Parameter estimation by multi-channel photon counting

Alexander Holm Kiilerich and Klaus Mølmer
Department of Physics and Astronomy, Aarhus University,
Ny Munkegade 120, DK 8000 Aarhus C. Denmark
(Dated: November 14, 2014)

The physical parameters governing the dynamics of a light emitting quantum system can be estimated from the photon counting signal. The information available in the full detection record can be analysed by means of the distribution of waiting times between detection events. Our theory allows calculation of the asymptotic, long time behaviour of the sensitivity limit, and it applies to emission processes with branching towards different final states accompanied by the emission of distinguishable photons. We illustrate the theory by application to a laser driven Λ-type atom.

PACS numbers: 03.65.Wj, 03.65.Yz, 02.50.Tt, 42.50.Lc

I. INTRODUCTION

Atoms and atom-like systems with discrete energy states are widely used for precision measurements of time and frequency and as sensitive probes of fields or other influences on the system behaviour. The random character of measurements on quantum systems fundamentally limits the information achievable, but quantum states with squeezed uncertainty of particular observables, and entangled states of multi-particle systems have been identified as particularly sensitive initial states for (repeated) single-shot experiments, see e.g. [1, 2].

Rather than many repeated experiments we have the situation in mind of a single quantum system probed continuously over time. One must then take the measurement back action into account at all measurement steps, and this is conveniently done in the quantum trajectory formalism. This provides, conditioned on the measurement record [3–5], both the state of the quantum system and, via Bayes’ rule, the probabilities of different candidate values of the estimated parameter. If the system is subject to damping and decoherence, and behaves in an ergodic manner, one may regard data obtained at sufficiently well separated moments of time as statistically independent. Continuous probing of the same system for a long time \(T\) can hence be thought of as a number of \(N\) independent experiments with \(N \propto T\), and we expect an estimation error scaling asymptotically as \(1/\sqrt{T}\).

To confirm this expectation and to identify the quantitative performance of continuous probing we shall address the Cramér-Rao bound (CRB) [6],

\[
|\Delta S(\theta)|^2 \geq \frac{1}{F(\theta)},
\]

which expresses the lower limit of the statistical variance \(|\Delta S(\theta)|^2\) of any unbiased estimator for an unknown quantity \(\theta\) by the Fisher information,

\[
F(\theta) = -\sum_D \frac{\partial^2 \ln L(D|\theta)}{\partial \theta^2} L(D|\theta)
\]

where \(L(D|\theta)\) in Eq. (2) is the likelihood to obtain measurement data \(D\) conditioned on the value \(\theta\). For \(n_{\text{rep}}\) repeated experiments, Eq. (1) is written with an extra factor \(1/n_{\text{rep}}\), and the bound applies in the limit of \(n_{\text{rep}} \gg 1\). In our case, however, \(D\) represents a single, time-dependent detection record, and the asymptotic convergence of our estimate should follow from the probing time dependence of \(F(\theta)\) in the long time limit.

In this article we consider detection by photon counting of the radiation emitted by a quantum light source. For a closed two-level transition, the discrete waiting times between detection events form independent and identically distributed stochastic variables, and we have previously shown [7] that this simplifies the evaluation of the Fisher information and the Cramér-Rao bound (CRB). Here, we generalize the approach of [7] to the case of multi-level systems with distinguishable emission processes and branching of the decay towards different final states. This situation is exemplified by the Λ-system depicted in Fig. 1, with an excited state from which spontaneous decay occurs towards two different ground states. Since the decay processes leave the atom in different states, subsequent time intervals between detector clicks are not independent. The purpose of this article is to derive a theory that allows calculation of the Fisher information and the Cramér-Rao bound for parameter estimation with two-channel (and more general multi-channel) counting signals.

The article is outlined as follows. In Sec. II, we discuss how single-channel and multi-channel photon counting records can be reorganized as the sampling of uncorrelated stochastic variables and how the Fisher information can be calculated from the distribution of waiting times between detections in different channels. In Sec. III we present a master equation analysis of the theoretical waiting time distribution functions that allow practical calculation of the Fisher information. In Sec. IV, we present the filter function that should be applied to multi-channel measurement data records to achieve parameter estimates that reach the Cramér-Rao bound. In Sec. V we show results for the Λ-system in Fig. 1 and in Sec. VI, we conclude the analysis.
II. SINGLE-CHANNEL AND MULTI-CHANNEL COUNTING SIGNALS

A photon counting detection record contains the discrete times of detection events \( D = \{ t_i \} \), and, if the emitter always jumps to the same state when a photon is detected, measurement intervals \( \tau_i = t_{i+1} - t_i \) between detector clicks are independent and identically distributed stochastic variables. A data record with \( N + 1 \) count events, thus, yields \( N \) independent realizations \( \tau_i \) of the waiting time probability distribution \( \nu(\tau) \). The waiting time distribution is a function of the parameters governing the system dynamics and it can be calculated by master equation or quantum trajectory theory. The governing system dynamics and it can be calculated by master equation or quantum trajectory theory. The event counting record contains the distribution \( \{ \tau_i \} \) sample this distribution, and if the variance of the total number of detection events grows only linearly with time, the number \( n_i \) of registered waiting times in short intervals \([\tau_i, \tau_{i+1}]\) will be a Poisson distributed stochastic variable with mean value \( \overline{n}_i \). The likelihood function for a given data record thus factorizes as the product of the Poisson distributions for the various \( n_i \)'s, and the Fisher information Eq. (2) can be evaluated by carrying out the sum over the distributions pertaining to each interval \( i \),

\[
F_p(\theta) = \sum_i \frac{1}{n_i(\theta)} \left( \frac{\partial n_i(\theta)}{\partial \theta} \right)^2. \tag{3}
\]

See, e.g., [8] [9] for similar arguments applied to high-resolution spatial measurements by scattering of coherent light or to probing of the motion of a Bose condensate. If we know how \( \overline{n}_i \) depends on the unknown parameter \( \theta \), and if we assume the limit of infinitesimal intervals with average counts \( \overline{n}_i(\theta) = \overline{n}(\tau, \theta) \Delta \tau \), we can evaluate Eq. (3) as an integral,

\[
F_p(\theta) = \int \frac{1}{\overline{n}(\tau, \theta)} \left( \frac{\partial \overline{n}(\tau, \theta)}{\partial \theta} \right)^2 d\tau. \tag{4}
\]

Since the total number of detection events, and, hence, \( \overline{n}(\tau, \theta) \) is proportional to the duration \( T \) of the experiment, the Fisher information is proportional to \( T \), and

we conclude from Eq. (1) that the estimation error, indeed, decreases as \( \sim 1/\sqrt{T} \).

In the long time limit, where \( \overline{n}_i \gg 1 \), the Poisson distributions are well approximated by Gaussian functions, and for uncorrelated Poisson distributions, the distribution of the vector of detection events \( \mathbf{n} = (n_1, n_2, \ldots) \) is approximately given by \( \propto \exp(-\frac{1}{2}(\mathbf{n} - \overline{\mathbf{n}})\Sigma^{-1}(\mathbf{n} - \overline{\mathbf{n}})^T) \), where \( \Sigma \) is the diagonal covariance matrix with elements \( \Sigma_{ii} = \text{Var}(n_i) = \overline{n}_i \), and \( T \) denotes the transpose of the vector of arguments. The Fisher information for a multi-variate normal distribution is known, and when \( \overline{n}_i \) grows linearly with time, the leading term yields

\[
F(\theta) = (\partial \overline{n}/\partial \theta)^T \Sigma^{-1}(\partial \overline{n}/\partial \theta),
\]

reproducing the exact expression \( F_p(\theta) \) in the case of Poisson statistics (3).

Now, the sum of uncorrelated Poisson distributed stochastic variables is itself Poisson distributed, while it is well known that the total number of emitted photons, \( N = \sum n_i \), from quantum light sources may show sub- or super-Poissonian counting statistics \([10]\). As long as the variance in the total number of detection events scales less than quadratically with the mean value, the small fraction of events \( n_i \ll N \) of specified duration will still show Poisson statistics, but we should correct Eq. (3) to account for the case of \( V \equiv \text{Var}(N) \neq N \). This is most readily done within the Gaussian approximation, since the constraint of a definite value of \( \text{Var}(N) \) can be simply accommodated by multiplying the uncorrelated Gaussian distribution with the function \( \propto \exp(-\frac{1}{2}(N - \overline{N})/Q) \), where \( 1/Q = 1/V - 1/N \). The resulting multi-variate Gaussian distribution is thus characterized by the modified covariance matrix \( \Sigma^{-1} = \Sigma^{-1} + (1/Q)\Pi \), where \( \Pi \) is the matrix with all elements equal to unity. The Fisher information with the modified covariance matrix \( \Sigma \) yields,

\[
F(\theta) = F_p(\theta) + \frac{N - V}{NV} \left( \frac{\partial V}{\partial \theta} \right)^2. \tag{5}
\]

By taking the total number of counts explicitly into account, the Fisher information increases if this number fluctuates less than the Poisson distribution assumed in \([3]\), while it acquires a smaller values if the distribution is super-Poissonian. Ref. [11] did not take this effect into account. This was justified by the focus in that article on a saturated transition, where the distribution of waiting times yields much more information than the total count, and where the first term in (5) therefore completely dominates the Fisher information. In general, and in the examples studied in Sec. VI in this article, the full count statistics term cannot be ignored.

Let us now turn to the case of signals from quantum emitters observed by photon counters that distinguish between different decay channels, e.g., by making use of the polarization or frequency of the emitted photons. For generality we assume that there are \( M \) such channels (for the \( \Lambda \)-system in Fig. 1, \( M = 2 \)). Our analysis is restricted to the case for which detection of a photon in channel \( m \) accompanies a jump of the emitter into a definite state
\(|\phi_m\rangle\), which is the initial state for the subsequent evolution of the system. This is not a requirement for the Bayesian analysis, but our calculation of the Fisher information relies on definite waiting time distributions after detection in each channel. These waiting time distributions until the next detection event, thus, depend on \(m\), the channel of the most recently detected photon, and we can sort the detection record into lists \(\{\tau_j\}_{mm'}\) containing the duration of time intervals between detection in channel \(m\) followed by subsequent detection in channel \(m'\). These lists, in turn, sample the corresponding waiting time distributions in an independent and uncorrelated manner, and, for our parameter estimation, they retain all the information available in the multi-channel detection record (apart from a modification due to the deviation from Poisson statistics that we shall address separately in the end of Sec. III).

The combinations \(m m'\) define \(M^2\) interval types, and for each \(m m'\), the number \(n_{mm',i}\) denotes the number of waiting times \(\tau\) registered in intervals \([\tau_i, \tau_{i+1}]\). As in the single channel case, if the variance of the total number of detected photons grows only linearly with time, the small fraction of events of specific type and duration are Poisson distributed,

\[
P(n_{mm',i}) = \frac{[\pi_{mm',i}]^{n_{mm',i}}}{n_{mm',i}!} e^{-\pi_{mm',i}},
\]

with a mean value \(\pi_{mm',i}\). The likelihood function in Eq. (2) now factorizes as a product of Poisson distributions, \(L(D|\theta) = \prod_{mm',i} P(n_{mm',i}|\theta)\), and the single channel result (1) is readily generalized,

\[
F_p(\theta) = \sum_{mm'} \int_0^\infty \frac{1}{\pi_{mm'}(\tau,\theta)} \left( \frac{\partial \pi_{mm'}(\tau,\theta)}{\partial \theta} \right)^2 d\tau
\]

where \(\pi_{mm'}(\tau,\theta)\) is the theoretically expected distribution of intervals of type \(mm'\) and duration \(\tau\).

III. WAITING TIME DISTRIBUTIONS

We obtain the distribution functions \(\pi_{mm'}(\tau,\theta)\) by solving effective master equations where the unknown quantity \(\theta\) is one of the Hamiltonian or damping parameters. With the understanding that our results may be finally evaluated and varied with respect to the parameter of interest, we suppress, in this section, the variable \(\theta\) from the equations.

The average behaviour of an atomic quantum system decaying by spontaneous emission of photons into broadband photon reservoirs is described by a master equation of the form (\(\hbar = 1\) [11]),

\[
\frac{d\rho}{dt} = -i[\hat{H}_0, \rho] + \sum_m \left( \hat{C}_m \rho \hat{C}^\dagger_m - \frac{1}{2} \{\hat{C}^\dagger_m \hat{C}_m, \rho\} \right),
\]

where the operators \(\hat{C}_m\) represent jump processes in the atom associated with decay and emission of different, distinguishable kinds of radiation. While decay processes may preserve, e.g., coherences between excited Zeeman states in the ground state after the emission of light of linear or circular polarization, we emphasize that our analysis of the Fisher information is restricted to the case in which a jump \(\hat{C}_m\) puts the system in a definite final state \(|\phi_m\rangle\), from which the dynamics proceeds. This is for example the case for the three level atom, shown in Fig. 1 where the two operators, \(\hat{C}_0 = \sqrt{\Gamma_0} |0\rangle \langle 2|\) and \(\hat{C}_1 = \sqrt{\Gamma_1} |1\rangle \langle 2|\) describe decay into the ground states \(|0\rangle\) and \(|1\rangle\) with rates \(\Gamma_0\) and \(\Gamma_1\), respectively.

With the interpretation of quantum trajectories or Monte Carlo wave functions [7] [12] [13] as the states of dissipative quantum systems conditioned on the outcome of continuous probing of their emitted radiation, it is possible to simulate realistic detection records. The jumps into state \(|\phi_m\rangle\) are governed by the rate \(\langle \hat{C}^\dagger_m \hat{C}_m \rangle\) where the expectation value is calculated as function of time for a given evolving wave function. On average, the stochastically evolving wave functions reproduce the master equation and therefore the average number of these jumps equals the value obtained by the density matrix describing the un-observed quantum system. For probing over long times \(T\), we thus get the average number of jumps into state \(|\phi_m\rangle\), \(\bar{N}_m = \text{Tr}(\hat{C}^\dagger_m \hat{C}_m \rho^s) T\), where \(\rho^s\) is the steady state density matrix solution to the master equation (8).

For the distributions of intervals between detector clicks we now have \(\pi_{mm'}(\tau) = \bar{N}_m w_{mm'}(\tau)\), where \(w_{mm'}(\tau)d\tau\) is the probability that after a jump into \(|\phi_m\rangle\), the next emission event is detected in channel \(m'\) in \([\tau, \tau + d\tau]\). To determine the function \(w_{mm'}(\tau)\), we note that the terms \(\sum_m \hat{C}_m \rho \hat{C}^\dagger_m\) in Eq. (8) account for the feeding of the system ground states associated with the emission process, i.e., they describe terms in the reduced system density matrix, correlated with single-photon excited states of the modes of the radiation field. If the system has just been put into the state \(|\phi_m\rangle\) due to detection of a photon in channel \(m\), the probability that no photon is detected until a certain later time \(\tau\) is equal to the population of the zero-photon component of the combined state of the system and the environment at that time. This is, in turn, given by the trace of the un-normalized density matrix, \(\hat{\rho}\), which evolves from the initial state \(\hat{\rho}(\tau = 0) = |\phi_m\rangle \langle \phi_m|\), omitting the ground state feeding term of the master equation,

\[
\frac{d\hat{\rho}}{dt} = -i[\hat{H}_0, \hat{\rho}] - \frac{1}{2} \sum_m \{\hat{C}^\dagger_m \hat{C}_m, \hat{\rho}\}.
\]

The resulting \(\hat{\rho}|m(\tau)\) is equivalent to the so-called no-jump wave function [13] evolving from the state \(|\phi_m\rangle\) by the non-hermitian Hamiltonian \(\hat{H}_\text{eff} = \hat{H}_0 - \frac{1}{2} \sum_m \hat{C}^\dagger_m \hat{C}_m\). The probability \(w_{mm'}(\tau)d\tau\) that after a detector click at time \(t\) of type \(m\), the next click is of type \(m'\) and occurs in the time interval \([t + \tau, t + \tau + d\tau]\), is now given by

\[
w_{mm'}(\tau)d\tau = \text{Tr}(\hat{C}^\dagger_m \hat{C}_m \rho(\tau))d\tau.
\]
It follows from the master equation that these waiting time distributions are normalized according to
\[
\sum_{m'} \int_0^\infty w_{mm'}(\tau) d\tau = 1.
\]
(11)

With the values thus found theoretically for \(\mathcal{N}_m\) and \(w_{mm'}(\tau)\), we know \(\pi_{mm}(\tau)\), and we can evaluate the Fisher information in (7).

If photons are detected with only finite efficiency \(\eta\), this is equivalent to a fraction \(1 - \eta\) of the quantum jumps passing unnoticed. The corresponding un-normalized state \(\tilde{\rho}\) conditioned on no detection events is then found by including a ground state feeding term, \((1 - \eta)\tilde{C}\rho\tilde{C}^\dagger\), in the no-jump master equation to account for the unobserved emission [7]. In the multi-channel case, if different channels are monitored with detector efficiencies \(\eta_m\), we obtain the no-detected-jump master equation (7).

The first term represents the temporal dependence of the detected emission events, \(\dot{\tilde{\rho}} = -i[\tilde{H}_0, \tilde{\rho}] + \sum_m \left((1 - \eta_m)\tilde{C}_m\rho\tilde{C}_m^\dagger - \frac{1}{2}(\tilde{C}_m\tilde{C}_m^\dagger, \tilde{\rho})\right)\).
\[
\frac{d\tilde{\rho}}{dt} = -i[\tilde{H}_0, \tilde{\rho}] + \sum_m \left((1 - \eta_m)\tilde{C}_m\rho\tilde{C}_m^\dagger - \frac{1}{2}(\tilde{C}_m\tilde{C}_m^\dagger, \tilde{\rho})\right).
\]
(12)

The solutions of this equation for initial states \(|\phi_m(\tau = 0)\rangle = |\phi_m\rangle|\phi_m\rangle\) yield the waiting time distributions between the detected emission events [7],
\[
w_{mm'}(\tau) d\tau = \eta_m' \text{Tr}(\tilde{C}_m\tilde{C}_m^\dagger|\phi_m\rangle\langle\phi_m|) d\tau,
\]
(13)
which are normalized as in Eq. (11). The average number of detected events in channel \(m\) during probing for time \(T\) is \(\mathcal{N}_m = \eta_m\text{Tr}(\tilde{C}_m\tilde{C}_m^\dagger)T\), and with the resulting \(\pi_{mm'}(\tau) = \mathcal{N}_m w_{mm'}(\tau) d\tau\), we can calculate the Fisher information according to (7).

It is instructive to employ \(\pi_{mm'}(\tau) = \mathcal{N}_m w_{mm'}(\tau)\) and expand the derivatives in Eq. (4) or (7) leading to
\[
F_p(\theta) = \sum_{mm'} \mathcal{N}_m \int \frac{(\partial w_{mm'}(\tau)/\partial \theta)^2}{w_{mm'}(\tau)} d\tau + \sum_m \left(\frac{\partial \mathcal{N}_m}{\partial \theta}\right)^2 \mathcal{N}_m.
\]
(14)

The first term represents the temporal dependence of the waiting time distributions on \(\theta\), while the second term represents the information retrieved from the total counts in each decay channel, assuming that these are Poisson distributed with variance \(\mathcal{N}_m\). A term with mixed derivatives drops out because of the normalization of \(w_{mm'}(\tau)\), Eq. (11) and hence independence on \(\theta\) of their integrals over time. The relative significance of the terms highly depends on the physical system.

As an example, consider a system with two open channels, e.g., the \(\Lambda\)-system Fig. [1] with branching ratio \(r : 1 - r\). The temporal distribution of waiting times holds negligible information on \(r\) while the last term in Eq. (14) yields \(F_p(r) = \mathcal{N}/(r(r - 1))\) where \(\mathcal{N}\) is the total number of detections. Indeed, this is equivalent to a single-parameter Bernoulli experiment, e.g., a biased coin toss, for which the Fisher information is well-known.

With reference to our discussion in Sec. [11] leading to Eq. (6) we note that a correction applies in the case of non-Poissonian total count statistics, and that this precisely corresponds to the replacement in Eq. (14) of \(\mathcal{N}_m\) by \(V_m \equiv \text{Var}(N_m)\) in the denominator of the last term,
\[
F(\theta) = \sum_{mm'} \mathcal{N}_m \int \frac{(\partial w_{mm'}(\tau)/\partial \theta)^2}{w_{mm'}(\tau)} d\tau + \sum_m \left(\frac{\partial \mathcal{N}_m}{\partial \theta}\right)^2 \mathcal{N}_m.
\]
(15)

To evaluate our final expression for the Fisher information (15), we need to calculate the variance in the photon counts. This is one of the founding problems of quantum optics [10], and we give here a simple recipe relying on our previous discussion. Consider the duration \(T_N = \sum_{i=1}^N \tau_i\) of \(N\) waiting time intervals. \(T_N\) has a mean value \(\bar{T}_N = N\tau\) and a variance \(\text{Var}(T_N) = N\text{Var}(\tau)\) The corresponding uncertainty in the number of detection events in a definite time interval follows, \(\sqrt{\text{Var}(N)} = \langle dN/dT_N \rangle \cdot \sqrt{\text{Var}(T_N)} = \sqrt{N\sqrt{\text{Var}(\tau)/\tau}}\) where, for an exponential waiting time distribution with \(\text{Var}(\tau) = \tau^2\), we recover the Poissonian statistics.

The quantities \(\tau\) and \(\text{Var}(\tau)\) can be evaluated from the waiting time distribution functions, and in the multi-channel case, the \(k\)th moment of \(\tau\) pertaining to the channel \(m\) is given as
\[
\langle \tau^k \rangle_m = \int \tau^k w_{mm}(\tau) d\tau,
\]
(16)
where \(w_{mm}(\tau)\) is the distribution function for waiting times between photo detection events in the channel \(m\), and is obtained by solving Eq. (12) with efficiencies \(\eta_m\) and \(\eta_m' \neq m\). One may then calculate
\[
V_m = \frac{\text{Var}(\tau)_m}{\tau^2_m} \mathcal{N}_m,
\]
(17)
clearly identifying whether \(\mathcal{N}_m\) follows sub- or super-Poissonian statistics.

**IV. ACHIEVING THE CRAMÉR-RAO BOUND**

The CRB concerns the asymptotic sensitivity, and we assume that the value of \(\theta\) is already known to within a small error \(\delta\theta\) from an offset value which, for convenience, we redefine as \(\theta = 0\). For single channel Poisson distributed counting signals, a simple linear filter achieves the CRB [18][14] and motivates an ansatz for the multi-channel estimator of the form,
\[
S_p(n_{mm'}(\tau)) = \sum_{mm'} \left(\int g_{mm'}(\tau) n_{mm'}(\tau) d\tau + C_{mm'}\right),
\]
(18)
which weighs the actual recorded distributions of waiting times \(n_{mm'}(\tau)\) with gain functions \(g_{mm'}(\tau)\) and constant
offsets $C_{m'n'}$, chosen to ensure the correct mean value and to minimize the statistical variance of the estimator.

We assume that $\delta \theta$ is sufficiently small that the corresponding change in the expected waiting time distribution $\pi_{m'n'}(\tau, \delta \theta)$ in [18] is well represented by a first order Taylor expansion. To cancel the zeroth order terms in [18], we then pick

$$C_{m'n'} = - \int g_{m'n'}(\tau) \pi_{m'n'}(\tau, 0) \, d\tau,$$  \hspace{1cm} (19)

and for data in complete accordance (no noise) with the expected mean, we obtain to first order

$$S_P(\pi_{m'n'}(\tau, \delta \theta)) = \delta \theta \sum_{m'm'} \int g_{m'n'}(\tau) \left( \frac{\partial \pi_{m'n'}(\tau, \theta)}{\partial \theta} \right) \bigg|_{\theta=0} \, d\tau.$$

The uncorrelated, Poisson distributed count signals allow calculation of the variance of the estimator [18].

$$(\Delta S_P)^2 = \sum_{m'm'} \int g_{m'n'}^2(\tau) \pi_{m'n'}(\tau, 0) \, d\tau.$$  \hspace{1cm} (20)

Next, the signal-to-noise ratio,

$$(\text{SNR})^2 = \frac{S_P^2(\pi_{m'n'}(\tau, \delta \theta))}{(\Delta S_P)^2},$$

can be maximized by the Cauchy Schwarz inequality,

$$\langle |\langle v_k(\tau), u_k(\tau) \rangle|^2 \rangle \leq \langle \langle v_k(\tau), u_k(\tau) \rangle \rangle \langle u_k(\tau) \rangle,$$

where $v_k(\tau)$ and $u_k(\tau)$ are functions of the continuous variable $\tau$ and the discrete variable $k = (m'n')$.

Applying the inequality with $v_k(\tau) = \delta \theta \frac{\partial \pi_{m'n'}(\tau, \delta \theta)}{\partial \delta \theta} \bigg|_{\delta \theta=0} \pi_{m'n'}^{-1}(\tau, 0)$ and $u_k(\tau) = g_{m'n'}(\tau) \pi_{m'n'}^{1/2}(\tau, 0)$, we obtain

$$(\text{SNR})^2 \leq (\delta \theta)^2 F_P(\theta)$$  \hspace{1cm} (21)

with $F_P(\theta)$ given in Eq. [7]. The Cauchy Schwarz inequality is saturated when the functions $v_k(\tau)$'s and $u_k(\tau)$'s are proportional, which occurs when

$$g_{m'n'}(\tau) = \frac{\beta}{\pi_{m'n'}(\tau, 0)} \left( \frac{\partial \pi_{m'n'}(\tau, \theta)}{\partial \theta} \right) \bigg|_{\theta=0},$$  \hspace{1cm} (22)

where the constant $\beta$ is the same for all $m'n'$.

The requirement that data in complete accordance with the expected distributions $\pi_{m'n'}(\tau, \theta)$ should lead to $S(\pi_{m'n'}(\tau, \delta \theta)) = \delta \theta$ establishes that, in fact, $\beta$ must be the inverse Fisher information $\beta = F_P^{-1}(\theta)$.

The shot noise limit, $\text{SNR} = 1$ in [21], defines the lowest distinguishable value of $\delta \theta = 1/\sqrt{F_P(\theta)}$, and collecting the results provides the linear estimator Eq. [18] in terms of the expected and the actually measured distribution of time intervals between the detector clicks,

$$S_P(n_{m'n'}(\tau)) = F_P^{-1}(\theta) \sum_{m'm'} \int \left( \frac{\partial n_{m'n'}(\tau)}{\partial \theta} \right) \bigg|_{\theta=0} \, d\tau.$$

The prior estimate is adjusted according to the discrepancy between the recorded waiting times and those expected from that prior. The Fisher information appears as a normalizing factor which reflects that larger adjustments may apply when the uncertainty is large. Still, we recall that this expression only applies asymptotically and that it is valid only if the first order Taylor expansions in the deviation from our prior guess are accurate enough, see also [15].

In the general case of non-Poissonian counting statistics $V_m \neq N_m$, and in the derivation above we must explicitly treat the $N_m$’s as independent stochastic variables that can themselves have a $\theta$-dependence. Factorizing the waiting time distributions $\pi_{m'n'}(\tau) = N_m(\theta) w_{m'n'}(\tau, \theta)$ allows us to employ separate gains for each $N_m$ in Eq. [18], and when the variance of the estimator Eq. [20] is corrected to include the proper variances $V_m$, the arguments given in this section carries over and the estimator acquires an extra term depending on the photon counts $N_m$,

$$S(n_{m'n'}(\tau)) = F_P^{-1}(\theta) \left[ \sum_{m'm'} \int \left( \frac{\partial n_{m'n'}(\tau, \theta)}{\partial \theta} \right) \bigg|_{\theta=0} \right] \times \left( \frac{n_{m'n'}(\tau)}{n_{m'n'}(\tau, \theta)} - 1 \right) \, d\tau$$

$$+ \sum_m \int \left( \frac{\partial N_m(\tau, \theta)}{\partial \theta} \right) \bigg|_{\theta=0} \left( \frac{N_m - N_m(\theta)}{V_m(\theta)} \right) \, d\tau.$$  \hspace{1cm} (24)

The Fisher information is given in Eq. [15], and Eq. [24] constitutes a linear estimator that exhausts the information in the multi-channel photon counting data record and, hence, achieves the Cramér-Rao Bound asymptotically.

V. PHOTON COUNTING FROM A LASER DRIVEN $\Lambda$-TYPE ATOM

As an example, we apply the formalism to a $\Lambda$-type system coupled to two laser fields, as shown in Fig. [1].

The couplings are described by Rabi frequencies $\Omega_0$ and $\Omega_1$ and laser-atom detunings $\delta_0$ and $\delta_1$ as indicated in the figure. We assume no direct coupling between $|0\rangle$ and $|1\rangle$, and that the decay into these two ground states is distinguishable, either by the polarization or by well-separated frequencies of the emitted photons.

In the rotating wave approximation, the Hamiltonian of the system can be written in matrix form as ($\hbar = 1$),

$$\hat{H}_0 = \begin{pmatrix} \delta_1 & 0 & \Omega_1^* \\ 0 & \Omega_0 & \delta_1 \\ \Omega_1 & \delta_1 & 0 \end{pmatrix}.$$  \hspace{1cm} (25)

The decay from $|2\rangle$ to $|0\rangle$ with rate $\Gamma_0$ and from $|2\rangle$ to $|1\rangle$ with rate $\Gamma_1$ (Fig. [1]) lead to a measurement record of photo detection events, and the intervals between the
associated quantum jumps can be sorted according to the corresponding four different types \((mn')\):

\[
\begin{align*}
(00): & \quad |2\rangle \rightarrow |0\rangle \text{ after } |2\rangle \rightarrow |0\rangle \\
(10): & \quad |2\rangle \rightarrow |0\rangle \text{ after } |2\rangle \rightarrow |1\rangle \\
(01): & \quad |2\rangle \rightarrow |1\rangle \text{ after } |2\rangle \rightarrow |0\rangle \\
(11): & \quad |2\rangle \rightarrow |1\rangle \text{ after } |2\rangle \rightarrow |1\rangle
\end{align*}
\]

Most physical systems are prone to dephasing, e.g., due to fluctuating magnetic fields, and we model this by introducing a decoherence term in the master equations \([8, 9, 12]\) corresponding to the operator \(\hat{C}_D = \sqrt{\gamma}(|0\rangle \langle 0| - |1\rangle \langle 1| + |2\rangle \langle 2|)\). The effect of this is to flip the sign of the \(|1\rangle\) amplitude relative to those of the two other states with a rate \(\gamma\).

In Fig. 2 we show two examples of the four delay functions \(w_{mm'}(\tau)\) for the \(\Lambda\)-system assuming perfect detection in both channels (physical parameters are given in the figure caption). For resonant coupling on both transitions (blue, solid lines) all four waiting time distributions resemble those of a two level system, (see \([7]\)). For finite detuning (red dashed lines) of the \(|1\rangle \leftrightarrow |2\rangle\) transition, the waiting time distribution functions after decay into \(|0\rangle\) largely maintain the same form, while, after decay into \(|1\rangle\) the distributions reflect the off-resonant \(|1\rangle \rightarrow |2\rangle\) excitation process.

In Fig. 3 we show in the upper panel the Fisher information per unit time for estimation of the laser-atom detuning \(\delta_1\) by photon counting. Results are shown for different values of the Rabi frequency, from weak \(\Omega_1 = 0.5\Gamma_0\) to strong \(\Omega_1 = 6\Gamma_0\), and the other parameters are \(\Omega_1 = 5\Gamma_0\), \(\delta_0 = 0\), \(\Gamma_1 = 0.5\Gamma_0\), and \(\gamma = 0.1\Gamma_0\). For \(\delta_0 = 0\), all statistical properties of the counting signal, and hence the Fisher information, are even functions of \(\delta_1\). Lower panel: The ratio \(\tau^2_m/\text{Var}(\tau)_m\) for the waiting times in the two channels as function of \(\delta_1\). We assume \(\Omega_1 = 3\Gamma_0\), while the remaining parameters are as in the upper panel.
Let us also investigate the parameter estimation sensitivity for a system with multiple decay channels of which only one is being observed. This situation occurs, e.g., in solid state emitters, which may relax both optically and by non-radiative coupling to the host material, and in the case of atoms which decay by emission of light in very different wave length regions. To describe this situation, we introduce hypothetical observers, Alice and Bob, holding only partial detection records. Alice has a perfect detector that monitors only the $|2\rangle \rightarrow |0\rangle$ channel. Her record of waiting times must then be matched to the distribution $w_{00}(\tau)$ found from Eq. (12), solved for the initial state $|0\rangle$ with $\theta_0 = 1$ and with $\eta_1 = 0$. Bob, on the other hand, monitors the $|2\rangle \rightarrow |1\rangle$ channel only, and his record of waiting times must be matched to the distribution $w_{11}(\tau)$ found from Eq. (12) solved for the initial state $|1\rangle$ with $\theta_0 = 0$ and $\eta_1 = 1$. The middle time-line in Fig. 4(a) illustrates a full detection record while the upper (lower) line shows the detection record of Alice (Bob).

In Fig. 4(b), we show the waiting time distributions for two values of the detuning $\delta_1$ (other physical parameters are given in the figure caption). The achievements of optimal frequency estimation strategies based on the individual records of Alice and Bob are given by the Fisher information Eq. (7), where the sum only has one term, $(mm') = (00)$ for Alice and $(mm') = (11)$ for Bob. Combining their records of waiting times, however, Alice and Bob may achieve a higher level of sensitivity. The Fisher information is then the sum of the individual Fisher informations according to Eq. (7). We show in Fig. 5 the Fisher information per time for estimation of $\delta_1$ by the separate records of Alice and Bob and by combining their registered distribution of waiting times. In Fig. 4(b), we observe that the delay function connected to the channel $|2\rangle \rightarrow |0\rangle$ is less sensitive to changes in detuning than the one pertaining to the $|2\rangle \rightarrow |1\rangle$ channel. This explains why Bob outperforms Alice at estimating the value of $\delta_1$.

The Fisher information for the full detection record (dash-dotted line in Fig. 5) is higher than that of Alice and Bob, even when they combine their waiting time records. This is because it makes use of all detection events and for example recognizes the first interval in Alice's record in Fig. 4(a) as two subsequent $(mm') = (01)$ and $(10)$ intervals rather than a single $(00)$ interval.

Consider, finally, an observer who has only access to the total, accumulated photon count. For a general multi-channel emitter the mean photo current in the asymptotic limit is $\hat{N}/T = \sum_m \text{Tr}(\hat{C}_m^\dagger \hat{C}_m \rho^{st}(\tau))$. For general counting statistics, we have $\Delta N = \sqrt{\hat{N}}$. This implies an uncertainty on $\theta$ given by $\Delta \theta = (\partial \hat{N}/\partial \theta)^{-1} \sqrt{\hat{N}}$, i.e., for detuning estimation in our $\Lambda$-atom,

$$\frac{(\Delta \delta_1)^{-2}}{T} = \frac{(\Gamma_0 + \Gamma_1)^2}{\Gamma_0 \text{Var}(\tau)_0 + \Gamma_1 \text{Var}(\tau)_1} \frac{(\partial \rho_{22}^{st}/\partial \delta_1)^2}{\rho_{22}^{st}},$$

where we have used Eq. (17) and $V = V_0 + V_1$. By Eq. (1)

\[
\begin{align*}
\delta_1 = & 1\Gamma_0, & \delta_1 = 2\Gamma_0
\end{align*}
\]

this can be directly compared to the Fisher information per time, and the result of Eq. (26) is included as the purple, long-dashed curve in Fig. 5. As expected, parameter estimates obtained from the full record and from the combined waiting time records of Alice and Bob achieve higher sensitivity on the whole detuning range.

\[ \text{VI CONCLUSION} \]

The full photon detection record of a quantum emitter contains more information about its dynamics than the mean signal. In this article, we have formulated a theory that quantifies this by calculating the Cramér-Rao sensitivity limit for multi-channel quantum light emitters: The information in the full photon detection record may be represented as waiting time distributions for which
Fig. 5. (Color online) The Fisher information per time for estimation of the laser-atom detuning $\delta_1$ in a $\Lambda$-type system by photon counting by Alice (green, solid line) and Bob (blue, dotted line), and by use of their combined records of waiting times (red, short-dashed line). The Fisher information from the complete detection record of both channels is shown as the dashed-dotted black curve, while the sensitivity obtained by only utilizing the total photon count (Eq. (26)) is shown by the purple, dashed curve. The results are calculated for the parameters $\Omega_0 = 5\Gamma_0$, $\Omega_1 = 2\Gamma_0$, $\delta_0 = 0$, $\Gamma_1 = 1$, and $\gamma = 0.1\Gamma_0$.

Eqs. (12,13) provide theoretical results, and which, by Eq. (15), supply the fundamental sensitivity limit Eq. (1). This optimal limit may be achieved via the linear estimator Eq. (24) or by a maximum likelihood estimate [5]. We exemplified the theory by the estimation of a detuning parameter in a driven $\Lambda$-type system with two distinct decay channels.

Our theory assumes an ergodic emitter, i.e., the system has a steady state which does not depend on the initial state of the system and which is not a dark state, such that the amount of accumulated data grows linearly with time. We also assumed that the decay of the system always feeds the same discrete set of final states, so that the data record can be analysed by a finite number of waiting time functions. Both the ergodicity assumption and the restriction to a finite number of final states are technical conditions for our method to apply, while our underlying Bayesian description is readily applied and several of the concepts introduced in this paper can be modified to account for the sensitivity limit in more general cases.

VII. ACKNOWLEDGMENTS

The authors acknowledge financial support from the Villum Foundation and helpful comments on the manuscript from C. K. Andersen, D. D. Bhaktavatsala Rao, P. Haikka, X. Qing, M. C. Tichy, and A. J. C. Wade.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330 (2004).
[2] G. Tóth and I. Apellaniz, J. Phys. A: Math. Theor. 47, 424006 (2014).
[3] H. Mabuchi, Quantum Semiclass. Opt. 8, 1103 (1996).
[4] J. Gambetta and H. M. Wiseman, Phys. Rev. A 64, 042105 (2001).
[5] S. Gammelmark and K. Mølmer, Phys. Rev. A 87, 032115 (2013).
[6] H. Cramér, Mathematical methods of statistics, Princeton mathematical series No. 9 (Princeton University Press, Princeton) (1954).
[7] A. H. Kiilerich and K. Mølmer, Phys. Rev. A 89, 052110 (2014).
[8] V. Delaubert, N. Treps, C. Fabre, H. A. Bachor, and P. Réfrégier, EPL 81, 44001 (2008).
[9] A. Negretti, C. Henkel, and K. Mølmer, Phys. Rev. A 78, 023630 (2008).
[10] H. Paul, Rev. Mod. Phys. 54, 1061 (1982).
[11] H. M. Wiseman and G. J. Milburn, Quantum Measurement and Control, (Cambridge University Press, Cambridge) (2010).
[12] H. Carmichael, An Open Systems Approach to Quantum Optics, Lectures Presented at the Université Libre de Bruxelles October 28 to November 4, 1991, (Springer Berlin Heidelberg) (1993).
[13] J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. 68, 580 (1992).
[14] S. Gammelmark and K. Mølmer, New J. Phys 13, 053035 (2011).
[15] R. D. Gill and M. I. Guta, On asymptotic quantum statistical inference in From Probability to Statistics and Back: High-Dimensional Models and Processes – A Festschrift in Honor of Jon A. Wellner, (Institute of Mathematical Statistics, Beachwood) (2013).