

Bayesian minimum mean square error for transmissivity sensing

Boyu Zhou,¹ Boulat A. Bash,^{2,3} Saikat Guha,^{3,2} and Christos N. Gagatsos^{2,3}

¹*Department of Physics, The University of Arizona 1118 E. Fourth Street, Tucson, AZ 85721, USA*

²*Department of Electrical and Computer Engineering,
The University of Arizona, Tucson AZ, 85721, USA*

³*Wyant College of Optical Sciences, The University of Arizona,
1630 E. University Blvd., Tucson, AZ 85721, USA*

We address the problem of estimating the transmissivity of the pure-loss channel from the Bayesian point of view, i.e., we consider that some prior probability distribution function (PDF) on the unknown variable is available and we employ methods to compute the Bayesian minimum mean square error (MMSE). Specifically, we consider two prior PDFs: the two-point and the beta distributions. By fixing the input mean photon number to an integer, for the two-point PDF we prove analytically that the optimal state is the Fock state and the optimal measurement is photon-counting, while for the beta PDF our numerical investigation provides evidence on the optimality of the Fock state and photon-counting. Moreover, we investigate the situation where the input mean photon number is any (non-negative) real number. For said case, we conjecture the form of the optimal input states and we study the performance of photon-counting, which is a sub-optimal yet practical measurement. Our methods can be applied for any prior PDF. We emphasize that we compute the MMSE instead of Bayesian lower bounds on the mean square error based on the Fisherian approach.

I. INTRODUCTION

Estimation of the transmissivity $\tau \in [0, 1]$ (or equivalently of loss) is of paramount importance in quantum communications as it determines the fundamental limit of any point-to-point quantum protocol [1]. In the quantum sensing setting, estimation of loss of power is connected to the reflectance of a target, and it affects the sensing system's ability to attain Heisenberg scaling in many scenarios [2, 3]. Moreover, transmissivity estimation plays a crucial role in quantum network diagnostics: Given a network where information-carrying light is transmitted through, one wishes to estimate what the unavoidable losses are due to imperfections, for experimental, manufacturing, or other practical reasons. Lastly, estimation of transmissivity and the questions that come with it (e.g. optimal states and measurements for said task) is interesting in its own right from the theory point of view.

To our knowledge, the problem at hand has been addressed through the Fisherian approach [4–8], i.e., considering the quantum and classical Fisher information (denoted as $F(\tau)$ and $C(\tau)$) whose inverses serve as lower bounds on the mean square error (MSE), with $F^{-1}(\tau)$ being the fundamental limit. These bounds are the well-known Cramér-Rao bounds. In this work, we do not consider the unknown parameter as having a fixed value, but we consider it as being a random variable that follows a prior probability distribution (PDF) $P(\tau)$, which gives rise to Bayesian quantum estimation [9, 10]. The use of a prior PDF

can be justified as a result of prior measurements of the system [11] or of being based upon natural assumptions, e.g., the manufacturer can provide information, in the form of a PDF, on the loss of an optical fiber by performing diagnostics on random samples, but without specifying the loss rate of the equipment at hand.

The pure-loss channel is a successful model for lossy optical elements and is defined as a beam splitter with transmissivity τ , whose lower input mode is set to vacuum and its lower output mode is traced out. The upper input and output modes are referred to as input and output of the channel respectively. For said channel, and within the Fisherian approach, one of the results in [5] was that for fixed input mean photon number, the Fock state is the optimal probe state to sense τ when the input mean photon number is an integer, while for that case photon-counting was proven to be the optimal measurement. Results for fixed mean photon number to non-integers were also derived.

If a prior PDF is available, one can construct Bayesian versions of the Fisherian Cramér-Rao bounds: Let us consider a single unknown variable x that follows a (prior) PDF $P(x)$. Also, let us consider the optimal measurement in the sense that the QFI is equal to the CFI, i.e., $F(x) = C(x)$, where we note that such a measurement always exists in the single-parameter setting. Then, the Bayesian version of the Cramér-Rao bound is [12],

$$\delta_B \geq J_E^{-1} \geq J_B^{-1}, \quad (1)$$

where,

$$\delta_B = \int dx P(x) \text{tr} [\hat{\rho}(x)(\hat{H} - x\hat{I})^2] \quad (2)$$

is the Bayesian mean square error for a measurement given by the eigenvectors of the Hermitian operator $\hat{H}$ on the final state $\hat{\rho}(x)$,

$$J_E^{-1} = \int dx P(x) F^{-1}(x), \quad (3)$$

$$J_B = J_D + J_P, \quad (4)$$

where J_D is the Fisher information related to data,

$$J_D = \int dx P(x) F(x) \quad (5)$$

and J_P is the Fisher information related to the prior PDF,

$$J_P = \int dx P(x) \left(\frac{\partial \ln P(x)}{\partial x} \right)^2. \quad (6)$$

In inequality (1), J_B is not always attainable and J_E^{-1} is attainable in the limit where the collected data goes to infinity [12]. Therefore, said bounds are not guaranteed to fully capture the impact of prior information. For any given Bayesian sensing task, even if examined using the Fisherian approach, a rigorous analysis must be re-done in a genuinely Bayesian fashion [9, 10]. However, some physical intuition can be drawn from results obtained in the Fisherian approach.

In this paper we set $x = \tau$ and we consider the formalism developed in [9]: We work with the minimum mean square error (MMSE) where the minimization is performed over all possible protective measurements, i.e., all possible Hermitian operators $\hat{H}$ of Eq. (2),

$$\delta \equiv \min_{\hat{H}} \delta_B. \quad (7)$$

This paper is organized as follows: In Section II we describe the pure-loss channel and review the Bayesian approach to single-variable quantum estimation [9]. In Section III we derive a lower bound on the MMSE which will be useful to the technical development of later sections. In Section IV we prove that for fixed input mean photon number to an integer and for the two-point PDF as prior, the Fock state is optimal and also find that the optimal measurement is photon-counting. In Section V, we consider the two-point prior PDF and we do a numerical investigation on what states are optimal for non-integer fixed input mean photon and we examine

the behavior of photon-counting as a feasible (yet sub-optimal) measurement. In Section VI, we repeat the same analysis, this time based only on numerics, for the beta distribution and we find similar results to the previous sections. Finally, in Section VII, we discuss further our results, we do a comparison between the bounds of inequality (1) and the MMSE δ , and we discuss future research directions.

II. THE MATHEMATICAL TOOLS AND THE SETUP

We denote the input of the pure-loss channel as $\hat{\rho}_0$. Then, the output state is $\hat{\rho}(\tau) = \mathcal{L}_\tau[\hat{\rho}_0]$. The time evolution of the input state $\hat{\rho}_0$ is governed by the following master equation and its initial condition,

$$\frac{\partial \hat{\rho}(t)}{\partial t} = \frac{\gamma}{2} [2\hat{a}\hat{\rho}(t)\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho}(t) - \hat{\rho}(t)\hat{a}^\dagger\hat{a}], \quad (8)$$

$$\hat{\rho}(t=0) = \hat{\rho}_0, \quad (9)$$

where the $\hat{a}$, $\hat{a}^\dagger$ are respectively the annihilation and creation operators of a single bosonic mode satisfying the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$ and through which we define the number operator $\hat{n} = \hat{a}^\dagger\hat{a}$. The loss parameter γ is related to the transmissivity τ of the pure-loss channel as follows,

$$\tau = \exp(-\gamma t). \quad (10)$$

Under the parameterization of Eq. (10), the solution to Eq. (8) under the initial condition of Eq. (9) is,

$$\hat{\rho}(\tau) = \sum_{l=0}^{\infty} \frac{(1-\tau)^l}{l!} \sqrt{\tau}^{\hat{n}} \hat{a}^l \hat{\rho}_0 \hat{a}^{\dagger l} \sqrt{\tau}^{\hat{n}}. \quad (11)$$

The problem set up we consider in this work is as follows: We consider pure input states $\hat{\rho}_0 = |\Psi_{\bar{n}}\rangle\langle\Psi_{\bar{n}}|$ with fixed mean-photon number $\langle\Psi_{\bar{n}}|\hat{n}|\Psi_{\bar{n}}\rangle = \bar{n}$. We consider as unknown parameter the transmissivity τ and on top of this we assume that τ is a random variable following a prior PDF, $P(\tau)$. Then, the task is to find the optimal input state that minimizes the minimum mean square error (MMSE). Under this Bayesian approach, Personick derived the formulas that give the MMSE δ and which projective measurement attains it [9]. For our purposes, said formulas are written as,

$$\delta = \text{tr}\hat{\Gamma}_2 - \text{tr}(\hat{B}\hat{\Gamma}_1). \quad (12)$$

where,

$$\hat{\Gamma}_k = \int_0^1 d\tau P(\tau) \tau^k \hat{\rho}(\tau), \quad k = 0, 1, 2. \quad (13)$$

and

$$\hat{B} = 2 \int_0^\infty dz e^{-z\hat{\Gamma}_0} \hat{\Gamma}_1 e^{-z\hat{\Gamma}_0}, \quad (14)$$

where the MMSE δ is always attainable and the optimal projective measurement is given by the eigenvectors of $\hat{B}$.

In this work we specifically consider the two-point prior PDF,

$$P(\tau) = q\delta(\tau - \tau_0) + (1 - q)\delta(\tau - \tau_1), \quad (15)$$

where $0 \leq q \leq 1$, i.e., the transmissivity can take two possible values τ_0 and τ_1 with probabilities q and $1 - q$ respectively. The two-point prior PDF allows for two possible transmissivities, occurring with a probabilities q and $1 - q$. This problem is akin to target detection and quantum reading but viewed from the angle of the MMSE instead of the minimum probability of error [13, 14].

We also discuss briefly the beta distribution

$$P(\tau) = \frac{\tau^{\alpha-1}(1-\tau)^{\beta-1}}{\mathcal{B}(\alpha, \beta)}, \quad (16)$$

where $\mathcal{B}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$, $\Gamma(\cdot)$ is the gamma function, and the values of $\alpha > 0$, $\beta > 0$ cover a great variety of PDF's behavior.

We note here that if the prior PDF is simply a Dirac delta $P(\tau) = \delta(\tau - \tau_0)$ then one can prove that the MMSE is $\delta = 0$, which makes perfect sense since such a prior PDF gives full knowledge on the parameter.

Also, we note that in our analysis we assume that the optimal input states are pure since classical mixing involves throwing away information.

III. LOWER BOUND ON THE MMSE

The first term of Eq. (12) depends only on the prior PDF, i.e., since $\text{tr}\hat{\rho} = 1$, $\text{tr}\hat{\Gamma}_2 = \int_0^1 P(\tau)\tau^2$. Since we fix the prior PDF, when lower-bounding the MMSE we only consider the input state which affects the second term of Eq. (12).

Upper bounding the second term of Eq. (12) will yield a lower bound on δ . We write the second term of Eq. (12) as,

$$\text{tr}(\hat{B}\hat{\Gamma}_1) = 2 \int_0^\infty dz \text{tr} \left(e^{-z\hat{\Gamma}_0} \hat{\Gamma}_1 e^{-z\hat{\Gamma}_0} \hat{\Gamma}_1 \right). \quad (17)$$

For Hermitian operators $\hat{X}$ and $\hat{Y}$, we have the inequality $\text{tr}(\hat{X}\hat{Y}\hat{X}\hat{Y}) \leq \text{tr}(\hat{X}^2\hat{Y}^2)$ [15]. The operators

$\hat{\Gamma}_k$ are Hermitian, therefore using said inequality we write,

$$\text{tr}(\hat{B}\hat{\Gamma}_1) \leq 2 \int_0^\infty dz \text{tr} \left(e^{-2z\hat{\Gamma}_0} \hat{\Gamma}_1^2 \right) \Leftrightarrow \quad (18)$$

$$\text{tr}(\hat{B}\hat{\Gamma}_1) \leq \text{tr}(\hat{\Gamma}_0^{-1}\hat{\Gamma}_1^2) = \text{tr}(\hat{\Gamma}_1^2\hat{\Gamma}_0^{-1}), \quad (19)$$

where this lower bound is meaningful as long as $\hat{\Gamma}_0$ is invertible. Under said condition, we get a lower bound on the MMSE δ ,

$$\delta \geq \delta_{\text{LB}}, \quad (20)$$

$$\delta_{\text{LB}} = \int_0^1 d\tau P(\tau)\tau^2 - \text{tr}(\hat{\Gamma}_0^{-1}\hat{\Gamma}_1^2). \quad (21)$$

In inequality (20), equality is achieved if and only if $[\hat{\Gamma}_0, \hat{\Gamma}_1] = 0$.

IV. INTEGER MEAN PHOTON NUMBER: OPTIMALITY OF FOCK STATES

In this section we prove that for input mean photon number fixed to an integer, i.e., $\bar{n} = n \in \mathbb{N}$, and for the prior PDF of Eq. (15), the only pure state that attains the lower bound of Eq. (21) is the Fock state $|n\rangle$, i.e., the Fock state is optimal in the sense that attains the lowest MMSE. For this case, we also prove that the optimal measurement is photon-counting or photon number resolving (PNR) detection, i.e., projection on the Fock basis.

Using Eq. (13), the condition $[\hat{\Gamma}_0, \hat{\Gamma}_1] = 0$ implies,

$$\int_0^1 d\tau \int_0^1 d\tau' \tau' P(\tau) P(\tau') [\hat{\rho}(\tau), \hat{\rho}(\tau')] = 0. \quad (22)$$

For the prior of Eq. (15), Eq. (22) gives,

$$q(1 - q)(\tau_0 - \tau_1)[\hat{\rho}(\tau_0), \hat{\rho}(\tau_1)] = 0, \quad (23)$$

where the integrals disappeared because of the Dirac deltas, and two terms gave zero as they are proportional to $[\hat{\rho}(\tau_0), \hat{\rho}(\tau_0)] = 0$ and $[\hat{\rho}(\tau_1), \hat{\rho}(\tau_1)] = 0$. From Eq. (23), there are three options: (i) $q = 0$ or $q = 1$, which means that the prior is simply a single Dirac delta and for that case we get that the MMSE is $\delta = 0$ regardless of the input state as full information on the parameter is provided by the prior PDF, (ii) $\tau_0 = \tau_1$, in which case the prior PDF is again a single Dirac delta, and (iii) $[\hat{\rho}(\tau_0), \hat{\rho}(\tau_1)] = 0$.

Since τ_0 and τ_1 are arbitrary values of the transmissivity, the non-trivial option $[\hat{\rho}(\tau_0), \hat{\rho}(\tau_1)] = 0$ means that any two outputs of the pure-loss channel must commute so that the lower bound of Eq. (21) is attained.

It is known that if the input of the pure-loss channel is a Fock state, the output will remain Fock diagonal. Therefore, Fock states satisfy the condition $[\hat{\Gamma}_0, \hat{\Gamma}_1] = 0$ and as a result they attain the lower bound of Eq. (21).

At the same time, the condition $[\hat{\Gamma}_0, \hat{\Gamma}_1] = 0$ must be true for all τ_0 and τ_1 , including the choice $\tau_0 = 0$ and $0 < \tau_1 \leq 1$. For said choice, from Eq. (11) we get $\hat{\rho}(\tau_0 = 0) = |0\rangle\langle 0|$, which means that $\hat{\rho}(\tau_1)$ must be Fock-diagonal for all $0 < \tau_1 \leq 1$. In Appendix A we prove that a Fock-diagonal output of the pure-loss channel can only correspond to a Fock-diagonal input, and since we restrict our search to pure states, the only Fock-diagonal input state is the Fock state.

Furthermore, let a state $|\Psi_n\rangle$ and a Fock state $|n\rangle$, with mean photon number $\langle \Psi_n | \hat{n} | \Psi_n \rangle = \langle n | \hat{n} | n \rangle = n$. Since all states in a Hilbert space can be connected through at least one unitary operator, there is always at least one unitary $\hat{U}_n$, such that $\hat{U}_n |\Psi_n\rangle = |n\rangle$. Therefore, the state preparation scheme $\hat{U}_n |\Psi_n\rangle = |n\rangle$ results to a lower MMSE for all pure states with the same input energy.

We now derive the expression of the MMSE for a Fock state input and the optimal measurement that attains it. In general, the main difficulty with the MMSE formulation is to compute analytically the term $\exp(-z\hat{\Gamma}_0)$ required by Eq. (14). Since for a Fock input state, the output remains Fock diagonal, per Eq. (13) $\hat{\Gamma}_0$ is Fock-diagonal, rendering $\exp(-z\hat{\Gamma}_0)$ Fock-diagonal as well. In Appendix C we derive the MMSE,

$$\begin{aligned} \delta^{(n)} &= \delta_{\text{LB}}^{(n)} \\ &= q\tau_0^2 + (1-q)\tau_1^2 - \sigma^{(n)}(q, \tau_0, \tau_1), \end{aligned} \quad (24)$$

where,

$$\sigma^{(n)}(q, \tau_0, \tau_1) = \sum_{l=0}^n \frac{[q\tau_0 e_l^{(n)}(\tau_0) + (1-q)\tau_1 e_l^{(n)}(\tau_1)]^2}{qe_l^{(n)}(\tau_0) + (1-q)e_l^{(n)}(\tau_1)}$$

and

$$e_l^{(n)}(\tau) = \binom{n}{l} \tau^{n-l} (1-\tau)^l. \quad (25)$$

Since the $\hat{\Gamma}_k$ operators are diagonal on the Fock basis, the operator $\hat{B}$ is also Fock-diagonal, therefore the optimal projective measurement consists of projections on the Fock basis, i.e., PNR detection. In Appendix B, we find the form of the $\hat{B}$ operator,

$$\hat{B} = \sum_{l=0}^n b_l^{(n)}(q, \tau_0, \tau_1) |n-l\rangle\langle n-l|, \quad (26)$$

where,

$$b_l^{(n)}(q, \tau_0, \tau_1) = \frac{q\tau_0 e_l^{(n)}(\tau_0) + (1-q)\tau_1 e_l^{(n)}(\tau_1)}{qe_l^{(n)}(\tau_0) + (1-q)e_l^{(n)}(\tau_1)}.$$

We note that even though the symmetric logarithm derivative found in [5] (derived within the Fisherian context) and the $\hat{B}$ presented here are both Fock-diagonal, their exact form is different, i.e., their eigenvalues are different, reflecting the incomparability of the Bayesian and the Fisherian approaches. Moreover, as we discuss in Section VII, it is not obvious what is the optimal state for arbitrary prior PDFs.

V. NON-INTEGHER MEAN PHOTON NUMBER

We now consider the input mean photon number $\bar{n}$ as a real non-negative, not necessarily an integer. For this case, we provide numerical evidence that the optimal state has the form,

$$|\Phi_{\bar{n}}\rangle = |a(\bar{n})| |\lceil \bar{n} \rceil - 1\rangle + |c(\bar{n})| |\lceil \bar{n} \rceil\rangle \quad (27)$$

where,

$$|c(\bar{n})| = \sqrt{1 - \lceil \bar{n} \rceil + \bar{n}} \quad (28)$$

$$|a(\bar{n})| = \sqrt{1 - |c(\bar{n})|^2} \quad (29)$$

and $\lceil \bar{n} \rceil$ is the ceiling of $\bar{n}$. We refer to the state of Eq. (27) as *in-between state* as it is a superposition of the two nearest Fock states for a given $\bar{n}$ and it reverts to a Fock state when $\bar{n}$ is an integer.

For our numerical simulations, we create random states that satisfy the following conditions,

$$|\Psi_{\bar{n}}\rangle = \sum_{n=0}^N d_n |n\rangle, \quad (30)$$

$$\sum_{n=0}^N |d_n|^2 = 1, \quad (31)$$

$$\sum_{n=0}^N n |d_n|^2 = \bar{n}, \quad (32)$$

for $d_n \in \mathbb{C}$ and by choosing N and $\bar{n}$. For the two-point prior PDF (15), we verify that for integer $\bar{n}$ the optimal state is the Fock state with the same photon number. For real $\bar{n}$, our numerical results support the conjecture that the optimal state has the form of the state (27), up to a phase, i.e.,

$$|\Phi'_{\bar{n}}\rangle = e^{i\phi\bar{n}} |\Phi_{\bar{n}}\rangle. \quad (33)$$

By removing the global phase, the state in Eq. (33) can be rewritten as,

$$|\Phi'_n\rangle = |a(\bar{n})|[\lceil\bar{n}\rceil - 1] + |c(\bar{n})|e^{i\phi}[\lceil\bar{n}\rceil], \quad (34)$$

which in turn is (up to a global phase and by setting $\phi = \theta_2 - \theta_1$) equal to the general state $\sqrt{1 - |c(\bar{n})|^2}e^{i\theta_1}[\lceil\bar{n}\rceil - 1] + |c(\bar{n})|e^{i\theta_2}[\lceil\bar{n}\rceil]$. However, the phase operator in Eq. (33) commutes with the pure-loss channel [16]. Therefore, the state of Eq. (27) interacts with the pure-loss, i.e., it picks up the unknown variable τ , and then the phase operator is applied on the channel's output. Since the MMSE is invariant under unitary operators that do not depend on the unknown variable and were applied after the unknown variable has been inscribed on the state [17], the MMSE for the states of Eqs. (27) and (34) is the same.

We note here that we performed our numerical computations by letting the Fock coefficients of our random states to be complex numbers (see Eqs. (30)). In Fig. 1 we show an example of our numerical results for $N = 4$ (i.e. a superposition of five Fock states) and 200 states for each fixed $\bar{n}$. The blue lines represent the MSE of the states of Eq. (27), the black solid dots represent random pure states with fixed mean photon number $\bar{n}$, and the black empty dots represent Fock states. We see that the black dots are above the blue lines.

A. Sub-optimal measurement

For the in-between states of Eq. (27), $\hat{B}$ is not diagonal and we were not able to find a useful (in the sense of feasibility) optimal measurement via analytical and numerical means. However, we examine the behavior of PNR detection: We find that, albeit sub-optimal (except if $\bar{n} \in \mathbb{N}$, where it becomes optimal), its values are close to the conjectured MMSE (i.e. the MSE for the states of (27)). Also, we see that the MSE for PNR detection comes closer to the conjectured MMSE as the differences $|\tau_0 - \tau_1|$ and $|q_0 - q_1|$ become smaller; as said differences go to zero, the two-point prior PDF approaches a delta function, for which all MSEs are zero. We show our results in Fig. 1. For PNR detection, the MSE $\tilde{\delta}$ can be computed as follows,

$$\tilde{\delta} = \int_0^1 d\tau \sum_{k=0}^n P(\tau) P(k|\tau) (\Pi_k - \tau)^2, \quad (35)$$

where k is the detected photon number, $P(\tau)$ is the prior probability, $P(k|\tau)$ is the conditional probability,

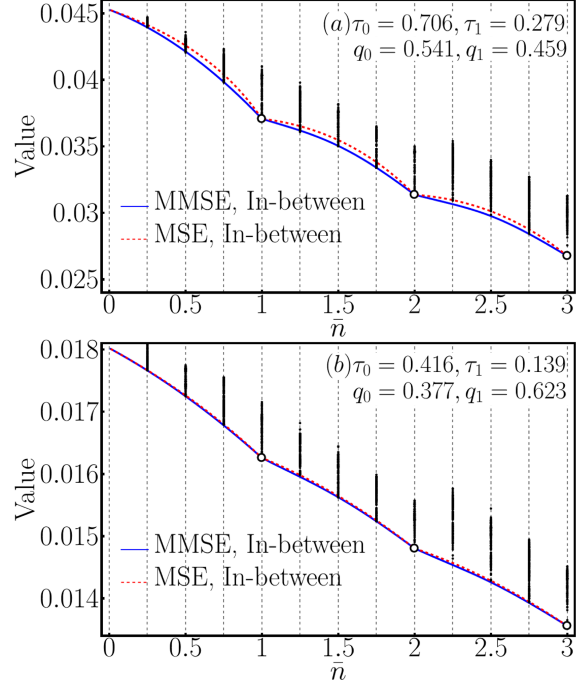


FIG. 1. For the two-point prior PDF, we plot the MSE of states with mean photon number $\bar{n}$. The blue lines represent the in-between states of Eq. (27), the red dashed lines represent the MSE by choosing PNR detection, the black solid dots represent random states with fixed $\bar{n}$, and the black empty dots represent Fock states. (a) Transmissivities $\tau_0 = 0.706$, $\tau_1 = 0.279$, and prior PDF with parameters $q_0 = 0.541$, $q_1 = 0.459$; (b) transmissivities $\tau_0 = 0.416$, $\tau_1 = 0.139$, and prior PDF with parameters $q_0 = 0.377$, $q_1 = 0.623$.

ity, and Π_k is the expected value,

$$\begin{aligned} \Pi_k &= \int_0^1 d\tau P(\tau|k) \tau \\ &= \frac{\int_0^1 d\tau P(k|\tau) P(\tau) \tau}{\int_0^1 d\tau P(k|\tau) P(\tau)}, \end{aligned} \quad (36)$$

where we used Bayes' rule for $P(\tau|k)$. For the two-point prior distribution of Eq. (15) and the output state $\hat{\rho}_n(\tau)$ if the input is the state of Eq. (27), we find,

$$\Pi_k = \frac{q_0 \tau_0 P(k|\tau_0) + (1 - q_0) \tau_1 P(k|\tau_1)}{q_0 P(k|\tau_0) + (1 - q_0) P(k|\tau_1)}, \quad (37)$$

$$P(k|\tau) = \langle k | \hat{\rho}_n(\tau) | k \rangle = P_1^{(k,\tau)} + P_2^{(k,\tau)} \quad (38)$$

where,

$$P_1^{(k,\tau)} = |c(\bar{n})|^2 (1-\tau)^{\lceil \bar{n} \rceil - k} \tau^k \binom{\lceil \bar{n} \rceil}{k},$$

$$P_2^{(k,\tau)} = (1 - |c(\bar{n})|^2) (1-\tau)^{n-k-1} \tau^k \binom{\lceil \bar{n} \rceil - 1}{k},$$

and $0 \leq k \leq \lceil \bar{n} \rceil$ with the convention $\binom{\lceil \bar{n} \rceil - 1}{0} = 0$.

VI. BETA DISTRIBUTION AS PRIOR PDF

We now consider the prior PDF of Eq.(16). For that case, we can analytically calculate the MSE for a Fock state as input to the pure-loss channel (see Appendix C),

$$\delta = \frac{\alpha\beta}{(\alpha + \beta)(\alpha + \beta + 1)(\alpha + \beta + \bar{n})}. \quad (39)$$

For non-integer mean photon number and non-Fock states, we compute the MSE numerically using the same methods as in Section V. The numerical results indicate that the optimal input states with non-integer mean photon number, have the form of Eq. (27), which means that for $\bar{n} \in \mathbb{N}$ the optimal states are Fock states. We show our results in Fig. 2.

VII. DISCUSSION AND CONCLUSIONS

For the task of sensing the transmissivity of a pure-loss channel, the QFI for a Fock state $|n\rangle$ has been found to be [5][18],

$$F(\tau) = \frac{n}{\tau(1-\tau)}, \quad (40)$$

also $F(\tau)$ is attainable by photon-counting. Let us discuss and revisit the main results of present work.

Two-point prior PDF: For the prior PDF of Eq. (15), the Fisherian-inspired, Bayesian bound of Eq. (3) gives,

$$J_E^{-1} = \frac{\tau_1(1-\tau_1) + q(\tau_0 - \tau_1)(1-\tau_0 - \tau_1)}{n} \quad (41)$$

which in general is non-negative, including the case where $q = 0$ (or $q = 1$) for which the prior PDF becomes a delta function and therefore the Personick MMSE gives zero. Therefore, even if it is known that the Fock state is optimal in the Fisherian approach, the Fisherian-inspired bounds have room for improvement begging the question of what the optimal states

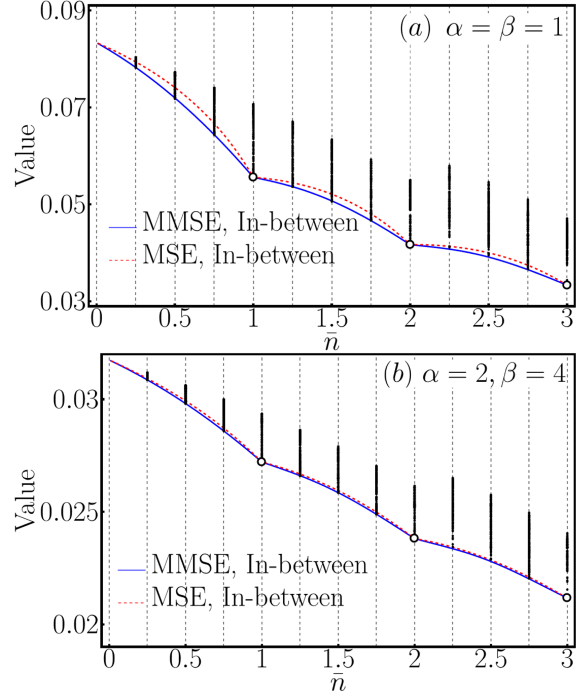


FIG. 2. For the beta prior PDF, we plot the MSE as a function of $\bar{n}$. The blue lines represent the MSE of in-between states of Eq. (27), the red dashed lines represent the MSE for PNR detection, the black solid dots represent 200 random states for each fixed $\bar{n}$; the black empty dots represent Fock states. For all random states we used random superpositions of the first five Fock states. (a) parameters $\alpha = \beta = 1$ (uniform distribution); (b) parameters $\alpha = 2, \beta = 4$.

and measurements are, that we considered in this work. Surprisingly, we found that the optimal state is unique and is the Fock state, while the optimal measurement is photon-counting. Moreover, we conjectured the form of the optimal states for $\bar{n} > 0$ and we examined the performance of photon-counting for such states. For completeness, we also give the formulas,

$$J_D = n \left[\frac{q}{\tau_0(1-\tau_0)} + \frac{1-q}{\tau_1(1-\tau_1)} \right] \quad (42)$$

for a Fock state $|n\rangle$ as input, while for J_P the integral is problematic since it involves derivative and inverse on the delta function. However, we believe that J_P diverges to $+\infty$, meaning that $J_B \rightarrow +\infty$, which results to a in general non-attainable bound. We present an example of our computations in Fig. 3, where instead of J_B^{-1} , we plot J_D^{-1} which is higher than the Personick MMSE.

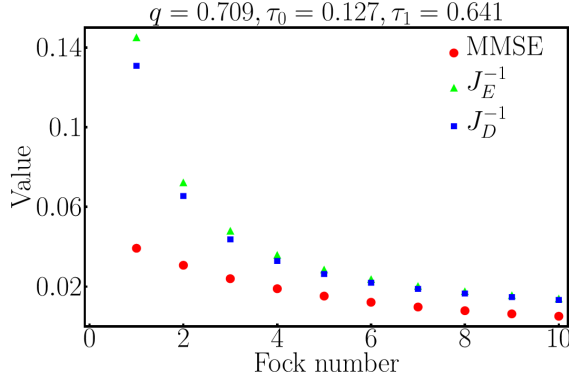


FIG. 3. For Fock states and the two-point prior PDF with parameters $q = 0.79$, $\tau = 0.127$, $\tau = 0.641$, we plot the Fisherian-inspired Bayesian bounds (triangles and squares) and the MMSE. The MMSE is the lowest one and by definition attainable.

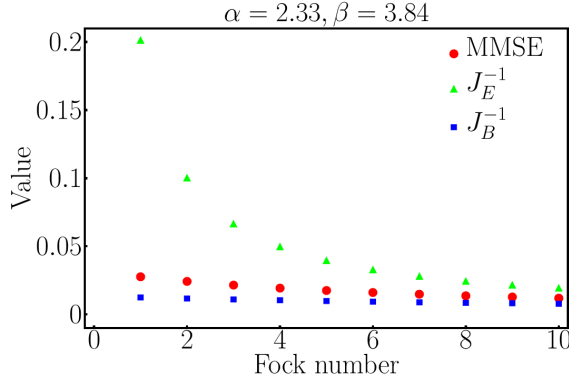


FIG. 4. For Fock states and the beta prior PDF with parameters $\alpha = 2.33$, $\beta = 3.84$, we plot the Fisherian-inspired Bayesian bounds (triangles and squares) and the MMSE. The MMSE is the lowest attainable quantity, emphasizing the fact that J_B^{-1} is not attainable by a measurement.

Beta prior PDF: For that case we find,

$$J_E^{-1} = \frac{\alpha\beta}{n(\alpha + \beta)(\alpha + \beta + 1)} \quad (43)$$

$$J_B = \mathcal{C}(\alpha, \beta) [nf(\alpha, \beta) + g(\alpha, \beta)] \quad (44)$$

where $\mathcal{C}(\alpha, \beta) = \mathcal{B}(\alpha - 2, \beta - 2)/\mathcal{B}(\alpha, \beta)$, $f(\alpha, \beta) = (\alpha - 2)(\beta - 2)$, and $g(\alpha, \beta) = (\alpha - 1)(\beta - 1)(\alpha + \beta - 4)$. The computation of J_P required for Eq. (44) converges only if $\alpha > 2$ and $\beta > 2$, reflecting the problematic nature of the Fisherian-inspired bounds. We present an example of our computations in Fig. 4.

We note that for the beta prior PDF, we were not able to prove analytically that the Fock states or the in-between states of Eq. (27) states are optimal, even though our numerical computations support such conjecture. To prove this, following the lower bound approach of Eq. (21) and for $\bar{n} \in \mathbb{N}$, one should show that the lower bound is a function of the mean input photon number only; That would mean that even if states other than Fock, albeit with the same mean-photon number, satisfy Eq. (22), these states give the same MMSE. Another route would be to prove that for all prior PDFs, Eq. (22) necessarily gives $[\hat{\rho}(\tau), \hat{\rho}(\tau')] = 0$, however we note that this seems to be an onerous task (and probably non-true).

Aforesaid difficulties make the genuine Bayesian approach interesting and non-trivial. Future directions, could include the study of optimal Gaussian states and optimal measurements for sensing the transmissivity of a pure-loss channel. For that case, the difficulty in calculating the MMSE is found in the fact that even if $\hat{\rho}(\tau)$ is Gaussian, $\hat{\Gamma}_0$ is not. Other future directions are sensing the transmissivity including a thermal environment, and sensing multiple parameters of a general Gaussian transformation (or channel) by devising algorithms in the same spirit as in [11].

-
- [1] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, *Nature Communications* **8**, 15043 (2017).
 - [2] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, *Nature Communications* **3**, 1063 (2012).
 - [3] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *Nature Physics* **7**, 406 (2011).
 - [4] R. Nair, *Phys. Rev. Lett.* **121**, 230801 (2018).
 - [5] G. Adesso, F. Dell'Anno, S. De Siena, F. Illuminati, and L. A. M. Souza, *Phys. Rev. A* **79**, 040305 (2009).
 - [6] A. Monras and M. G. A. Paris, *Phys. Rev. Lett.* **98**, 160401 (2007).
 - [7] Z. Gong, N. Rodriguez, C. N. Gagatsos, S. Guha, and B. A. Bash, *IEEE Journal of Selected Topics in Signal Processing*, 1 (2022).
 - [8] R. Jonsson and R. D. Candia, *Journal of Physics A: Mathematical and Theoretical* **55**, 385301 (2022).
 - [9] S. Personick, *IEEE Transactions on Information Theory* **17**, 240 (1971).
 - [10] J. Rubio and J. Dunningham, *Phys. Rev. A* **101**, 032114 (2020).
 - [11] K. K. Lee, C. N. Gagatsos, S. Guha, and A. Ashok, *IEEE Journal of Selected Topics in Signal Processing*

- , 1 (2022).
- [12] H. L. Van Trees, K. L. Bell, and Z. Tian, *Detection estimation and modulation theory, part I*, 2nd ed. (Wiley-Blackwell, Hoboken, NJ, 2013).
 - [13] S. Pirandola, *Phys. Rev. Lett.* **106**, 090504 (2011).
 - [14] R. Nair, *Phys. Rev. A* **84**, 032312 (2011).
 - [15] D.-W. Chang, *Journal of Mathematical Analysis and Applications* **237**, 721 (1999).
 - [16] A. S. Holevo, *Quantum systems, channels, information*, De Gruyter Studies in Mathematical Physics (De Gruyter, Berlin, Germany, 2012).
 - [17] This is proven by using Eqs. (13) and (14) and showing that the transformed $\hat{\Gamma}_k$ operators are $e^{i\phi\hat{n}}\hat{\Gamma}_ke^{-i\phi\hat{n}}$ and the transformed $\hat{B}$ operator is $e^{i\phi\hat{n}}\hat{B}e^{-i\phi\hat{n}}$. Then, using Eq. (12) and the cyclic permutation property of operators within a trace one can show that the unitary will have no effect on the MMSE.
 - [18] Taking into account the different parameterization between this work and [5], the results match.
 - [19] W. Magnus, *Communications on Pure and Applied Mathematics* **7**, 649 (1954).

Appendix A: Fock-diagonal outputs of pure-loss channel necessarily correspond to Fock-diagonal inputs

Although the result of this appendix may be known, we could not find a formal proof. Thus, we provide one here.

We consider a generic single-mode density operator written on the Fock basis,

$$\hat{\rho}(t) = \sum_{n,m} c_{n,m}(t) |n\rangle\langle m|. \quad (\text{A1})$$

Then, Eq. (8) can be written as,

$$\begin{aligned} \dot{c}_{m,n} = & \frac{\gamma}{2} [2\sqrt{(m+1)(n+1)}c_{m+1,n+1} \\ & - (m+n)c_{m,n}], \end{aligned} \quad (\text{A2})$$

where $c_{m,n}(t) \equiv c_{m,n}$ are time-dependent (unless otherwise specified) and the dot represents derivative with respect to time.

We can rewrite Eq. (A2) as $\dot{\mathbf{c}} = \mathbf{A}\mathbf{c}$, where $\mathbf{c}$ is a vectorized version of the matrix with elements $c_{m,n}$, and $\mathbf{A}$ is a matrix which by inspection of Eq. (A2) is upper triangular. For that reason, we choose to rewrite $c_{m,n}$ as $c_{n+l,n}$, i.e.,

$$\begin{aligned} \dot{c}_{n+l,n} = & \gamma\sqrt{(n+1)(n+l+1)}c_{n+l+1,n+1} \\ & - \gamma\left(n + \frac{l}{2}\right)c_{n+l,n}, \end{aligned} \quad (\text{A3})$$

where $l = 0, 1, 2, \dots$, and $l = 0$ gives the time evolution of the diagonal elements of Eq. (A1). For any fixed l , we denote a vector,

$$\mathbf{c}^{(l)} = \begin{pmatrix} c_{0+l,0} \\ c_{1+l,1} \\ \vdots \\ c_{n+l,n} \\ \vdots \end{pmatrix}. \quad (\text{A4})$$

Under these considerations, we write Eq. (A3),

$$\begin{pmatrix} \dot{\mathbf{c}}^{(l=0)} \\ \dot{\mathbf{c}}^{(l=1)} \\ \dot{\mathbf{c}}^{(l=2)} \\ \vdots \end{pmatrix} = \mathbf{A} \begin{pmatrix} \mathbf{c}^{(l=0)} \\ \mathbf{c}^{(l=1)} \\ \mathbf{c}^{(l=2)} \\ \vdots \end{pmatrix}, \quad (\text{A5})$$

where $\mathbf{A}$ is a block-diagonal matrix or the form,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^{(l=0)} & & & \\ & \mathbf{A}^{(l=1)} & & \\ & & \mathbf{A}^{(l=2)} & \\ & & & \ddots \end{pmatrix}, \quad (\text{A6})$$

while all empty entries are equal to zero. The matrix $\mathbf{A}^{(l)}$ has the form,

$$\mathbf{A}^{(l)} = \mathbf{A}_D^{(l)} + \mathbf{A}_{UD}^{(l)}, \quad (\text{A7})$$

where the matrix $\mathbf{A}_D^{(l)}$ is diagonal,

$$\mathbf{A}_D^{(l)} = -\gamma \mathbf{D}, \quad (\text{A8})$$

where $\mathbf{D} = \text{diag} [\frac{l}{2} + 0, \frac{l}{2} + 1, \frac{l}{2} + 2, \dots]$, and the matrix $\mathbf{A}_{UD}^{(l)}$ is upper-diagonal (i.e. the only non-zero elements are in its upper diagonal),

$$\mathbf{A}_{UD}^{(l)} = \gamma \mathbf{D}_u, \quad (\text{A9})$$

where $\mathbf{D}_u = \text{udiag} [\sqrt{l+1}, \sqrt{2(l+2)}, \sqrt{3(l+3)}, \dots]$, where by udiag we denote an upper diagonal matrix.

The solution to each differential equation (i.e. for each l) represented in Eq. (A5) is given by,

$$\mathbf{c}^{(l)} = \exp(\mathbf{A}^{(l)} t) \mathbf{c}_0^{(l)}, \quad (\text{A10})$$

where $\mathbf{c}^{(l)}$ is time-dependent and $\mathbf{c}_0^{(l)} \equiv \mathbf{c}^{(l)}(t=0)$ is given by the initial condition of Eq. (9) on the Fock basis.

We note that since $\mathbf{A}^{(l)}$ is upper-triangular, the matrix $\exp(\mathbf{A}^{(l)} t)$ is upper triangular as well. Let us work out the specific form of $\exp(\mathbf{A}^{(l)} t)$. We have,

$$\exp(\mathbf{A}^{(l)} t) = \exp(\mathbf{A}_D^{(l)} t + \mathbf{A}_{UD}^{(l)} t). \quad (\text{A11})$$

To apply the Zassenhaus formula [19], we first calculate the commutator,

$$\mathbf{C}_N = [\mathbf{A}_D^{(l)}, \mathbf{C}_{N-1}], \quad (\text{A12})$$

with $\mathbf{C}_0 = \mathbf{A}_{UD}^{(l)}$. We find,

$$\mathbf{C}_N = \gamma^N \mathbf{A}_{UD}^{(l)}. \quad (\text{A13})$$

We also find,

$$[\mathbf{A}_{UD}^{(l)}, \mathbf{C}_N] = 0. \quad (\text{A14})$$

Applying the Zassenhaus formula we get,

$$\begin{aligned} \exp(\mathbf{A}^{(l)} t) &= \exp(\mathbf{A}_D^{(l)} t) \\ &\times \exp\left(\sum_{j=1}^{\infty} \frac{t^j}{j!} (-1)^{j+1} \gamma^{j-1} \mathbf{A}_{UD}^{(l)}\right), \end{aligned}$$

which gives,

$$\exp(\mathbf{A}^{(l)} t) = \exp(\mathbf{A}_D^{(l)} t) \exp(\mathbf{S}^{(l)}), \quad (\text{A15})$$

where,

$$\mathbf{S}^{(l)} = (1 - e^{-t\gamma}) \mathbf{D}_u,$$

from which we find the ij (we enumerate as $i, j = 1, 2, \dots$ everywhere) element of $\exp(\mathbf{S}^{(l)})$ for $j \geq i$,

$$\exp(\mathbf{S}^{(l)})_{ij} = \frac{(1 - e^{-t\gamma})^{j-i}}{(j-i)!} \prod_{r=i}^{j-1} \sqrt{r(l+r)},$$

with the convention $\prod_{r=i}^{i-1} \sqrt{r(l+r)} = 1$. For $j < i$ we find $\exp(\mathbf{S}^{(l)})_{ij} = 0$. Since $\mathbf{A}_D^{(l)}$ is diagonal, the ij element of $\exp(\mathbf{A}_D^{(l)})$ is,

$$\exp(\mathbf{A}_D^{(l)})_{ij} = \delta_{ij} \exp\left[-t\gamma \left(\frac{l}{2} + i - 1\right)\right].$$

Multiplying a diagonal matrix with an upper-triangular matrix as per the left hand side of Eq. (A15), we get an upper-triangular matrix, as expected. Also, all elements of $\exp(\mathbf{A}^{(l)} t)$ are non-zero for $l \neq 0$ (i.e. for non-diagonal entries of the density operator on the Fock basis), unless when $t\gamma \rightarrow 0$, which is the case of no-interaction ($\tau = e^{-t\gamma} = 1$). We demand Eq. (A10) to be element-wise equal to zero for $l \neq 0$, i.e.,

$$\mathbf{c}^{(l \neq 0)} = \exp(\mathbf{A}^{(l \neq 0)} t) \mathbf{c}_0^{(l \neq 0)} = 0. \quad (\text{A16})$$

In Eq. (A16) the upper-triangular matrix $\exp(\mathbf{A}^{(l)} t)$ multiplies the vector $\mathbf{c}_0^{(l)}$, rendering all coefficients of Eq. (A4), starting from the bottom and moving upwards equal to zero. Therefore, if the time-evolved off-diagonal coefficients of Eq. (A1) are zero, then the input density operator is Fock-diagonal.

If $\tau = e^{-t\gamma} = 1$, then the input state is retrieved at the output and trivially such an output is Fock-diagonal if the input is Fock-diagonal.

The converse, i.e., if the input is Fock-diagonal, then the output is also Fock-diagonal is a well known fact referred to as *gauge* or *phase* invariance of the pure-loss channel.

Appendix B: The optimal projective measurements for Fock states under the two-point prior and beta PDFs

For input state $\hat{\rho}_0 = |n\rangle\langle n|$, Eq. (11) gives,

$$\hat{\rho}(\tau) = \sum_{l=0}^n e_l^{(n)}(\tau) |n-l\rangle\langle n-l|, \quad (\text{B1})$$

where $e_l^{(n)}(\tau)$ is given in Eq. (25). For the prior PDF of Eq. (15), from Eq. (13) we find,

$$\begin{aligned}\hat{\Gamma}_k &= q \sum_{l=0}^n e_l^{(n)}(\tau_0) \tau_0^k |n-l\rangle \langle n-l| \\ &\quad + (1-q) \sum_{l=0}^n e_l^{(n)}(\tau_1) \tau_1^k |n-l\rangle \langle n-l|. \quad (\text{B2})\end{aligned}$$

Since $\hat{\Gamma}_0$ is Fock-diagonal we find,

$$\begin{aligned}e^{-z\hat{\Gamma}_0} &= \sum_{l=0}^n \exp \left[-z \left(q e_l^{(n)}(\tau_0) + (1-q) e_l^{(n)}(\tau_1) \right) \right] \\ &\quad \times |n-l\rangle \langle n-l| \quad (\text{B3})\end{aligned}$$

From Eqs. (B2) for $k=1$ and (B3), Eq. (14) gives Eq. (26) which is a closed form relation expressed as a finite sum.

The calculation is similar if we use the beta distribution of Eq. (16). In fact, for every density operator of Eq. (B1) and for *any* prior PDF $P(\tau)$ we can write,

$$\begin{aligned}\hat{\Gamma}_k &= \int_0^1 d\tau P(\tau) \tau^k \sum_{l=0}^n e_l^{(n)}(\tau) |n-l\rangle \langle n-l| \\ &= \sum_{l=0}^n g_l^{(n,k)}[P] |n-l\rangle \langle n-l|, \quad (\text{B4})\end{aligned}$$

where,

$$g_l^{(n,k)}[P] = \int_0^1 d\tau P(\tau) \tau^k e_l^{(n)}(\tau). \quad (\text{B5})$$

The square bracket denotes the functional dependence of $g_l^{(n,k)}[P]$ on $P(\tau)$.

From Eqs. (14) and (B4) for $k=1$, we get,

$$\hat{B} = \sum_{l=0}^n \frac{g_l^{(n,1)}[P]}{g_l^{(n,0)}[P]} |n-l\rangle \langle n-l|, \quad (\text{B6})$$

meaning that for a Fock state input, projection on Fock states is always the optimal measurement.

Appendix C: The MMSE for Fock states under the two-point prior and beta PDFs

Let us consider the two-point prior PDF of Eq. (15). Using Eq. (B2) for $k=2$ we find,

$$\text{tr}(\hat{\Gamma}_2) = q\tau_0^2 + (1-q)\tau_1^2. \quad (\text{C1})$$

From Eqs. (26) and (B2) for $k=1$ we find,

$$\text{tr}(\hat{B}\hat{\Gamma}_1) = \sum_{l=0}^n \frac{[q\tau_0 e_l^{(n)}(\tau_0) + (1-q)\tau_1 e_l^{(n)}(\tau_1)]^2}{q e_l^{(n)}(\tau_0) + (1-q) e_l^{(n)}(\tau_1)},$$

which together with Eqs. (12) and (C1), gives Eq. (24).

The calculation is similar if we use the beta distribution of Eq. (16) to derive Eq. (39). In fact, for every density operator of Eq. (B1) and for *any* prior PDF $P(\tau)$, using Eqs. (12) and (B4) for $k=1, 2$ we find,

$$\delta^{(n)}[P] = \sum_{l=0}^n \left[\left(g_l^{(n,2)}[P] \right)^2 - \frac{\left(g_l^{(n,1)}[P] \right)^2}{g_l^{(n,0)}[P]} \right],$$

where the square brackets in $\delta^{(n)}[P]$ denote the functional dependence on $P(\tau)$.