Chapter 7

Quantum Bayesian estimation

In this chapter we will follow the Bayesian paradigm and develop methods to find optimal measurement and estimation strategies for quantum estimation models. A Bayesian quantum estimation model consists of the familiy of states ρ_θ and the prior distribution $p(\theta)$. The goal is to find a measurement $\{\Pi_x\}$ and an estimator $\tilde{\theta}(x)$ that minimize the average Bayesian cost:

$$\bar{C} = \int \mathrm{d}\theta\, p(\theta) \int \mathrm{d}x\, \mathrm{Tr}(\rho_\theta \Pi_x) C[\theta, \tilde{\theta}(x)], \tag{7.1}$$

where $C(\theta, \tilde{\theta})$ is the cost function penalizing for the deviation of the estimator from the true value. In particular, if we chose $C(\theta, \tilde{\theta}) = (\theta - \tilde{\theta})^2$ we return to the standard Bayesian variance cost function.

From chapter 5 we now that once the measurement is fixed and hence we can write the conditional probability $p(x|\theta)$ it is clear how to find the optimal Bayesian estimator. Still, the issue of determining the optimal measurement remains non-trivial.

Note that we can formally relabel the measurement operators Π_x to $\Pi_{\tilde{\theta}(x)}$, so that in fact the label represents the estimated value of parameter, $\Pi_{\tilde{\theta}} = \int \mathrm{d}x\, \Pi_x \delta(\tilde{\theta} - \tilde{\theta}(x))$. We do not loose any generality here, but thanks to this we can combine the double minimization over the measurement and the estimator to a single optimization over the measurements only:

$$\min_{\{\Pi_{\tilde{\theta}}\}} \bar{C}, \quad \Pi_{\tilde{\theta}} \geq 0, \quad \int \mathrm{d}\tilde{\theta}\, \Pi_{\tilde{\theta}} = \mathbb{1} \tag{7.2}$$

$$\bar{C} = \int \mathrm{d}\theta \mathrm{d}\tilde{\theta} p(\theta) \mathrm{Tr}(\rho_\theta \Pi_{\tilde{\theta}}) C(\theta, \tilde{\theta}). \tag{7.3}$$

Of course this in general is a untractable problem, as the space of all allowed generalized measurements is enormous. Still, as demonstrated below with some additional assumptions on the cost function or the set of states, the problem may be solved. Note that the above reformulation makes the classical results on the optimal Bayesian estimation not really very helpful in deriving fundamental bounds on precision, as we have incorporated the estimator in the labeling of the measurement operators and hence in some sense the optimal estimator is given for free once we solve the above search for the optimal measurement. This is not to say, that we will never utilize classical results. Quite contrary, whenever we will desire to provide a practical protocol that performs optimally we will eventually be forced to write down a standard (projective) measurement and the explicit form of estimator that achieves the bound derived using formal approach formulated above, and then we will definitely make use of the optimal Bayesian estimator construction known from classical theory.

7.1 Quadratic cost problems

Let us start by restricting ourselves to the quadratic cost function $C(\theta, \tilde{\theta}) = (\theta - \tilde{\theta})^2$. For simplicity of the formulas that follow we redefine the parameter θ so that the expectation value of the prior distribution is zero, $\int \mathrm{d}\theta p(\theta)\theta = 0$. The Bayesian variance to be minimized takes the form:

$$\overline{\Delta^2\tilde{\theta}} = \int \mathrm{d}\theta \mathrm{d}\tilde{\theta} p(\theta) \mathrm{Tr}[\rho_\theta \Pi_{\tilde{\theta}} (\theta - \tilde{\theta})^2] =$$
$$\int \mathrm{d}\theta p(\theta)\theta^2 + \mathrm{Tr}\left[\int \mathrm{d}\theta p(\theta)\rho_\theta \int \mathrm{d}\tilde{\theta}\Pi_{\tilde{\theta}}\tilde{\theta}^2\right] +$$
$$- 2\mathrm{Tr}\left[\int \mathrm{d}\theta p(\theta)\theta\rho_\theta \int \mathrm{d}\tilde{\theta}\Pi_{\tilde{\theta}}\tilde{\theta}\right] =$$
$$= \Delta^2\theta + \mathrm{Tr}(\bar{\rho}\Lambda_2) - 2\mathrm{Tr}(\bar{\rho}'\Lambda_1), \quad (7.4)$$

where we $\Delta^2\theta = \int \mathrm{d}\theta\, p(\theta)\theta^2$ represents the variance of the prior distribution, $\bar{\rho} = \int \mathrm{d}\theta\, p(\theta)\rho_\theta$ is the average state, $\bar{\rho}' = \int \mathrm{d}\theta\, p(\theta)\theta\rho_\theta$ and $\Lambda_k = \int \mathrm{d}\tilde{\theta}\, \Pi_{\tilde{\theta}}\tilde{\theta}^k$. The following theorem determines the minimum of the above quantity optimized over all measurements $\Pi_{\tilde{\theta}}$.

Theorem 7.1. Given family of states ρ_θ and the priori distribution $p(\theta)$ (with expectation value at $\theta = 0$) the minimal Bayesian variance for estimation of θ is given by:

$$\overline{\Delta^2\tilde{\theta}} = \Delta^2\theta - \mathrm{Tr}\left(\bar{\rho}\bar{\Lambda}^2\right), \quad (7.5)$$

where $\bar{\Lambda}$ is defined by the following equation:

$$\bar{\rho}' = \frac{1}{2}\left(\bar{\Lambda}\bar{\rho} + \bar{\rho}\bar{\Lambda}\right) \quad (7.6)$$

and $\bar{\rho} = \int \mathrm{d}\theta\, p(\theta)\rho_\theta$, $\bar{\rho}' = \int \mathrm{d}\theta\, p(\theta)\theta\rho_\theta$.

Proof. Let us first prove that if a given POVM measurement $\{\Pi_{\tilde{x}}\}$ is optimal, then we may find a projective measurement yielding the same cost. Let us perform eigen-decomposition of Λ_1 operator:

$$\Lambda_1 = \int \mathrm{d}\tilde{\theta}\, \Pi_{\tilde{\theta}}\tilde{\theta} = \sum_i \tilde{\theta}_i |\tilde{\theta}_i\rangle\langle\tilde{\theta}_i|. \quad (7.7)$$

Consider now the following inequality:

$$\int \mathrm{d}\tilde{\theta}(\tilde{\theta} - \Lambda_1)\Pi_{\tilde{\theta}}(\tilde{\theta} - \Lambda_1) \geq 0, \quad (7.8)$$

which is true since $\Pi_{\tilde{\theta}} \geq 0$ while Λ_1 is hermitian. This implies:

$$\int \mathrm{d}\tilde{\theta}\, \Pi_{\tilde{\theta}}\tilde{\theta}^2 + \Lambda_1^2 - 2\Lambda_1^2 \geq 0 \quad (7.9)$$

and hence

$$\Lambda_2 \geq \Lambda_1^2. \quad (7.10)$$

Let us now replace the measurement $\{\Pi_{\tilde{\theta}}\}$ with the projective measurement, corresponding to the projection on the eigenbasis $|\tilde{\theta}_i\rangle$ of Λ_1. For this choice $\Lambda_2 = \Lambda_1^2$, which according to (7.10) is the smallest operator possible. Inspecting (7.4) we see that we want the term $\mathrm{Tr}(\bar{\rho}\Lambda_2)$ to be as small as possible, and hence it is always optimal to choose the projective measurement in the eigenbasis of Λ_1.

Assuming the measurement i projective, we may now introduce a single operator variable write $\bar{\Lambda} = \Lambda_1$, $\Lambda_2 = \bar{\Lambda}^2$ and the optimization problem amounts to minimization of the following cost function over a single hermitian operator $\bar{\Lambda}$:

$$\overline{\Delta^2\tilde{\theta}} = \Delta^2\theta - \mathrm{Tr}(\bar{\rho}\bar{\Lambda}^2) - 2\mathrm{Tr}(\bar{\rho}'\bar{\Lambda}). \quad (7.11)$$

Since the above formula is quadratic in matrix $\bar{\Lambda}$, the minimization can be performed explicitly and the condition for fanishing first derivative amounts to the following linear equation:

$$\bar{\Lambda}\bar{\rho} + \bar{\rho}\bar{\Lambda} - 2\bar{\rho}' = 0. \quad (7.12)$$

Multiplying the above equality by $\bar{\Lambda}$ and taking the trace of both sides we get that $\mathrm{Tr}(\bar{\rho}'\bar{\Lambda}) = \mathrm{Tr}(\bar{\rho}\bar{\Lambda}^2)$ and therefore we arrive at the formula stated in the theorem. ■

Gaussian prior distribution. Equations (7.5,7.6) remind of formulas used for calculation of the QFI. The main difference is that instead of the derivative of the state on the left hand side of (7.6) there appear the $\bar{\rho}'$ operator. In order to establish a closer relation between these two approaches consider a gaussian prior distirbution $p_{\theta_0} = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(\theta-\theta_0)^2}{2\sigma^2}}$, where θ_0 is a free parameter determining the center of the prior. We will consider the effect of variation of the center of the prior around

$\theta_0 = 0$. We now have:

$$\bar{\rho}'\big|_{\theta_0=0} = \int \mathrm{d}\theta\, p_{\theta_0=0}(\theta)\theta\rho_\theta = \int \mathrm{d}\theta\, \left.\frac{\mathrm{d}p_{\theta_0}(\theta)}{\mathrm{d}\theta_0}\right|_{\theta_0=0} \rho_\theta\sigma^2 = \frac{\mathrm{d}}{\mathrm{d}\theta_0}\int \mathrm{d}\theta\, p_{\theta_0}(\theta)\rho_\theta\Big|_{\theta_0=-} = \sigma^2 \left.\frac{\mathrm{d}\bar{\rho}_{\theta_0}}{\mathrm{d}\theta_0}\right|_{\theta_0=0} = \Delta^2\theta \left.\frac{\mathrm{d}\bar{\rho}_{\theta_0}}{\mathrm{d}\theta_0}\right|_{\theta_0=0}. \tag{7.13}$$

From the above formula we see that for Gaussian prior $\bar{\rho}'$ is proportional to the derivative of the averaged state $\bar{\rho}$ with respect to the shift of the prior distribution and the proportionality constant is the prior variance. Therefore:

$$\frac{1}{2}\left(\bar{\rho}\bar{\Lambda} + \bar{\Lambda}\bar{\rho}\right) = \Delta^2\theta \frac{\mathrm{d}\bar{\rho}_{\theta_0}}{\mathrm{d}\theta_0} \tag{7.14}$$

and consequently $\bar{\Lambda} = \Lambda\Delta^2\theta$, where Λ is the standard SLD for the QFI estimation approach where using $\bar{\rho}_{\theta_0}$ we want to estimate changes in the center of the prior θ_0 around point $\theta_0 = 0$. As a result: $\mathrm{Tr}(\bar{\rho}\bar{\Lambda}^2) = (\Delta^2\theta)^2 F_Q(\bar{\rho}_{\theta_0})$. Finally we can write:

$$\overline{\Delta^2\tilde{\theta}} = \Delta^2\theta\left[1 - \Delta^2\theta F_Q(\bar{\rho}_{\theta_0})\right], \tag{7.15}$$

and we have arrived at the relation between the cost in the quantum Bayesian estimation and the QFI of the corresponding problem of estimating the prior from the averaged state $\bar{\rho}_{\theta_0}$.

7.2 Covariant estimation problems

Even though the previous section has provided us with a general recipe how to find the optimal Bayesian strategy for quadratic cost functions this is not always enough in practical problems. Note, that unlike in the frequentist approach where we in fact work in the paradigm of small fluctuations of parameter value around a known value, in the Bayesian approach we are typically faced with priors which are by no mean narrow and hence the approximation of arbitrary cost function to be locally quadratic is no longer justified.

This is especially pronounced in quantum estimation theory where we face problems where angle-like parameters, or more generally rotations are to be estimated. In such cases theory restricted to quadratic cost function is not really helpful. Unfortunately, if in the consideration from the previous chapter we replace the quadratic cost with some other cost function arbitrary one in general will not be able to provide a closed solution to the problem of determining the optimal measurement and hence the minimal cost.

Fortunately, as we show below if the problem enjoys certain symmetry, we may utilize some powerful methods based on group theory considerations and attack the problem of determining the optimal Bayesian estimation strategy from another perspective. We will refer to these class of estimation problems as covariant with respect to representation of a certain group.

Definition 7.1 (Covariant estimation problem)**.** Let G be a a Lie group where the group element $g \in G$ is is the estimation parameter in our problem. Let U_g be a unitary representation of the group in some Hilbert space, $U_{g_1}U_{g_2} = U_{g_1 g_2}$. We say that the Bayesian estimation problem is covariant with respect to U_g if and only if the following conditions are satisfied:

a) The parameter to be estimated is an element of the group $g \in G$

b) The family of states is generated by the action of the group representation $\rho_g = U_g\rho_e U_g^\dagger$—is the orbit of the group.

c) The cost function is left invariant with respect to the action of the group: $\forall_{g_1,g_2,h\in G} C(hg_1, hg_2) = C(g_1, g_2)$.

d) The prior distribution is invariant with respect to the action of the group:

$dg p(g) = d(hg)p(hg)$—the prior is uniform with respect to the Haar measure on the group.

Even though the above conditions appear quite restrictive they are nevertheless satisfied in a number of important quantum estimation problem, and in particular the phase estimation example we have been studying in the previous chapter.

Example 7.1 Consider the estimation problem, where the family of states

$$\rho_\varphi = |\psi_\varphi\rangle\langle\psi_\varphi|, \quad |\psi_\varphi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \qquad (7.16)$$

and we want to estimate φ given flat prior distribution $p(\varphi) = \frac{1}{2\pi}$. Note that the flat prior is a natural least informative choice for such a problem, and hence the issue of choosing the right prior that haunts the Bayesian approach is not relevant here. In order to formulate the complete model we need to choose a cost function. Note that the standard squared distance cost function is not a sensible choice as we are not working in the local phase estimation approach and hence the differences between the estimated and the real phase may be significant. Therefore, we want to use a cost function that takes into account the periodic nature of the phase parameter and does not penalizes us if the phase difference is a multiple of 2π. One natural choice is $C(\varphi, \tilde{\varphi}) = 4\sin^2\left(\frac{\varphi-\tilde{\varphi}}{2}\right)$—it reduces to the standard squared error for small phase deviations, takes maximum value for $\varphi - \tilde{\varphi} = \pi$ and reflects the periodic nature of the phase parameter.

This problem is indeed an example of a covariant estimation problem, where the group behind is the $U(1)$ group, with the representation $U_\varphi = e^{i\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]\varphi}$ in the qubit space. The family of states forms indeed an orbit, where $|\psi_\varphi\rangle = U_\varphi|\psi_0\rangle$, $|\psi_0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Finally, the prior as well as the cost function are invariant under the action of the group, as it amounts to simple phase addition.

For covariant estimation problems the Bayesian cost is given by:

$$\bar{C} = \int dg\, d\tilde{g}\, \mathrm{Tr}(U_g \rho_e U_g^\dagger\, \Pi_{\tilde{g}}) C(g, \tilde{g}), \qquad (7.17)$$

where we assume that dg is the normalized Haar measure on the group, $\int dg = 1$, with respect to which the prior is trivial $p(g) = 1$.

Definition 7.2 (Covariant measurement). $\{\Pi_{\tilde{g}}\}$ is called a covariant measurement with respect to an action of group representation U_g if and only if

$$\forall_{\tilde{g},h} U_h \Pi_{\tilde{g}} U_h^\dagger = \Pi_{h\tilde{g}}. \qquad (7.18)$$

Remark. In particular for a covariant measurement

$$\Pi_{\tilde{g}} = U_{\tilde{g}} \Pi_e U_{\tilde{g}}^\dagger, \qquad (7.19)$$

so that all measurement operators are determined by a single *seed* operator Π_e.

Theorem 7.2 (Optimality of covariant measurements). For the covariant estimation problem, the optimal measurement can be found within the class of covariant measurements.

Proof. Let $\Pi_{\tilde{g}}^{\mathrm{opt}}$ be the optimal measurement minimzing $\bar{C}$:

$$\bar{C}^{\mathrm{opt}} = \int dg\, d\tilde{g} \mathrm{Tr}(\Pi_{\tilde{g}}^{\mathrm{opt}} \rho_g) C(g, \tilde{g}). \qquad (7.20)$$

Let us define

$$\Pi_{\tilde{g}}^{\mathrm{cov}} = \int dg'\, U_{g'}^\dagger \Pi_{g'\tilde{g}}^{\mathrm{opt}} U_{g'}. \qquad (7.21)$$

This is indeed a covariant measurement, since:

$$U_h \Pi_{\tilde{g}}^{\mathrm{cov}} U_h^\dagger = \int dg'\, U_{hg'^{-1}} \Pi_{g'\tilde{g}} U_{g'h^{-1}} = \overset{g' \to g'h}{=} \int dg'\, U_{g'}^\dagger \Pi_{g'h\tilde{g}}^{\mathrm{opt}} U_{g'} = \Pi_{h\tilde{g}}^{\mathrm{cov}}. \qquad (7.22)$$

Moreover, the corresponding cost:

$$\begin{aligned} \bar{C}^{\mathrm{cov}} &= \int dg\, d\tilde{g}\, \mathrm{Tr}\left(\Pi_{\tilde{g}}^{\mathrm{cov}} \rho_g\right) C(g, \tilde{g}) = \\ &\int dg\, d\tilde{g}\, \mathrm{Tr}\left(\int dg'\, U_{g'}^\dagger \Pi_{g'\tilde{g}}^{\mathrm{opt}} U_{g'} U_g \rho_e U_g^\dagger\right) C(g, \tilde{g}) = \\ &\int dg\, d\tilde{g}\, dg'\, \mathrm{Tr}\left(U_{g'g}^\dagger \Pi_{g'\tilde{g}}^{\mathrm{opt}} U_{g'g} \rho_e\right) C(g, \tilde{g}) = \\ \overset{\substack{g \to g'^{-1}g \\ \tilde{g} \to g'^{-1}\tilde{g}}}{=} &\int dg\, d\tilde{g}\, dg'\, \mathrm{Tr}\left(U_g^\dagger \Pi_{\tilde{g}}^{\mathrm{opt}} U_g \rho_e\right) C(g'^{-1}g, g'^{-1}\tilde{g}) = \\ &\int dg\, d\tilde{g}\, dg'\, \mathrm{Tr}\left(\Pi_{\tilde{g}}^{\mathrm{opt}} \rho_g\right) C(g, \tilde{g}) = \bar{C}^{\mathrm{opt}}. \end{aligned} \qquad (7.23)$$

So is equal to the optimal minimal cost. ■

Thanks to the above theorem, the problem of identifying the optimal estimation strategy may be significantly simplified. Note that thanks to the covariance property of the measurement we have:

$$\bar{C} = \int \mathrm{d}g\,\mathrm{d}\tilde{g}\,\mathrm{Tr}\,(\Pi_{\tilde{g}}\rho_g)\,C(g,\tilde{g}) =$$
$$= \int \mathrm{d}g\,\mathrm{d}\tilde{g}\,\mathrm{Tr}\left(U^\dagger_{\tilde{g}^{-1}g}\Pi_e U_{\tilde{g}^{-1}g}\rho_e\right) C(g,\tilde{g}) =$$
$$\overset{g\to\tilde{g}g}{=} \int \mathrm{d}g\,\mathrm{d}\tilde{g}\,\mathrm{Tr}\left(U^\dagger_{g}\Pi_e U_{g}\rho_e\right) C(g,e) =$$
$$= \int \mathrm{d}g\,\mathrm{Tr}\,(\Pi_e\rho_g)\,C(g,e). \quad (7.24)$$

The whole problem now amounts to minimization of the above quantity over a *single* operator Π_e with constraints $\Pi_e \geq 0$, $\int \mathrm{d}g\, U_g \Pi_e U_g^\dagger = \mathbb{1}$. This is a huge simplification of the original problem and often the optimal operator Π_e may be found analytically, as is demonstrated in the examples that follow.

Example 7.2 (Phase estimation on product qubit states) Let us consider the problem of estimating the phase φ given N qubits:

$$\rho^{(N)}_\varphi = U^{\otimes N}_\varphi |\psi_0\rangle\langle\psi_0|^{\otimes N} U^{\dagger\otimes N}, \quad |\psi_0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \quad (7.25)$$

where $U_\varphi = e^{i\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right]\varphi}$. We assume the flat prior $p(\varphi) = 1/2\pi$, and the cost function $C(\varphi,\tilde{\varphi}) = 4\sin^2\left(\frac{\varphi-\tilde{\varphi}}{2}\right)$. The goal is to maximize

$$\bar{C}^{(N)} = \frac{2}{\pi}\int_0^{2\pi} \mathrm{d}\varphi\,\mathrm{Tr}\left(\Pi^{(N)}_0 \rho^{(N)}_\varphi\right)\sin^2\varphi \quad (7.26)$$

over $\Pi^{(N)}_0$, such that $\Pi^{(N)}_0 \geq 0$, $\int \frac{\mathrm{d}\varphi}{2\pi} U^{\otimes N}_\varphi \Pi^{(N)}_0 U^{\dagger\otimes N}_\varphi = \mathbb{1}$.

Let us first solve the $N = 1$ case:

$$\bar{C}^{(1)} = \frac{1}{\pi}\mathrm{Tr}\left(\Pi^{(1)}_0 \int_0^{2\pi}\mathrm{d}\varphi\,\sin^2\varphi \begin{bmatrix} 1 & e^{-i\varphi} \\ e^{i\varphi} & 1\end{bmatrix}\right). \quad (7.27)$$

Taking into account that $\sin^2\varphi = \frac{1}{2} - \frac{1}{4}e^{i\varphi} - \frac{1}{4}e^{-i\varphi}$ and performing the integral in the above formula yields:

$$\bar{C}^{(1)} = \mathrm{Tr}\left(\Pi^{(1)}_0 \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1\end{bmatrix}\right). \quad (7.28)$$

We need to minimize the above quantity over $\Pi^{(1)}_0$, keeping in mind that $\Pi^{(1)}_0 \geq 0$, $\int \frac{\mathrm{d}\varphi}{2\pi} U_\varphi \Pi^{(1)}_0 U^\dagger_\varphi = \mathbb{1}$. The completness condition imply that:

$$\Pi^{(1)}_0 = \begin{bmatrix} 1 & a \\ a^* & 1\end{bmatrix} \quad (7.29)$$

where a is an arbitrary complex number. However, the positivity condition further implies that $|a| \leq 1$. We therefore need to perform the following minimization:

$$\min_{|a|\leq 1}\mathrm{Tr}\left(\begin{bmatrix} 1 & a \\ a^* & 1\end{bmatrix}\begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1\end{bmatrix}\right) =$$
$$= \min_{|a|\leq 1}(2 - \mathrm{Re}a) = 1 \quad (7.30)$$

for $a = 1$. This means that the optimal seed measurement

$$\Pi^{(1)}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1\end{bmatrix} = 2|+\rangle\langle+|. \quad (7.31)$$

Note that $\Pi^{(1)}_\varphi = 2|\psi_\varphi\rangle\psi_\varphi$ so the optimal measurement corresponds to POVM spanning all equatorial states. The resulting cost equals:

$$\bar{C}^{(1)} = 1. \quad (7.32)$$

One can check, that even though we have achieved the minimal cost using covariant measurement, the same cost would be obtained in this case using the simplest possible projective measurement with $\Pi_0 = |+\rangle\langle+|$, $\Pi_1 = |-\rangle\langle-|$ and the corresponding estimated values of the phase equal to $\tilde{\varphi}(0) = 0$, $\tilde{\varphi}(1) = \pi$.

We now move on to solve the general case $N > 1$. First of all note that $|\psi_\varphi\rangle^{\otimes N} \in \mathcal{H}^{\otimes N}_S$, where $\mathcal{H}^{\otimes N}_S$ is the fully symmetric (bosonic) subspace of the space of N qubits. Therefore without loosing generality we may restrict ourselves to this subspace. This space is $N+1$ dimensional so it significantly reduces the size of the considered Hilbert space. Let us denote by $|n\rangle$ ($n = 0,\ldots,N$) symmetric states with n qubits in states $|1\rangle$ and $N-n$ qubits in state 0:

$$|n\rangle = \frac{1}{\sqrt{\binom{N}{n}}}\sum_{\mathrm{perm}} |\underbrace{0,\ldots,0}_{N-n},\underbrace{1,\ldots,1}_{n}\rangle, \quad (7.33)$$

where the sum is performed over all non-trivial permutations. Then

$$|\psi_\varphi\rangle = \frac{1}{\sqrt{2^N}}\sqrt{\binom{N}{n}}e^{in\varphi}|n\rangle. \quad (7.34)$$

After the integration over φ is performed the formula for the cost to be minimized reduces to

$$\bar{C}^{(N)} = \frac{1}{2^{N-1}}\mathrm{Tr}\left(\Pi^{(N)}_0 A\right), \quad (7.35)$$

where A is the following $N+1 \times N+1$ tridiagonal matrix

$$A = \sum_{n=0}^N \binom{N}{n}|n\rangle\langle n| +$$
$$-\frac{1}{2}\sum_{n=1}^N\sqrt{\binom{N}{n}\binom{N}{n-1}}\left(|n\rangle\langle n-1| + |n-1\rangle\langle n|\right). \quad (7.36)$$

Within the symmetric subspace the group representation acts according to $U^S_\varphi = \sum_n e^{in\varphi}|n\rangle\langle n|$. The completeness condtion (restrcited to the symmetric subspace) implies that $(\Pi_0^{(N)})^n_n = 1$, while off-diagonal entries are arbitrary. Still, positivity condition requires that all off-diagonal elements have absolute values less or equal 1. Since off-diagonal terms in A will all contribute negaitvely to the final cost it is optimal to choose all off diagonal elements in the $\Pi_0^{(N)}$ operator to be 1. Hence the optimal

$$\Pi_0^{(N)} = \left(\sum_{n=0}^{N}|n\rangle\right)\left(\sum_{m=0}^{N}\langle m|\right) \tag{7.37}$$

and the resulting cost:

$$\bar{C}^{(N)} = \frac{1}{2^{N-1}}\left(\sum_{n=0}^{N}\binom{N}{n} - \sum_{n=1}^{N}\sqrt{\binom{N}{n}\binom{N}{n-1}}\right) = \\ = 2 - \frac{1}{2^{N-1}}\sum_{n=1}^{N}\sqrt{\binom{N}{n}\binom{N}{n-1}}. \tag{7.38}$$

One can check that in the limit of large N, $\bar{C}^{(N)} \to \frac{1}{N}$, see Problem 7.2, and hence we recover the identical asymptotic result as obtained within the frequentist approach.

Example 7.3 (Estimation of a completely unknown qubit state) We consider a model in which are given N copies of a completely unknowns qubit state and our goal is to estimate it. Formally the state reads:

$$\rho_\psi^{(N)} = |\psi\rangle\langle\psi|^{\otimes N},\ |\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle, \tag{7.39}$$

where we parameterize the state of a qubit using Bloch sphere spherical angles. In order to think of this problem as a covariant estimation problem we may view $|\psi\rangle$ as obtained by rotating a fixed state $|0\rangle$ be a representation of the SU(2) group. More precisely, since the initial state $|0\rangle$ will not change under rotations around the z axis, the relevant group will be $G = \mathrm{SU}(2)/\mathrm{U}(1)$.

$$|\psi\rangle = U_\psi|0\rangle = e^{i\varphi\sigma_z/2}e^{i\theta\sigma_y/2}|0\rangle. \tag{7.40}$$

The Haar measure in this case corresponds to the natural measure on the sphere so we will integrate over the states using

$$\mathrm{d}\psi = \frac{1}{4\pi}\mathrm{d}\theta\mathrm{d}\varphi\sin\theta \tag{7.41}$$

As a cost function we choose:

$$C(\psi,\tilde{\psi}) = 4(1 - |\langle\psi|\tilde{\psi}\rangle|^2). \tag{7.42}$$

Note that this represents a distance derived from the fidelity measure, and for infitesimally close states will reduce to standard squared distance on the sphere, see Problem.... We have now formulated our problem as a covariant estimation problem. Using (7.24) we have

$$\bar{C}^{(N)} = \int \mathrm{d}\psi\, \mathrm{Tr}\left(\Pi_e|\psi\rangle\langle\psi|^{\otimes N}\right)4(1 - |\langle\psi|0\rangle|^2), \tag{7.43}$$

which we need to minimize over Π_e, keeping in mind $\Pi_e \geq 0$, $\int \mathrm{d}\psi\, U_\psi^{\otimes N}\Pi_e U_\psi^{\dagger\otimes N} = \mathbb{1}$. Let us rewrite the expression for the cost

$$\bar{C}^{(N)} = 4\left[1 - \underbrace{\int \mathrm{d}\psi\, \mathrm{Tr}\left(\Pi_e|\psi\rangle\langle\psi|^{\otimes N}\right)|\langle\psi|0\rangle|^2}_{\mathcal{F}}\right], \tag{7.44}$$

where we have introduced $\mathcal{F}$ which may be viewed as the fidelity of estimation, which needs to be maximized. We can equivalently write $\mathcal{F}$ as:

$$\mathcal{F} = \int \mathrm{d}\psi\, \mathrm{Tr}\left(\Pi_e \otimes |0\rangle\langle 0|\, |\psi\rangle\langle\psi|^{\otimes N+1}\right), \tag{7.45}$$

where we have formally extended the space to $N+1$ qubit space in order to incorporate the cost function inside the trace operator. Note that

$$\mathcal{F} = \mathrm{Tr}\left(\Pi_e \otimes |0\rangle\langle 0| \underbrace{\int \mathrm{d}\psi|\psi\rangle\langle\psi|^{\otimes N+1}}_{A}\right). \tag{7.46}$$

Let us study the properties of the A operator. This operator clearly is suported on the fully symmetric subspace of $N+1$ qubits and has trace 1. Moreover, this operator is invariant under the action of $U_\psi^{\otimes N+1}$, and since the fully-symmetric subspace carries the irreducible representation of the $U_\psi^{\otimes N+1}$ representation (with total angular momentum $j = (N+1)/2$), then by Schur lemma this operator must be proportional to identity on this subspace. The fully symmetric subspace of $N+1$ qubits has dimension $N+2$ and as a result:

$$A = \frac{1}{N+2}\mathbb{1}_{\mathcal{H}_S^{\otimes N+1}}. \tag{7.47}$$

We now need to find Π_e such that $\mathrm{Tr}\left(\Pi_e \otimes |0\rangle\langle 0|\,\mathbb{1}_{\mathcal{H}_S^{\otimes N+1}}\right)$ is maximal. We may restrict the Π_e operator to act solely on the symmetric subspace $\mathcal{H}_S^{\otimes N}$ as this the subspace where states $|\psi\rangle^{\otimes N}$ live. Let us denote $U_\psi^{j=N/2}$ to be the irreducible representation of $SU(2)$ acting on this subspace. Taking into account the completeness condition for Π_e:

$$\int \mathrm{d}\psi U_\psi^{j=N/2}\Pi_e U_\psi^{j=N/2\dagger} = \mathbb{1}_{\mathcal{H}_S^{\otimes N}} \tag{7.48}$$

we see that $\mathrm{Tr}\Pi_e = N+1$. It is clear that in order to have the largest overlap between $\Pi_e \otimes |0\rangle\langle 0|$ and $\mathbb{1}_{\mathcal{H}_S^{\otimes N+1}}$, we would like to have $\Pi_e \otimes |0\rangle\langle 0|$ operator fully supported on $\mathcal{H}_S^{\otimes N+1}$. This will be so provided we choose

$$\Pi_e = |0\rangle\langle 0|^{\otimes N}(N+1). \tag{7.49}$$

As a result we get $\mathcal{F} = \frac{N+1}{N+2}$ and finally

$$\bar{C}^{(N)} = 4\left(1 - \frac{N+1}{N+2}\right) \tag{7.50}$$