ and $\psi_b(x)$ are of the same shape but are separated in space and then expand to reach the far-field regime. Another requirement is that the coefficients of the initial state

$$|\psi_0\rangle = \sum_{j=0}^N C_j^{(0)} |j\rangle \tag{5.50}$$

are real and posses the symmetry $C_j^{(0)} = C_{N-j}^{(0)}$. States having these properties naturally appear in the context of quantum interferometry with ultra-cold gas trapped in the double-well potential. Namely, the ground state of the symmetric

two-mode Bose-Hubbard Hamiltonian for every ratio of the interaction strength U to the tunneling rate J has real and symmetric coefficients $C_j^{(0)}$.

According to Eq. (5.33), the rotation around the z axis transforms the state into

$$|\psi\rangle = e^{-i\theta \hat{J}_z} |\psi_0\rangle = \sum_{j=0}^{N} C_j |j\rangle, \qquad (5.51)$$

where $C_j = C_j^{(0)} e^{-i\theta\left(j-\frac{N}{2}\right)}$. As argued in detial in [19], the CFI from Eq. (5.42) can be calculated under the aforementioned assumptions and the outcome is

$$F = 4 \sum_{j} |C_j|^2 \eta_j^{(z)2} = 4\Delta^2 \hat{J}_z = F_Q, \qquad (5.52)$$

where $\eta_j^{(z)}$ was defined in Eq. (5.35c). This shows that the estimation from the N-th body correlation in the far field is optimal.

5.4   Estimation from population imbalance with mixed states

In the previous Section we have identified the set of pure states for which estimation from the population imbalance is optimal. Here we show, how various methods of estimation from the measured population-imbalance are affected by the presence of noise during the interferometric sequence.

To this end, we focus on the MZI transformation as in Eq. (5.3). For simplicity, we assume that the interferometer is fed with a pure state

$$|\psi_{\text{in}}\rangle = \sum_{n=0}^{N} C_n |n, N-n\rangle. \qquad (5.53)$$

We now incorporate the noise in the MZI transformation by assuming that the imprinted phase fluctuates from shot to shot around the true value θ. In consequence, the output state is not pure anymore and reads

$$\hat{\rho} = \int_{-\pi}^{\pi} d\varphi \, P(\varphi) \, |\psi_\varphi\rangle \langle\psi_\varphi|, \qquad (5.54)$$

where $|\psi(\varphi)\rangle = e^{-i(\theta+\varphi)\hat{J}_y} |\psi_{\text{in}}\rangle$ is a pure state after a "single" transformation. We assume that the probability that the acquired phase differs from θ by φ is Gaussian, i.e.

$$P(\varphi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\varphi^2}{2\sigma^2}\right]. \qquad (5.55)$$

It is normalized when $\sigma \ll 2\pi$, which is a reasonable assumption in the context of phase estimation.

We focus on states $|\psi_{\text{in}}\rangle$, which are the usefully entangled for the $\hat{J}_y$ interferometer. These are the spin-squeezed states [10, 11], where the squeezing is
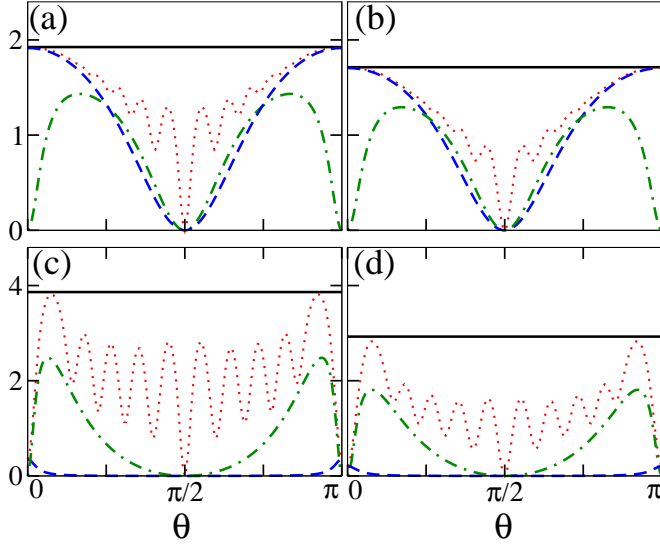
Figure 5.3: "Fisher information" normalized to the number of particles with $N = 10$. Solid black line – QFI, dotted red line – CFI, dashed blue line – spin squeezing, dot-dashed green line – second moment. (a) $w = 0.4, \sigma = 0.09$, (b) $w = 0.4, \sigma = 0.12$, (c) $w = 0.036, \sigma = 0.09$, (d) $w = 0.036, \sigma = 0.12$,

related to the reduced fluctuations of the population imbalance between the two modes. Such entangled multi-particle states have been recently generated in a number of experiments [12, 13, 14, 15, 16]. The spin-squeezed states can be modeled by a Gaussian as follows

$$C_n = \frac{\sqrt{2}}{\sqrt{\pi w N}} \exp\left[-\frac{(n - \frac{N}{2})^2}{w\,N} \right]. \tag{5.56}$$

In such case, the spin-squeezing parameter is equal to

$$\xi_s^2 = N\frac{\langle \hat{J}_z^2 \rangle_{\mathrm{in}}}{\langle \hat{J}_x \rangle_{\mathrm{in}}^2}, \tag{5.57}$$

where the average values are calculated with $|\psi_{\mathrm{in}}\rangle$ and using coefficients (5.56). When $w < 1$, we obtain $\xi_s^2 < 1$ so the state is spin-squeezed and entangled.

To picture what is the relation between spin-squeezing and the MZI, let us for a moment consider a noise-less case, when the output state is $e^{-i\theta\hat{J}_y} |\psi_{\mathrm{in}}\rangle$. In such scenario, if the phase is estimated from the average population imbalance, then the sensitivity is given by the error propagation formula

$$\Delta^2\theta = \frac{1}{m}\frac{\Delta^2\hat{J}_z}{\left|\frac{\partial\langle\hat{J}_z\rangle}{\partial\theta}\right|^2} \geqslant \frac{1}{m}\frac{\xi_s^2}{N}. \tag{5.58}$$

Here, the variance and the average value of the $\hat{J}_z$ operator are calculated with the θ-dependent state $e^{-i\theta\hat{J}_y}|\psi_{\text{in}}\rangle$ and the above inequality is saturated only around $\theta = 0$. Therefore, in the absence of noise, spin-squeezed input states are usefully entangled.
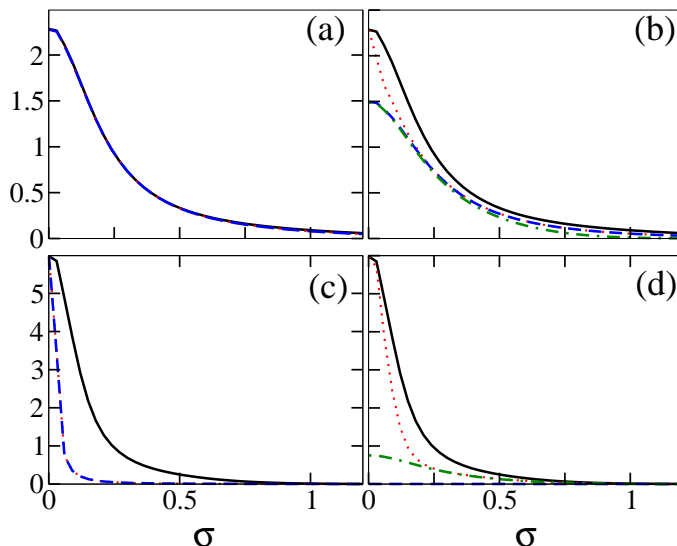

Figure 5.4: "Fisher information" normalized to the number of particles with $N = 10$. Solid black line – QFI, dotted red line – CFI, dashed blue line – spin squeezing, dot-dashed green line – second moment. (a) $w = 0.4, \theta = 0$, (b) $w = 0.4, \theta = \pi/4$, (c) $w = 0.036, \theta = 0$, (d) $w = 0.036, \theta = \pi/4$.

It is interesting to check, if these states are still useful, when the interferometer is noisy. To this end, we proceed as follows. We take $|\psi_{\text{in}}\rangle$ with various w and generate the output mixed states (5.54) for different values of the noise width σ and θ. We then numerically calculate the QFI and the CFI for the population imbalance measurement

$$F_{\text{imb}} = \sum_{n=0}^{N} \frac{1}{p(n|\theta)} \left(\frac{\partial p(n|\theta)}{\partial \theta} \right)^2, \tag{5.59}$$

where the probability of having n atoms in one mode and $N - n$ in the other is equal to

$$p(n|\theta) = \text{Tr}\left[\left[\hat{\rho}|n, N - n\rangle\langle n, N - n| \right] \right] \tag{5.60}$$

We compare these two results with the sensitivity obtained from the average population imbalance as in Eq. (5.58). Finally, we calculate the sensitivity of the estimation from the second moment of the population imbalance

$$\Delta^2\theta = \frac{1}{m} \frac{\Delta^2\left(\hat{J}_z^2\right)}{\left| \frac{\partial\langle\hat{J}_z^2\rangle}{\partial\theta} \right|^2}. \tag{5.61}$$

This last estimation method was recently used to prove the entanglement of the twin-Fock state generated in the experiment [7].
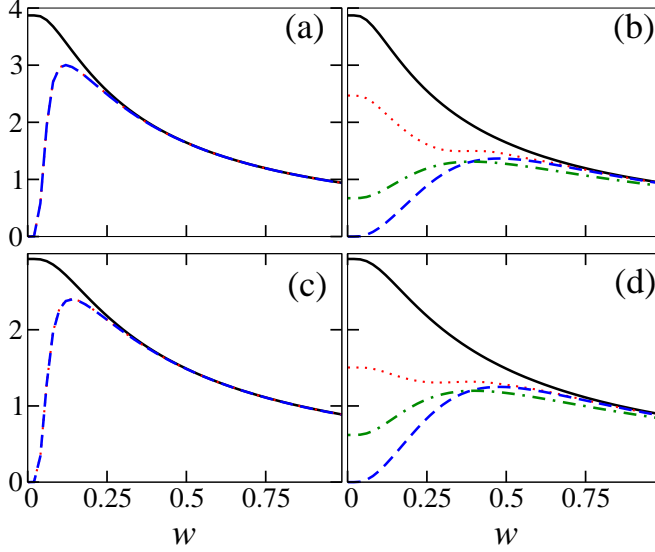

Figure 5.5: "Fisher information" normalized to the number of particles with $N = 10$. Solid black line – QFI, dotted red line – CFI, dashed blue line – spin squeezing, dot-dashed green line – second moment. (a) $\sigma = 0.09, \theta = 0$, (b) $\sigma = 0.09, \theta = \pi/4$, (c) $\sigma = 0.12, \theta = 0$, (d) $\sigma = 0.12, \theta = \pi/4$.

In all the following figures, we plot the inverse sensitivity squared, normalized to $m\,N$, i.e. $\frac{1}{\Delta^2\theta mN}$ with $N = 10$ particles. In case of the CFI and the QFI this gives the corresponding Fisher informations normalized to the shot-noise. In Fig. 5.3 we plot these four quantities as a function of θ for two different values of σ and w. The QFI does not depend on θ and exceeds the shot-noise limit in all cases. Note however, that even a small amount of noise ($\sigma = 0.09$) has a dramatic impact on other three sensitivities. When the input state is close to coherent ($w = 0.4$), the CFI is optimal only for small angles and develops some complicated oscillatory structure for other values of θ. The estimation from the average population imbalance works well also only for small angles and the sensitivity smoothly deteriorates as $\theta \to \frac{\pi}{2}$. To contrary, estimation from the second moment gives best sensitivity far from $\theta = 0$. The situation changes abruptly for a strongly squeezed state ($w = 0.036$). In this case, the CFI is not optimal at $\theta = 0$ and reaches a value close to the bound set by the QFI at some intermediate angles.

In Fig. 5.4 we plot the normalized inverse of the squared sensitivity as a function of the noise width σ for two values of θ and w. For $\theta = 0$ the estimation from the second moment gives no sensitivity whatsoever, so it is not displayed. We observe that for $\theta = 0$ and $w = 0.4$, both the CFI and the first moment saturate the QFI, which is consistent with Fig. 5.3. For other values of w the

CFI (and consequently other estimation methods) are far from optimal as soon as the noise enters the system.

In Fig. 5.5 we plot the normalized inverse of the squared sensitivity as a function of the squeezing of the initial state w for two values of θ and σ. We observe, that although the QFI gets better as w drops, the other sensitivities have some optimal squeezing point, after which either drop (estimation from two lowest moments for $\theta = 0$ and $\theta = \frac{\pi}{4}$ and the CFI for $\theta = 0$) or saturate (CFI for $\theta = \frac{\pi}{4}$).
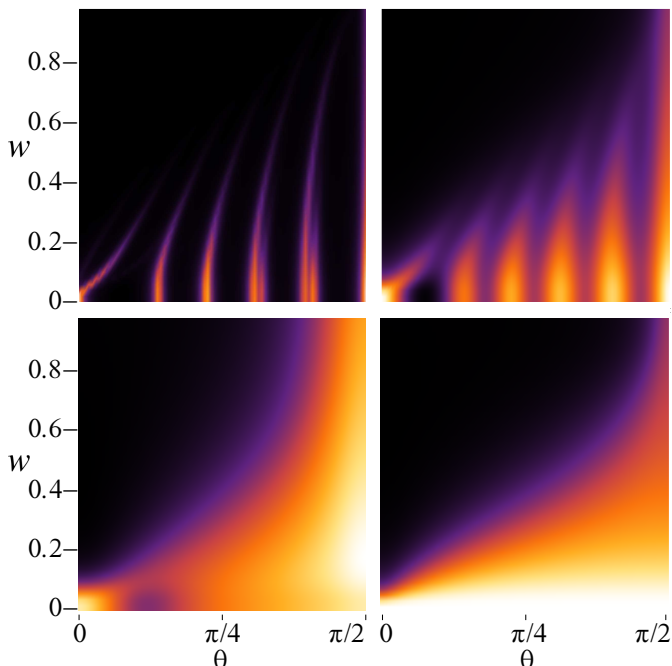

Figure 5.6: Density plots of the difference between the QFI and inverse sensitivities calculated with various estimation methods as a function of w and θ. The brighter the color, the bigger the difference is. The top left shows the difference between F_Q and the CFI calculated from the population imbalance F_{imb} from Eq. (5.59) with $\sigma = 0.03$. Top right and bottom left – the difference between the same two quantities, but calculated with $\sigma = 0.09$ and $\sigma = 0.27$ (correspondingly). Bottom right – the difference between the F_Q and the inverse sensitivity square obtain from the lowest moment of the population imbalance using Eq. (5.61) with $\sigma = 0.09$.

Finally in Fig. 5.6 we use the density plots to show how the inverse of the sensitivity calculated with various estimation protocols differ from the QFI. We display these results as a function of squeezing of the initial state w and the phase θ. The brighter the color is, the bigger the difference between the QFI and a given inverse sensitivity. The black color stands for a zero difference, and

thus denotes optimal estimation strategy.

The top left plot of Fig. 5.6 shows the difference between F_Q and the CFI calculated from the population imbalance F_{imb}) with $\sigma = 0.03$. Clearly, apart from small regions, the CFI is optimal for most phases and values of the squeezing w. When the amount of noise is increased to $\sigma = 0.09$, these region of optimality shrinks substantially (top right). For even larger $\sigma = 0.27$, the area where the CFI is not optimal anymore starts to dominate (bottom left). In the bottom right plot we show the analogous difference between the F_Q and the inverse sensitivity square obtain from the lowest moment of the population imbalance and for $\sigma = 0.09$. We see how this estimation method is outperformed by the full Fisher information calculated with the same amount of noise (top right).

Chapter 6

Entanglement with ultra-cold atoms

In this chapter we review some recent experiments, which demonstrate multi-particle entanglement in cold-atomic systems using entanglement measures borrowed from quantum interferometry. Before we focus on some particular results, we need however to understand, what is the way to describe such ultra-cold gas in a two-mode setup. Typically such bi-partite configuration is introduced by means of a double well potential, and we will now demonstrate that a state of a Bose-Einstein Condensate (BEC) trapped in this external field can be treated as an input to an interferometer (hopefully) operating below the shot noise limit.

To begin the detailed analyzis, note that when talking about BECs, we are dealing with a system of N idistinguishable bosons. Therefore, it is convenient to use the language of the second quantization, where the central object is the field operator $\hat{\Psi}(\mathbf{r}, t)$. This object, when acting on some state $|\psi\rangle$, anihilates a single particle located at position x. The key property of the field operator is the commutation relation, which in case of bosons reads

$$\left[\hat{\Psi}(\mathbf{r}, t), \hat{\Psi}^{\dagger}(\mathbf{r}', t)\right] = \delta^{(3)}(\mathbf{r} - \mathbf{r}'). \tag{6.1}$$

In order to create a complete quantum theory, we now need to express the 1st quantization Hamiltonian of N particles, which (say) interact via some two-body potential

$$H = \sum_{i=1}^{N} \left(-\frac{\hbar^2 \nabla_i^2}{2m} + V_1(\mathbf{r}_i)\right) + \frac{1}{2} \sum_{i \neq j} V_2(\mathbf{r}_i - \mathbf{r}_j) \tag{6.2}$$

in terms of the field operators. In turns out that the proper transition from the 1st to the 2nd quantization gives

$$\hat{H} = \int d\mathbf{r}\, \hat{\Psi}^{\dagger}(\mathbf{r}, t) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_1(\mathbf{r})\right) \hat{\Psi}(\mathbf{r}, t) \tag{6.3a}$$

$$+ \frac{1}{2} \iint d\mathbf{r}\, d\mathbf{r}'\, \hat{\Psi}^{\dagger}(\mathbf{r}, t) \hat{\Psi}^{\dagger}(\mathbf{r}', t) \hat{\Psi}(\mathbf{r}', t) \hat{\Psi}(\mathbf{r}, t) V_2(\mathbf{r} - \mathbf{r}'). \tag{6.3b}$$

This Hamiltonian together with the Heisenberg equation of motion

$$i\hbar\partial_t\hat{\Psi}(\mathbf{r},t) = \left[\hat{H},\hat{\Psi}(\mathbf{r},t)\right] \tag{6.4}$$

and the quantum state of the system $|\psi\rangle$ form a full quantum theory. For instance, a one-body density, which can be deduced from the symmetrized N-body wave-function $\Phi(\mathbf{r}_1\ldots\mathbf{r}_N;t)$ as follows

$$\rho(\mathbf{r};t) = N\int d\mathbf{r}_2\ldots d\mathbf{r}_N\,|\Phi(\mathbf{r}_1\ldots\mathbf{r}_N;t)|^2 \tag{6.5}$$

can be equivalently calculated in the second quantization as

$$\rho(\mathbf{r};t) = \langle\psi|\hat{\Psi}^\dagger(\mathbf{r},t)\hat{\Psi}(\mathbf{r},t)\,|\psi\rangle. \tag{6.6}$$

In an alagous way, one can reproduce all the moments of the symmetrized many-body wave function, thus prooving the equivalence of both the languages.

Our aim now is to employ this formalism to describe an ultra-cold gas in a double-well potential $V_{\mathrm{dw}}(\mathbf{r})$. One property of an ultra-cold gas is that atoms collide with very low mutual velocity. Therefore the whole two-body potential can be approximated with an effective contact interaction potential, i.e. $V_2(\mathbf{r} - \mathbf{r}') \simeq g\delta^{(3)}(\mathbf{r} - \mathbf{r}')$. Here, g is the strength of the coupling. The reason, why this approximation might be valid is that when the mutual energy of the colliding pair is low, they cannot probe the details of the interaction. All they can feel is some effective force, which in the crudest approximation is modelled by the delta potential. The only point, where the contact interaction is somehow linked with the true two-body force is through the coupling strength g, which is deduced from the full $V_2(\mathbf{r} - \mathbf{r}')$ in such a way that the effective potential reproduces the low-energetical scattering properties of the true force. It turns out, that the correct expression for the coupling constant is $g = 4\pi\hbar^2 a_s/m$, where a_s is the s-wave scattering length derived from the full force [20]. With this approximation, the Hamiltonain (6.3) with the one-body trapping potential $V_{\mathrm{dw}}(\mathbf{r})$ reads

$$\hat{H} = \int d\mathbf{r}\,\hat{\Psi}^\dagger(\mathbf{r},t)\left(-\frac{\hbar^2\nabla^2}{2m} + V_{\mathrm{dw}}(\mathbf{r})\right)\hat{\Psi}(\mathbf{r},t) \tag{6.7a}$$

$$+ \frac{g}{2}\int d\mathbf{r}\,\hat{\Psi}^\dagger(\mathbf{r},t)\hat{\Psi}^\dagger(\mathbf{r},t)\hat{\Psi}(\mathbf{r},t)\hat{\Psi}(\mathbf{r},t). \tag{6.7b}$$

Now, it is important to assume that the inter-well barrier of $V_{\mathrm{dw}}(\mathbf{r})$ is sufficiently steep, so that the states with lowest energy are localized in one of the wells. If this is the case, then the bosonic field operator can be written in a following form

$$\hat{\Psi}(\mathbf{r}) = \sum_i \psi_a^{(i)}(\mathbf{r})\hat{a}_i + \sum_i \psi_b^{(i)}(\mathbf{r})\hat{b}_i, \tag{6.8}$$

where $\psi_{a/b}^{(i)}(\mathbf{r})$ is a wave-packet localized in the left/right well with a corresponding anihilation operator $\hat{a}_i/\hat{b}_i$. When the trapped gas is very cold, one can in the

crudest approximaton assume, that only the lowest pair of modes is occupied, so the above field operator simplifies to

$$\hat{\Psi}(\mathbf{r}) = \psi_a(\mathbf{r})\hat{a} + \psi_b(\mathbf{r})\hat{b}. \tag{6.9}$$

This is what is called the two-mode approximation, and our point now is to derive the Hamiltonian for the operators $\hat{a}$ and $\hat{b}$. The open question remains what particular shape of the wave-packets $\psi_{a/b}(\mathbf{r})$. In the simplest approach, we will assume that these are the two lowest eigen-functions of the one-body part of the Hamiltonian. From elementary quantum mechanics we know, that the two lowest-lying states of a double-well Hamiltonian are $\psi_{\text{sym}}(\mathbf{r})$, which is symmetric with respect to reflection around the center of the trap, while the other $\psi_{\text{asym}}(\mathbf{r})$ is asymmetric. We construct the localized wave-packets as follows

$$\psi_{a/b}(\mathbf{r}) = \frac{\psi_{\text{s}}(\mathbf{r}) \pm \psi_{\text{asym}}(\mathbf{r})}{\sqrt{2}} \tag{6.10}$$

We will now substitute this decomposition into the Hamiltonian (6.7a) and evaluate the spatial integrals. First, let us concentrate on the one-body part. We have that

$$\hat{H}_1 = \int d\mathbf{r}\, \hat{\Psi}^{\dagger}(\mathbf{r}, t) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \hat{\Psi}(\mathbf{r}, t) = \tag{6.11a}$$

$$= \int d\mathbf{r}\, \psi_a^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \psi_a(\mathbf{r})\, \hat{a}^{\dagger}\hat{a} \tag{6.11b}$$

$$+ \int d\mathbf{r}\, \psi_b^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \psi_b(\mathbf{r})\, \hat{b}^{\dagger}\hat{b} \tag{6.11c}$$

$$+ \int d\mathbf{r}\, \psi_b^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \psi_a(\mathbf{r})\, \hat{b}^{\dagger}\hat{a} \tag{6.11d}$$

$$+ \int d\mathbf{r}\, \psi_a^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \psi_b(\mathbf{r})\, \hat{a}^{\dagger}\hat{b}. \tag{6.11e}$$

The first two terms, due to symmetry, are equal to $\left(\hat{a}^{\dagger}\hat{a} + \hat{b}^{\dagger}\hat{b}\right) E = \hat{N} \cdot E$. If we consider only the states with fixed number of atoms, $\hat{N}$ is a constant and therefore it can be safely removed from the Hamiltonian. The other two spatial integrals are equal, thanks to the symmetry between a and b. So let us analyze one of these integrals. It reads

$$\int d\mathbf{r}\, \psi_a^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \psi_b(\mathbf{r}) = \tag{6.12a}$$

$$\int d\mathbf{r}\, \frac{\psi_{\text{s}}^*(\mathbf{r}) + \psi_{\text{asym}}^*(\mathbf{r})}{\sqrt{2}} \left(-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{dw}}(\mathbf{r}) \right) \frac{\psi_{\text{s}}(\mathbf{r}) - \psi_{\text{asym}}(\mathbf{r})}{\sqrt{2}} \tag{6.12b}$$

$$= \frac{1}{2}(E_{\text{s}} - E_{\text{asym}}) = -\frac{1}{2}E_J, \tag{6.12c}$$

where $E_J > 0$ is called the Josephson energy. Therefore, we obtain that the one-body part of the Hamiltonian (6.3) reads

$$\hat{H}_1 = -\frac{1}{2}E_J(\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a}) = -E_j\hat{J}_x. \tag{6.13}$$

Now, we just need to identify the two-body part of the Hamiltonian. Note that since the wave-packets $\psi_{a/b}$ are separated in space (they are localized in the opposite wells of the potential), then after substituting the field operator in the form (6.9) into the two-body part, we can neglect the overlaps between the wave-packets and obtain that

$$\hat{H}_2 = \frac{g}{2}\int d\mathbf{r}\ \hat{\Psi}^\dagger(\mathbf{r},t)\hat{\Psi}^\dagger(\mathbf{r},t)\hat{\Psi}(\mathbf{r},t)\hat{\Psi}(\mathbf{r},t) \simeq \tag{6.14a}$$

$$\simeq \frac{g}{8}\int d\mathbf{r}\ \left(|\psi_a(\mathbf{r})|^2\hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + |\psi_b(\mathbf{r})|^2\hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{b}\right) \tag{6.14b}$$

$$= \frac{g}{8}\int d\mathbf{r}|\psi_a(\mathbf{r})|^2\left(\hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{b}\right). \tag{6.14c}$$

We now note, that this above expression can be written in terms of the $\hat{J}_z^2$ operator to give (after neglecting the constant terms proportional to N or N^2)

$$\hat{H}_2 = U\hat{J}_z^2, \quad \text{with} \quad U = \frac{g}{4}\int d\mathbf{r}|\psi_a(\mathbf{r})|^2 = \frac{g}{4}\int d\mathbf{r}|\psi_b(\mathbf{r})|^2. \tag{6.15}$$

This way we have derived what is called the two-mode Bose-Hubbard Hamiltonian, which reads

$$\hat{H} = -E_J\hat{J}_x + U\hat{J}_z^2. \tag{6.16}$$

In order to understand how useful entanglement is created in recent experiments with ultra-cold atoms, we will now focus on the family of ground states of this Hamiltonian. Since for whatever the values of the parameters E_J and U, the state of N particles in two modes can always be written as

$$|\psi\rangle = \sum_{n=0}^{N}\mathcal{C}_n|n, N-n\rangle. \tag{6.17}$$

Obviously, the amplitudes $\mathcal{C}_n$ fully determine the state. Note that the crucial quantity which characterizes the Hamiltonian (6.16) is actually the ratio between the two energies E_J and U. It is convinient to introduce a parameter $\alpha = NU/E_J$ and use the Hamiltonian in a form

$$\hat{H}_{\text{bh}} = -\hat{J}_x + \frac{\alpha}{N}\hat{J}_z^2. \tag{6.18}$$

The N in the denominator is introduced in order to accomodate for different spectra of $\hat{J}_x$ and $\hat{J}_z^2$. While the first one has a spectrum varying from $-\frac{N}{2}$ to $\frac{N}{2}$, the other naturally has a spectrum contained in the range from $-\frac{N^2}{4}$ to $\frac{N^2}{4}$.

Therefore this N in denominator somehow makes the two spectra "compatible". We are now prepared to analyze the ground states of the Hamiltonian (6.18) in terms of the coefficient α.

First, we will begin with the case of an non-interacting gas, having $g = 0$ and so $U = 0$ and in consequence $\alpha = 0$. In this situation, the Hamiltonian consists only of the $\hat{J}_x$ state and the ground state is particularly simple to identify. Namely, since the Hamiltonian is $-E_J \hat{J}_x$, and $E_J > 0$ then the ground state is an eigen-state of $\hat{J}_x$ with *maximal* eigen-value equal to $\frac{N}{2}$. Such state is referred to as the Coherent Spin-State and reads

$$|\psi\rangle_{\text{css}} = \frac{1}{\sqrt{N!}} \left(\frac{\hat{a}^\dagger + \hat{b}^\dagger}{\sqrt{2}} \right)^N |0,0\rangle. \tag{6.19}$$

This state is called "cohererent" because it is a product state of N particles in a coherent supperposition of being in the left and in the right well. Since it is a product, naturally the particles are not entangled. This state can also be written in a form (6.17) and by using the Newton binomial expansion we imediatelly obtain that

$$\mathcal{C}_n = \sqrt{\frac{1}{2^N} \binom{N}{n}}. \tag{6.20}$$

Since the probabilities $p_n = \mathcal{C}_n^2$ are given by the binomial distribution, we can naturally expect that the fluctuations of the population imbalance operator $\hat{J}_z$ are Poissonian. Indeed, for this state we obtain that

$$\left\langle \Delta^2 \hat{J}_z \right\rangle = \frac{N}{4}. \tag{6.21}$$

Also, for large N the binomial statistics can be reproduced with a gaussian. Therefore, as one can check, the probabilites p_n can be approximated with

$$p_n = |\mathcal{C}_n|^2 = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(n - \frac{N}{2}\right)^2}{2\sigma^2}}, \quad \text{with} \quad \sigma = \frac{\sqrt{N}}{2}. \tag{6.22}$$

Therefore, for the CSS, we have that $\sigma^2 = \left\langle \Delta^2 \hat{J}_z \right\rangle$. This is not a coincidence, because using this Gaussian model, one can calculate the variance of the $\hat{J}_z$ operator for any state as follows

$$\left\langle \Delta^2 \hat{J}_z \right\rangle \equiv \langle\psi|\hat{J}_z^2|\psi\rangle - \langle\psi|\hat{J}_z|\psi\rangle^2 = \sum_{n=0}^{N} p_n \left(n - \frac{N}{2} \right)^2 \tag{6.23a}$$

$$\simeq \int_{-\infty}^{\infty} p_n \left(n - \frac{N}{2} \right)^2 dn = \sigma^2. \tag{6.23b}$$

We are now ready to anticipate what other ground states of the Hamiltonian (6.18) with $\alpha > 0$ will look like. Recall, that the sign of α is related to the

kind of two-body interaction. Namely, when the coupling constant g is positive, than it means that the scattering length a_s is also positive. This sign of the scattering length is related to the *repulsive* two-body interactions between the particles. When the value of α is positive and grows, it means that the two-body
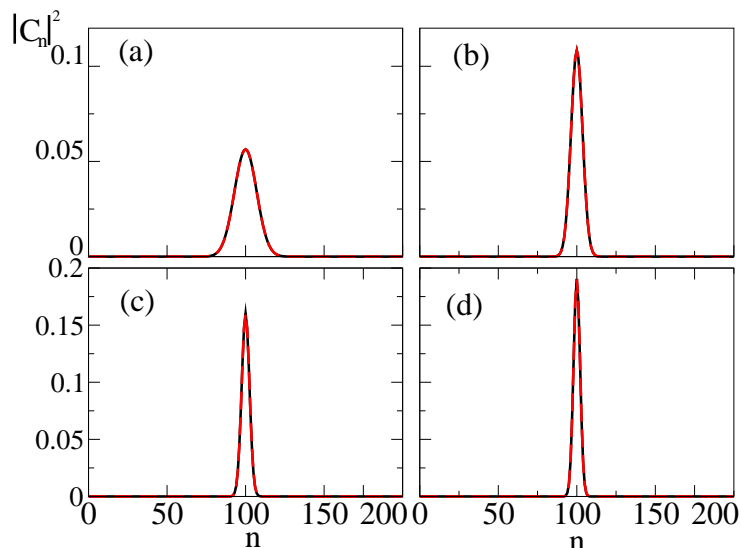

Figure 6.1: Coefficients $|\mathcal{C}_n|^2$ (solid black line) of the ground states of the Bose-Hubbard Hamiltonian (6.18) for different positive values of the parameter α equal to (a) $\alpha = 0$, (b) $\alpha = 25$, (c) $\alpha = 125$ and (d) $\alpha = 250$. The dashed red lines are the Gaussian fits, which fully overlap with the numerical results.

interactions become stronger and stronger and in particular when $\alpha \gg 1$, they dominate over the tunneling energy. In such case, the fluctuations of the atom number difference between the two sites, which are governed by the variance of the operator $\hat{J}_z$, should drop. This is because when the repulsion grows, it is energetically infavorable to have large number of particles on one of the sites. When the interaction tends to infinity, a ground state should be fully ballanced, i.e. have equal number of particles on each site and no fluctuations (which cost a lot of energy). Such state is called the Twin-Fock state and reads

$$|\psi\rangle_{\text{tf}} = \left| \frac{N}{2}, \frac{N}{2} \right\rangle. \tag{6.24}$$

This is an eigen-state of $\hat{J}_z$ with zero eigen-value, therefore in particular it has vanishing variance of the population imbalance operator, i.e.

$$\left\langle \Delta^2 \hat{J}_z \right\rangle_{\text{tf}} = 0. \tag{6.25}$$

Comparing this expression with Equations (6.21) and (6.23) it seems (if the Gaussian model is valid for the ground states of the Bose-Hubbard Hamiltonian

for all $\alpha > 0$) the that the width of the Gaussian has simply shrinked down to zero. To check if our understanding of the ground states is correct, in Fig. 6.1 we plot the coefficients $|\mathcal{C}_n|^2$ of the numerically calculated ground states of the Hamiltionian (6.18) for $N = 200$ particles and different values of the parameter α. Clearly, the coefficients can be well modelled by gaussian functions for all $\alpha > 0$. This analysis shows, that the ground states of the Hamiltonian (6.18) with
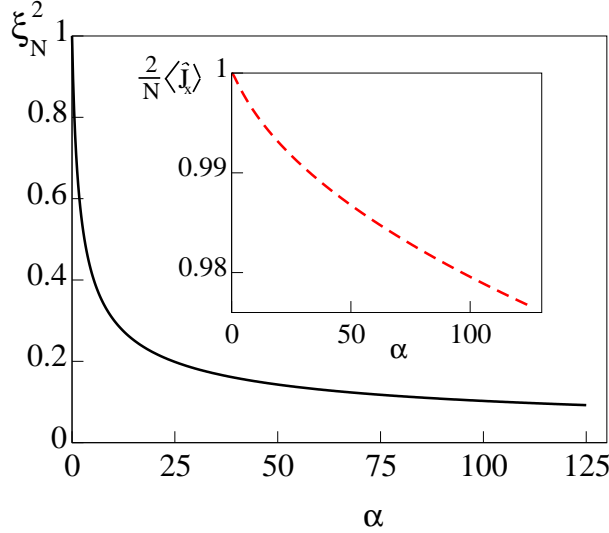

Figure 6.2: The spin-squeezing parameter defined in Eq. (6.26) calculcated with numerically obtained ground states of the Hamiltonian (6.18) with $N = 200$ particles plotted as a function of α. Since it drops below unity, the system is usefully entangled. The inset show the behavior of the denominator of ξ_N^2 normalized to the shot-noise level. Clearly, it hardly changes for the values of α taken.

$\alpha > 0$ are number-squeezed, because they have reduced fluctuations between the two wells of the trap. Note however, that the number-squeezing is closely related to another paramter, introduced in the Chapter 4, namely the spin-squeezing parameter

$$\xi_N^2 = N \frac{\left\langle \Delta^2 \hat{J}_z \right\rangle}{\left\langle \hat{J}_x \right\rangle^2}. \tag{6.26}$$

Recall that when $\xi_N^2 < 1$, then the state is particle-entangled. Note that the spin-squeezing parameter can be re-written in a following way

$$\xi_N^2 = \frac{\frac{4}{N} \left\langle \Delta^2 \hat{J}_z \right\rangle}{\left(\frac{2}{N} \left\langle \hat{J}_x \right\rangle \right)^2}. \tag{6.27}$$

Let us now concentrate on the nominator. According to Eq. (6.21), for the CSS, this nominator is equal to 1. Other ground states for $\alpha > 0$ have reduced fulctuations of the $\hat{J}_z$ operator, therefore the nominator drops below unity. Thus suggests that this family of ground states has $\xi_N^2 < 1$ and so it is spin-squeezed. To prove this, however, we have to demonstrate that $\left\langle \hat{J}_x \right\rangle$ remains high for these states, so that the denominator ensures that the whole ratio in Eq. (6.26) is not to big. For the CSS, which is an eigen-state of $\hat{J}_x$ with the eigen-value equal to $\frac{N}{2}$, the denominator is equal to 1, and so $\xi_N^2 = 1$ (no entanglement). But how does it behave for other states? We check it, by plotting in Fig. 6.2 the spin-squeezing parameter for numerically obtained ground states of the double-well Hamiltonian (6.18) as a function of α for $N = 200$. Clearly, ξ_N^2 drops below unity and so this family of states is usefully entangled from the metrological point of view. This observation motivated a number seminal experiment with ultracold atoms trapped in a double-well potential. We will now discuss how the particle-entangled state was created by reaching (approximately) the ground state of a double-well potential with repulsive interactions.

6.0.1   Experiment of Oberthaler

We will now explain the results of the paper [12] in much detail. The group of Markus Oberthaler was working on a Bose-Einstein Condensate of ^{87}Rb atoms. Initially, the gas was prepared in a harmonic trap. In the next step, however, an external 1-dimensional optical lattice was imprinted on the system using a set of two counter-propagating laser beams. Apart from this, the external harmonic trap is still present. By changing the steepness of this trap, they could control how many sites of the lattice were occupied, see Fig. 6.3. In particular, a two-
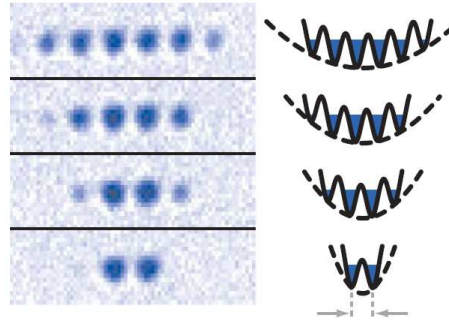

Figure 6.3: The experimental setup used to create a double-well setup. The gas is trapped in an optical lattice combined with an additional harmonic potential. By changing the depth of the trap, one can control the number of occupied lattice sites.

well setup could be prepared, as shown at the bottom of the figure.

The crucial point is, what is the state of the system after the imprint of

the optical lattice. Since the BEC is a gas of N particles roughly in the same quantum state, and since the optical lattice potential acts on each particle in the same way, then we can expect that the state after the splitting of the cloud into M sites is approximately given by the generalized CSS, i.e.

$$|\psi\rangle = \frac{1}{\sqrt{N!}} \left(\frac{\hat{a}_1^\dagger + \ldots + \hat{a}_M^\dagger}{\sqrt{M}} \right)^N |0\rangle , \qquad (6.28)$$

where the operator $\hat{a}_i^\dagger$ creates a particle in the i-th well. This expression for the state of the BEC agter splitting is justified only if during the splitting, some thermal excitations are not created in the system. Otherwise, more than a signle state per lattice site could be populated, and the above M-mode approximation would not be valid.

When the harmonic potential allows atoms to occupy only two sites of the lattice, i.e. $M = 2$, then we recover the expression for the CSS from Eq. (4.16). As we argued in the previous Section, this state is not particle entangled, because it is equivalent to the ground state of a two-well Bose-Hubbard Hamiltonian (6.18) with $\alpha = 0$. The question is, if using our knowledge from the previous Section, we could tell how a multi-particle entangled state could be generated in such system? The answer is: try to prepare the system in the ground state of (6.18) with $\alpha > 0$. But how could this be done, when staring from the CSS? The answer is simple. When the BEC is split into the two wells, the tunneling dominates over the two-body interactions, or in other words $\alpha \ll 1$. But this coefficient could be increased by rising the inter-well barrier. In this way, the tunneling rate drops, simply because it is more difficult for the particles to tunnel between the two sites. So although the strength of the two-body interactions remains constant, α will raise, because E_J drops. If the process of rising the inter-well barrier is adiabatical with comparison to other time-scales of the problem, than the system will follow the ground state of the Hamiltonian (6.18) with changing α. This way, one can create states, which are spin-squeezed, and thus particle-entangled.

Chapter 7

Time-of-flight interferometer

Chapter 8

Multi-mode interferometer

Bibliography

[1] Carl W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press (1976)

[2] H. Cramér, *Mathematical Methods of Statistics*, (Princeton Univ. Press, Princeton, NJ, 1946).

[3] A. Sørensen, L.-M. Duan, J. I. Cirac and P. Zoller, Nature **409**, 63 (2001)

[4] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004)

[5] L. Pezzé and A. Smerzi, Phys. Rev. Lett. **102**, 100401 (2009)

[6] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[7] B. Lücke, M. Scherer, J. Kruse, L. Pezzé, F. Deuretzbacher, P. Hyllus, O. Topic, J. Peise, W. Ertmer, J. Arlt, L. Santos, A. Smerzi and C. Klempt, Science **11**, 773 (2011)

[8] L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics: Theory and Applications* (Cambridge University Press, Cambridge, 1984).

[9] H. F. Hofmann, Phys. Rev. A **79**, 033822 (2009)

[10] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993)

[11] D. J. Wineland, J. J. Bollinger, W. M. Itano and D. J. Heinzen, Phys Rev. A **50**, 67 (1994)

[12] J. Estève, *et al.*, Nature **455**, 1216 (2008)

[13] M. F. Riedel, *et al.*, Nature **464**, 1170 (2010)

[14] C. Gross, *et al.*, Nature **464**, 1165 (2010)

[15] J. Appel, *et al.* , PNAS **106**, 10960 (2009)

[16] T. Berrada, S. van Frank, R. Bücker, T. Schumm, J.-F. Schaff, and J. Schmiedmayer, arxiv:1303.1030

[17] J. Grond, U. Hohenester, I. Mazets and J. Schmiedmayer, New J. Phys. **12**, 065036 (2010)

[18] J. Chwedeńczuk, P. Hyllus, F. Piazza and A. Smerzi, New J. Phys. **14**, 093001 (2012)

[19] J. Chwedeńczuk,, F. Piazza and A. Smerzi, New J. Phys. **13**, 065023 (2012)

[20] Franco Dalfovo, Stefano Giorgini, Lev P. Pitaevskii, and Sandro Stringari, Rev. Mod. Phys. **71**, 463 (1999)