

Parameter estimation for quantum jump unraveling

Marco Radaelli^{1,2}, Joseph A. Smiga³, Gabriel T. Landi³, and Felix C. Binder^{1,2}

¹School of Physics, Trinity College Dublin, Dublin 2, Ireland

²Trinity Quantum Alliance, Unit 16, Trinity Technology and Enterprise Centre, Pearse Street, Dublin 2, Ireland

³Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA

We consider the estimation of parameters encoded in the measurement record of a continuously monitored quantum system in the jump unraveling, corresponding to a single-shot scenario, where information is continuously gathered. Here, it is generally difficult to assess the precision of the estimation procedure via the Fisher Information due to intricate temporal correlations and memory effects. In this paper we provide a full set of solutions to this problem. First, for multi-channel renewal processes we relate the Fisher Information to an underlying Markov chain and derive a easily computable expression for it. For non-renewal processes, we introduce a new algorithm that combines two methods: the monitoring operator method for metrology and the Gillespie algorithm which allows for efficient sampling of a stochastic form of the Fisher Information along individual quantum trajectories. We show that this stochastic Fisher Information satisfies useful properties related to estimation on a single run. Finally, we consider the case where some information is lost in data compression/post-selection and provide tools for computing the Fisher Information in this case. All scenarios are illustrated with instructive examples from quantum optics and condensed matter.

In many situations of physical interest, an experimenter may want to obtain the value of a physical parameter that is not accessible via direct measurement. To do so, they must estimate the parameter based on a finite sample of outcomes. Estimation theory [1–4] provides tech-

niques for retrieving the value of the parameter based on such sampling, as well as bounds on the maximum precision that can be attained.

A common approach to parameter estimation in quantum systems considers a large number of independent and identically distributed (i.i.d.) copies of the system at hand [5–7]. A considerable research effort has been devoted to determining if and how genuinely quantum properties, such as entanglement and squeezing, can yield an advantage over fully classical estimation [8–10].

In contrast to the i.i.d. case, another scenario of experimental importance is that of a single system that is sequentially measured [11–14]. This can be implemented, for instance, using periodic measurements [15–17], collision models [18–22], or quantum continuous measurements [11, 12, 14, 23–34]. Sometimes the system can be reset after each measurement, leading to a scenario that is akin to the i.i.d. case. However, this is not always the case. As a consequence, each data point is no longer independent, leading to a (possibly intricate) memory structure in the detection record.

In this work we provide a full characterization of quantum metrology for quantum jump unravelings [35, 36]. In this scenario a system evolves according to abrupt jumps that occur at random times and in random channels, following statistics determined by the system dynamics. In between jumps the evolution is smooth and governed by a modified non-Hermitian Hamiltonian. Such quantum jump unravelings are an apt description of the common experimental case in which the measurement apparatus only records the type and time of emissions from the system, e.g. in atomic physics [37], superconducting systems [38, 39] or cavity QED [40, 41].

Note that the problem addressed here differs from metrology on the unconditional state of an open quantum system. In the latter case, the in-

Marco Radaelli: radaellm@tcd.ie

Joseph A. Smiga: joseph.smiga@rochester.edu

Gabriel T. Landi: gabriel.landi@rochester.edu

Felix C. Binder: felix.binder@tcd.ie

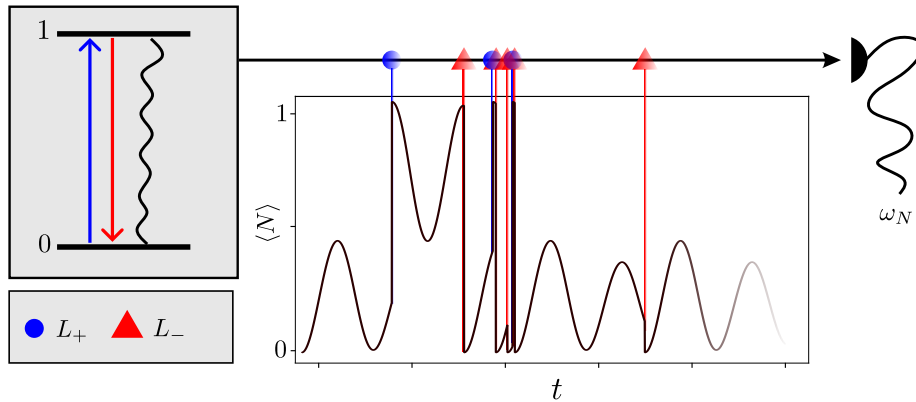


Figure 1: A quantum process emits a signal every time a jump occurs. There are two different jump channels (labeled as L_+ and L_-), corresponding to distinguishable signals. The experimenter has access only to the measurement record, constituted by jump time and jump channel for each emission. In the illustration, the quantum system is a qubit coupled to a thermal bath, whose behavior is discussed in detail in Sec. 5.1. The plot illustrates the expectation value of the number observable for the system as a function of time.

put of the estimation task is the full state of the open quantum system. Here, in contrast, the input is a measurement record as would be read on a detector during an experimental run. In this scenario, the observer has no control over the interaction between the system and the detector.

Restricted treatments of this problem were given in Refs. [26, 27, 33] for the specific case of single-channel photodetection. Here we extend this to arbitrary jump processes with multiple channels. Fig. 1 illustrates this scenario for the case of a qubit with two monitoring channels. The measurement record of such an experiment is generally of the form

$$\omega_{1:N} = \{(\tau_1, k_1), (\tau_2, k_2), \dots, (\tau_N, k_N)\}, \quad (1)$$

where k_i labels the respective channels and τ_i labels the time between jumps k_i and k_{i-1} . We assume that the system dynamics depend on some parameter θ , which may enter via the Hamiltonian, the jump operators, or both. The challenge is then to estimate θ to high precision. While elegant quantum bounds that maximize over all unravelings have been introduced in Refs. [11, 12, 22], these are generally not saturable in the jump unraveling. In contrast, we here focus on estimators $\hat{\theta}(\omega_{1:N})$ that only employ the (classical) measurement record in order to derive saturable precision bounds.

We remark that since the experimenter has no control over the measurement to be performed, we do not optimize over possible measurement strategies. Consequently, our problem is one of

(correlated) classical metrology, in contrast with the standard framework of quantum metrology, where the experimenter can choose the measurement to be applied [5, 6].

An interesting feature of the quantum jump scenario are the intricate memory patterns that different jumps retain of the past. The qubit in Fig. 1 is an example of a renewal process, where after each jump we know with certainty which state the system jumped to. Multi-channel renewal systems can be mapped to semi-Markov process [42, 43] so that the memory is short-lived. Many quantum systems, however, are not renewal. A simple example is an optical cavity subject to direct photodetection: each jump perturbs the system, but does not fully reset the state. Quite the contrary, the memory in these systems can be extremely long lived. This makes it difficult to estimate the Fisher Information (FI). Our goal in this paper is to provide a full set of tools for doing so. For multi-channel renewal we show that a simple-to-use formula can be derived (Sec. 2). Conversely, for non-renewal processes we introduce a new algorithm based on the monitoring operator [26] and the quantum Gillespie method [44]. This approach naturally introduces the concept of a stochastic FI which can be sampled along individual trajectories by our algorithm (Sec. 3). We show that the stochastic FI naturally fits into the description of single-shot estimation. A summary of these results is given in Table 1.

Process type	Fisher information
Single-channel renewal	$F_N(\theta) = -N \int_0^\infty d\tau W(\tau) \left(\frac{\partial^2}{\partial \theta^2} \ln W(\tau) \right)$ [27]
Multi-channel renewal	$F_N(\theta) = -N \sum_{m,n \in \mathbb{M}} p_m \left(\frac{\partial^2}{\partial \theta^2} \ln W(\tau, n m) \right)$ [Eq. (17)]
Multi-channel non-renewal	Numerical Gillespie algorithm [Sec. 3.1]

Table 1: Our work covers the calculation of the Fisher information contained in the WTD for a wide set of quantum processes, both renewal (where we extend a previously known result) and non-renewal, where we present a numerical strategy allowing for an efficient calculation.

Furthermore, as a relevant specification of our work, we discuss the loss of Fisher information under data compression/post-selection (Sec. 4). For example, what happens if one loses the information about the distinction among channels k_i in Eq. (1)? Or what happens if we build estimators based only on summary statistics, such as the average waiting times $E(\tau_i)$? Our results are illustrated with several examples in Sec. 5.

1 Setup

1.1 Estimation theory

Consider a string of random variables $X_{1:N}$ (here and in the following, the notation for strings is left- and right-inclusive: $X_{1:N} = X_1 X_2 \dots X_{N-1} X_N$), distributed according to a joint probability distribution $\Pr_\theta(X_{1:N})$ which depends on some unknown parameter $\theta \in \Theta$. We assume θ to be a scalar; the generalization to multi-parameters is straightforward. The goal of estimation theory is to construct an estimator $\hat{\theta}(X_{1:N})$ that provides an estimate of θ . The variance of an unbiased estimator ($\mathbb{E}(\hat{\theta}) = \theta$) is bounded from below by the Cramér-Rao bound [45, 46]

$$\text{Var} \left[\hat{\theta}(X_{1:N}) \right] \geq \frac{1}{F_{1:N}(\theta)}, \quad (2)$$

where the Fisher Information (FI)

$$F_{1:N}(\theta) = \sum_{x_{1:N}} \Pr_\theta(x_{1:N}) [\partial_\theta \log \Pr_\theta(x_{1:N})]^2, \quad (3)$$

represents the total information that $\Pr_\theta(X_{1:N})$ contains about θ . In the simplest scenario the X_i are independent and identically distributed (i.i.d.), causing the FI to simplify as $F_{1:N}(\theta) = N F_1(\theta)$ [47]. In this case, the quantity $F_1(\theta)$ represents the information rate — i.e., the FI acquired per symbol observed. In the presence of

correlations, in contrast, the calculation of Eq. (3) becomes significantly involved, owing to the high dimensionality of the sum.

The Cramér-Rao bound of Eq. (2) is always asymptotically saturable, via the maximum likelihood estimation technique detailed in Sect. 3.2. In the case of finite data, however, the bound may not be saturable. If it is, the choice of the appropriate estimator may be contingent on the specific form of \Pr_θ . For this reason, in the present work we focus on the computation of the Fisher information, and on the saturation of the bound in the asymptotic limit.

1.2 The waiting time distribution

In this paper we wish to investigate the above ideas in the context of continuously measured systems in the quantum jump unraveling. We consider a general Markovian evolution of a quantum state $\rho(t)$, as described by a quantum master equation of the form

$$\frac{d\rho}{dt} = \mathcal{L}\rho = -i[H, \rho] + \sum_{k=1}^r \mathcal{D}[L_k]\rho, \quad (4)$$

where \mathcal{L} is the Liouvillian, H is the system Hamiltonian and \mathcal{D} the dissipator, defined by the relation $\mathcal{D}[L]\bullet = L\bullet L^\dagger - \frac{1}{2}\{L^\dagger L, \bullet\}$. $\{L_k\}$ represents a set of jump operators. Equation (4) can be obtained from a closed-system microscopic description including both the system and the environment, in the limit of a weak interaction between them. We will refer to each k as a *channel*. In practical problems, each can represent a different physical entity. For example, photons might be emitted in different frequencies or polarizations. Conversely, in quantum dots one jump operator might represent the injection of electrons from a lead, while another represents extraction.

Equation (4) is a differential equation for the evolution of the state ρ . Equivalently, it can

be seen as the average over stochastic trajectories, where a different effect is applied at each timestep; this approach is known as *unraveling* of the master equation. The role of unravelings is two-fold: on one hand, they can be used to numerically solve the GKSL equation by sampling [48, 49]. Secondly, unravelings describe what an experimenter observes in the lab under continuous monitoring. In this work, we employ unravelings in the latter sense, as single-run measurement records.

Different unravelings of Eq. (4) have been proposed in the literature [36], with a prominent role being played by homodyne unraveling and quantum jump unraveling [48], corresponding to different experimental setups.

In this work, we focus on quantum jump unraveling, in which one considers quantum trajectories described by abrupt jumps occurring at random times and in random channels (as in Eq. (1)) [36]. In between jumps the system evolves smoothly according to a (non-unitary) no-jump dynamics. The jump channels k can thus be associated with “detectors” which record whenever a jump occurs in channel k . In many cases not all channels can be monitored. We therefore define the set of monitored channels \mathbb{M} . We may also assume that each channel has a finite efficiency η_k of recording a detection, such that $\eta_k = 1$ represents perfect efficiency. We define the superoperators

$$\mathcal{J}_k \rho = \eta_k L_k \rho L_k^\dagger, \quad (5)$$

as well as the no-jump superoperator

$$\mathcal{L}_0 = \mathcal{L} - \sum_{k \in \mathbb{M}} \mathcal{J}_k, \quad (6)$$

where the sum is only over the monitored set \mathbb{M}^I .

^I Note here that from a mathematical perspective, it is actually not necessary to use both efficiencies η_k and a subset \mathbb{M} of monitored channels. For example, we can treat all channels as monitored and then set $\eta_k = 0$ for those channels which are actually not. Likewise, if we prefer not to use efficiencies, we can decompose a Lindblad dissipator as $\mathcal{D}[L_k] = \mathcal{D}[\sqrt{\eta_k} L_k] + \mathcal{D}[\sqrt{1 - \eta_k} L_k]$ splitting each channel in two, with new jump operators $\sqrt{\eta_k} L_k$ and $\sqrt{1 - \eta_k} L_k$. Then we can include $\sqrt{\eta_k} L_k$ in the subset \mathbb{M} and leave $\sqrt{1 - \eta_k} L_k$ out. Keeping both concepts can be convenient, however, especially when making the connection with experiments.

The stochastic dynamics in the quantum jump unraveling can now be written as [25, 50]

$$d\rho = dt \mathcal{L} \rho + \sum_{k \in \mathbb{M}} \left(dN_k(t) - dt \text{Tr}(\mathcal{J}_k \rho) \right) \cdot \left(\frac{\mathcal{J}_k \rho}{\text{Tr}(\mathcal{J}_k \rho)} - \rho \right), \quad (7)$$

where dN_k is a Poisson increment — that is, a random variable that takes the value 1 if a jump occurs in channel k , and 0 otherwise. The probability of this happening is $\text{Pr}(dN_k(t) = 1) = dt \text{Tr}[\mathcal{J}_k \rho]$. Since these probabilities are infinitesimal, at each time step dt the probability of a jump taking place is vanishingly small. In other words, as is intuitive, most of the time $dN_k = 0$ and the system will evolve with no jumps.

Because of this structure, it is simpler (and entirely equivalent) to consider a stochastic process where we specify the time between jumps, as well as the channel in which the jump occurred. This is precisely the trajectory $\omega_{1:N}$ in Eq. (1) and is, in fact, the stochastic process that an experimentalist would observe. Thus, instead of dealing with the stochastic master equation (7), we can simply write down the joint probability density of observing the trajectory $\omega_{1:N}$. For concreteness, we assume that the system is initially in the state, ρ_0 . The probability density of observing the measurement record Eq. (1) is then

$$\text{Pr}(\omega_{1:N}) = \text{Tr} \{ \mathcal{J}_{k_N} e^{\mathcal{L}_0 \tau_N} \dots \mathcal{J}_{k_1} e^{\mathcal{L}_0 \tau_1} \rho_0 \}. \quad (8)$$

The question at hand is then the following: *assuming that H and L_k can depend on an unknown parameter θ , what is the precision with which we can estimate θ based only on the stochastic process in Eq. (1)?* The answer is given by the FI, Eq. (3), with the stochastic string $X_{1:N}$ replaced by $\omega_{1:N}$. In general, Eq. (8) can exhibit complex correlations: the effects of a given jump can persist, even after other jumps are measured. As a consequence, it cannot in general be decomposed in terms of simpler probabilities. This makes a direct evaluation of Eq. (3) extremely involved. Addressing this challenge is the main contribution of this paper. Before proceeding, however, some comments about Eq. (8) are in order:

- Equation (8) refers only to the monitored jumps $k \in \mathbb{M}$: in between them, any number of non-monitored jumps can occur. This is automatically taken into account via the definition of \mathcal{L}_0 .

- The probability given by Eq. (8) refers to a specific number of jumps N . This can be viewed as an ensemble where N is fixed, but the final time $t_f = \sum_{i=1}^N \tau_i$ is allowed to fluctuate. We call this the N -ensemble. One could also consider an ensemble where N fluctuates and t_f is fixed. We will call this the t_f -ensemble. In the limit of large N (or large t_f) results for the two ensembles is expected to coincide [51].
- Equation (8) is not necessarily normalized. Sub-normalization occurs in cases where there is a non-zero probability that a jump never happens. Here, we exclude this possibility by assumption. That is, we work with processes where jumps must always eventually occur (called *persistent* in the stochastic process literature [52]). Normalization is then given by $\int d\omega_{1:N} \Pr(\omega_{1:N}) = 1$ with

$$\int d\omega_{1:N} = \sum_{k_1, \dots, k_N} \int_0^\infty d\tau_1 \cdots \int_0^\infty d\tau_N. \quad (9)$$

2 Multi-channel renewal processes

A special subclass of processes exhibiting a much-simplified memory structure are the so-called renewal processes, in which the state is reset after each jump. This occurs whenever the jump operators satisfy the *renewal condition*,

$$\mathcal{J}_k \rho = \text{Tr}[\mathcal{J}_k \rho] \sigma_k, \quad (10)$$

where σ_k is a physical state that depends on the jump channel but not on ρ . Equation (10) causes Eq. (8) to factorize as

$$\Pr(\omega_{1:N}) = W(\tau_N, k_N | k_{N-1}) \cdots W(\tau_1, k_1 | k_0), \quad (11)$$

where

$$W(\tau, k | q) = \text{Tr} \{ \mathcal{J}_k e^{\mathcal{L}_0 \tau} \sigma_q \}, \quad (12)$$

is the waiting time distribution for the next jump to occur after time interval τ and in channel k , given that the previous jump occurred in channel q . Notice that this is normalized as $\sum_k \int_0^\infty W(\tau, k | q) d\tau = 1$ for all q . In terms of composite random variables $x_i = (\tau_i, k_i)$, it is clear that under the renewal condition $\omega_{1:N}$ itself is a Markov chain (Markov order 1 stochastic process [53]).

For renewal processes in which each measurement corresponds to a single jump operator L_k acting on the state, $\mathcal{J}_k \rho = \gamma_k L_k \rho L_k^\dagger$, satisfying the renewal condition [Eq. (10)], one may show that the jump operators themselves must always have the form $L_k = |\mu_k\rangle \langle \nu_k|$ (the constant rate γ_k can always be chosen to satisfy this condition), where $|\mu_k\rangle$ and $|\nu_k\rangle$ are generic normalized quantum states which need not be orthogonal (c.f. Appendix A for a proof). It then follows that the post-jump states in this renewal process are pure, $\sigma_k = |\mu_k\rangle \langle \mu_k|$, while the jump probabilities are $\text{Tr}(\mathcal{J}_k \rho) = \gamma_k \langle \nu_k | \rho | \nu_k \rangle$. The waiting time distribution in Eq. (12) can therefore be simplified to

$$W(\tau, k | q) = \gamma_k \langle \nu_k | e^{\mathcal{L}_0 \tau} (|\mu_q\rangle \langle \mu_q|) |\nu_k\rangle. \quad (13)$$

If all channels are monitored, the no-jump super-operator can be written as $\mathcal{L}_0 \rho = -i(H_e \rho - \rho H_e^\dagger)$, with a non-Hermitian Hamiltonian

$$H_e = H - \frac{i}{2} \sum_k L_k^\dagger L_k = H - \frac{i}{2} \sum_k \gamma_k |\nu_k\rangle \langle \nu_k|. \quad (14)$$

The WTD then simplifies further to

$$W(\tau, k | q) = |\Psi(\tau, k | q)|^2, \quad (15)$$

where $\Psi(\tau, k | q)$ is a “waiting time amplitude,” given by

$$\Psi(\tau, k | q) = \sqrt{\gamma_k} \langle \nu_k | e^{-iH_e \tau} |\mu_q\rangle. \quad (16)$$

2.1 Fisher information of multi-channel renewal processes

We now analyze how Eq. (3) simplifies under the factorization in Eq. (11). As shown in Appendix B, except for a small boundary term that is negligible when N is large, the FI for renewal processes can be written as

$$F_{1:N}(\theta) = N \sum_{k, q \in \mathbb{M}} p_k \int_0^\infty d\tau \frac{[\partial_\theta W(\tau, k | q)]^2}{W(\tau, k | q)}. \quad (17)$$

That is, the Fisher information of the entire measurement record reduces to the Fisher information of the WTD. Here

$$p_k := \frac{\text{Tr}[\mathcal{J}_k \rho_{\text{ss}}]}{\sum_j \text{Tr}[\mathcal{J}_j \rho_{\text{ss}}]}, \quad (18)$$

is the steady-state probability with which jumps occur in channel k . Equation (17) is our first

main result. Notice that the scaling of the Fisher information is still linear in the number of jumps; this is expected for processes of finite Markov order [54] and beyond [55], and is in line with Ref. [12], which studied the quantum Fisher information maximized over all possible unravelings of the master equation. We would like to alert the reader to a possible confusion with similar expressions sometimes found in the discussion of the Fisher information for classical master equations (e.g. [56]): W here is a function taking values over real numbers and a pair of jump labels, very different from a rate matrix.

The Cramér-Rao bound involving the Fisher information for a multi-channel renewal process can be asymptotically saturated by maximum likelihood estimation [2, 57], where the log-likelihood function takes the form

$$\ell(\tilde{\theta}|\omega_{1:N}) = \sum_{j=2}^N \log W_{\tilde{\theta}}(\tau_j, k_j|k_{j-1}). \quad (19)$$

Here, $W_{\tilde{\theta}}$ represents the waiting time distribution assuming that the value of the parameter is $\tilde{\theta}$, and border terms, that do not play any role in the asymptotic limit, are neglected. Other estimators generally perform worse than the MLE, at least in the limit of a large number of jumps on the trajectory. An example of an estimation task performed on the qubit thermometry system (as defined in Sec. 5.1) is discussed in Appendix C.

In the case of a single jump channel Eq. (11) will factor into a product of probabilities. The waiting times τ_i therefore become independent and identically distributed (the process becomes of Markov order 0). This is the scenario of renewal theory [58] and Eq. (17) simplifies to

$$F_{1:N}(\theta) = N \int_0^{\infty} d\tau \frac{[\partial_{\theta} W(\tau)]^2}{W(\tau)}, \quad (20)$$

where $W(\tau) = \text{Tr} \{ \mathcal{J} e^{\mathcal{L}_0 \tau} \sigma \}$. This agrees with a result derived in Ref. [12].

The Fisher information in Eq. (17) encompasses the information contained in both the jump channels and the jump times. It is possible to separate the two individual contributions using the basic decomposition law $F(X, Y) = F(Y) + F(X|Y)$, which holds for the Fisher information of any joint probability distribution [47, 54, 59]. As a result, one finds

$$F_{1:N}(\theta) = F_{1:N}^{\text{ch}}(\theta) + F_{1:N}^{\text{times|ch}}(\theta), \quad (21)$$

where

$$F_{1:N}^{\text{ch}}(\theta) = N \sum_{k,q \in \mathbb{M}} p(k|q) p_q (\partial_{\theta} \ln p(k|q))^2, \quad (22)$$

is the Fisher information of the sequence of jump channels only, with

$$p(k|q) = \int_0^{\infty} d\tau W(\tau, k|q) = -\text{Tr} \{ \mathcal{J}_k \mathcal{L}_0^{-1} \sigma_q \}, \quad (23)$$

being the probability distribution that the system transitions from $q \rightarrow k$, irrespective of when it happens. Here, we assume that the inverse of the no-jump Lindbladian exists; this is a sufficient condition for the absence of dark subspaces.

We mention in passing that the quantity p_k in Eq. (18) is also the solution of

$$\sum_q p(k|q) p_q = p_k, \quad (24)$$

as one may verify by explicit substitution. The second term in Eq. (21), on the other hand, is the conditional Fisher information of the time tags, given one knows the sequence of channels:

$$F_{1:N}^{\text{times|ch}}(\theta) = N \sum_{k,q \in \mathbb{M}} p(k|q) p_q \int_0^{\infty} d\tau \frac{[\partial_{\theta} W(\tau|k, q)]^2}{W(\tau|k, q)}, \quad (25)$$

where

$$W(\tau|k, q) = \frac{W(\tau, k|q)}{p(k|q)}, \quad (26)$$

is the probability that the jump takes a time τ , given that the sequence of jumps was $q \rightarrow k$. The quantity $F_{1:N}^{\text{times|ch}}$ is therefore the FI of the time tags, given that one knows the channels. The FI of the time tags when one is ignorant about the channels will be discussed in Sec. 4.

The two terms in Eq. (21) are individually non-negative, and have clear physical interpretations. Therefore, they can be used to assess how much information is contained in each aspect of the stochastic process (channels and times). The relative contribution of F^{ch} to the total Fisher information is illustrated in Fig. 3(b-c).

In passing, we mention that if we insert Eq. (26) into the general probability string Eq. (11), we can write it as

$$\begin{aligned} \Pr(\omega_{1:N}) &= [W(\tau_N|k_N, k_{N-1}) \cdots W(\tau_1|k_1, k_0)] \\ &\quad \times [p(k_N|k_{N-1}) \cdots p(k_1|k_0)]. \end{aligned} \quad (27)$$

Thus, we see that for renewal processes the path trajectory can be split into a probability for the sequence of channels $k_0 \rightarrow k_1 \rightarrow \dots$ times the conditional probability of the jump times τ_i , given this sequence of channels.

The integral appearing in Eq. (17) can be difficult to compute analytically, depending on the shape of the WTD. The same is true for Eq. (25). On the other hand, Eqs. (23) and (21) allow for an easy computation of the Fisher information when time tags are ignored. In Appendix D, we discuss the special case of a WTD that is a single decaying exponential, for which the integration can be analytically carried out; this case most frequently appears in incoherent systems described by classical master equations. In Appendix E, we introduce a bound on the Fisher information in terms of the mean and variance of the WTD, as

$$F_{1:N}(\theta) \geq F_{1:N}^{\text{ch}}(\theta) + N \sum_{k,q} \frac{[\partial_\theta \mu_{kq}]^2}{\sigma_{kq}^2} p(k|q) p_q, \quad (28)$$

where $\mu_{kq} = \mathbb{E}[\tau|k, q]$ and $\sigma_{kq}^2 = \text{Var}[\tau|k, q]$ are the conditional mean and variance of τ , calculated from Eq. (26). This bound can be useful in situations in which the integral of Eq. 17 is hard to compute, but the variance and the mean of the WTD is computable.

One may also consider the inverse decomposition of the Fisher information between channels and jumps, of the form $F_{1:N}(\theta) = F_{1:N}^{\text{times}}(\theta) + F_{1:N}^{\text{ch|times}}(\theta)$. While this decomposition is formally valid, the times-only process is, in general, non-renewal even though the original process is renewal. This prevents us from expressing $F_{1:N}^{\text{times}}$ analytically. The computation can be performed numerically, as detailed in Sect. 4.

3 Non-renewal processes

The calculation of the FI in Eq. (3) for a general measurement record $\omega_{1:N}$ is difficult because the probability does not factorize in any simple way as it did for renewal processes in Eq. (11). As a consequence, one must compute a large number of high-dimensional sums/integrals, rendering the problem generally intractable. In this section we start by reviewing the monitoring operator formalism [26], which allows one to compute the Fisher information by sampling over quantum trajectories. We then adapt this formal-

ism to our specific ensemble distribution given by Eq. (8), and present a new algorithm for efficient computation of the Fisher information over non-renewal quantum trajectories, based on the quantum Gillespie algorithm introduced in Ref. [44].

3.1 Monitoring operator for generic stochastic quantum mechanics

Consider a generic set of trace-non-increasing superoperators $\{\mathcal{M}_x\}$ that describe the evolution of the system after outcome x . Namely, if the system is initially in the normalized state ρ , then

$$\tilde{\rho}' = \mathcal{M}_x \rho \quad (29)$$

is the unnormalized state such that $\text{Tr} \tilde{\rho}'$ is the probability of measuring outcome x and $\rho' = \frac{\tilde{\rho}'}{\text{Tr} \tilde{\rho}'}$ is the corresponding normalized state. Stringing many measurements together, starting from an initial state ρ_0 , one defines the unnormalized density operator

$$\tilde{\rho}_N = \mathcal{M}_{x_N} \cdots \mathcal{M}_{x_1} \rho_0. \quad (30)$$

The corresponding probability is

$$\text{Pr}(x_1, \dots, x_N) = \text{Tr}\{\tilde{\rho}_N\}. \quad (31)$$

For the case of quantum jumps, $\mathcal{M}_x = \mathcal{J}_k e^{\mathcal{L}_0 \tau}$, so the outcomes are given by the pair $x = (k, \tau)$.

The *monitoring operator* for an unknown parameter θ is defined as as [26]

$$\xi_N = \frac{\partial_\theta \tilde{\rho}_N}{\text{Tr}\{\tilde{\rho}_N\}}. \quad (32)$$

One may verify that

$$F_{1:N}(\theta) = \mathbb{E}[\text{Tr}(\xi_N)^2], \quad (33)$$

is precisely the general FI in Eq. (3) (here the ensemble average is over Eq. (31)). Hence, stochastically sampling the monitoring operator provides us a method for computing the Fisher information in the non-renewal case. In addition, as we also discuss in Sec. 3.2, the monitoring operator can be used to perform maximum likelihood estimation tasks on quantum jump trajectories, solving many numerical instabilities issues with respect to a more traditional approach.

A major benefit of the monitoring operator formalism is that there is a simple, inductive process

to calculate it. First, for an initial state ρ_0 , the monitoring operator is initialized as

$$\xi_0 = \partial_\theta \rho_0. \quad (34)$$

Starting from the state, ρ_j , the next observation x_{j+1} is sampled from

$$\Pr(x_{j+1}|\rho_j) = \text{Tr}\{\mathcal{M}_{x_{j+1}}\rho_j\}, \quad (35)$$

and the state evolves to $\rho_{j+1} = \frac{\mathcal{M}_{x_{j+1}}\rho_j}{\text{Tr}\{\mathcal{M}_{x_{j+1}}\rho_j\}}$. Meanwhile, the monitoring operator evolves to

$$\xi_{j+1} = \frac{(\partial_\theta \mathcal{M}_{x_{j+1}})\rho_j + \mathcal{M}_{x_{j+1}}\xi_j}{\text{Tr}\{\mathcal{M}_{x_{j+1}}\rho_j\}}. \quad (36)$$

Note that this calculation only depends on the previous monitoring operator ξ_j , the superoperators $\{\mathcal{M}_x\}$, and the normalized density operator ρ_j . This process does not depend on keeping track of the unnormalized state, which leads to better numeric stability [31].

The inductive algorithm described above enables a computationally simple way to determine the evolution of the state ρ_j and monitoring operator ξ_j from their respective initial conditions. For calculating the Fisher information, this process is repeated many times to calculate $F_{1:N}(\theta) = \mathbb{E}[\text{Tr}(\xi_N)^2]$. In the next section, we show how to use the monitoring operator formalism to perform MLE estimation on general quantum stochastic processes.

We remark that the monitoring operator formalism is fully general with respect to the specific unraveling of a given master equation. For instance, it can be applied to homodyne unraveling [26, 31]. Consequently, most of the results presented below can be adapted to homodyne detection or other protocols, with the exception of the use of the Fisher-Gillespie algorithm, which crucially relies on the jump structure of the evolution.

3.2 MLE estimation

Consider a measurement record $\omega_{1:N}$, i.e. a string of outcomes (x_1, \dots, x_N) . This is the outcome of a single run of the experiment. The *likelihood* is the probability of obtaining $\omega_{1:N}$, as a function θ , the parameter to be estimated:

$$\ell(\tilde{\theta}) = \Pr[\omega_{1:N}|\tilde{\theta}]. \quad (37)$$

In MLE [4, 57] the estimator for θ on a single run is taken as the maximum of the $\ell(\tilde{\theta})$ function over all values of $\tilde{\theta}$:

$$\hat{\theta} = \underset{\tilde{\theta}}{\text{argmax}} \ell(\tilde{\theta}). \quad (38)$$

In other words, MLE assumes that the currently observed measurement record must come from the value of the parameter that makes it the most likely compared to other values of θ . It is possible to prove that, under suitable regularity conditions, MLE is asymptotically unbiased and saturates the Cramér-Rao bound.

In order to be able to perform the maximization, one has to be able to compute the likelihood $\ell(\tilde{\theta})$ for many different values of $\tilde{\theta}$. When considering long evolutions (that is, most situations of practical relevance in quantum metrology), we are faced with the problem that $\ell(\tilde{\theta})$ decays exponentially with the length of the measurement record. This renders the estimation task numerically unfeasible. In the context of finitely correlated processes (i.e., processes with a finite Markov order), the issue can be successfully tackled by maximizing the logarithm of the likelihood function, and exploiting the factorization properties of logarithms [54]. However, quantum processes in general exhibit an infinite Markov order. Here, we show a method, based on the monitoring operator formalism, that circumvents this issue. An alternative but equivalent method has been proposed in Ref. [26].

Given the measurement record $\omega_{1:N}$, and a candidate value for the parameter $\tilde{\theta}$, we can define a corresponding monitoring operator $\xi_{\omega_{1:N}}^{\tilde{\theta}}$, that is, the monitoring operator evolved along the measurement record $\omega_{1:N}$, assuming that the value of the parameter is $\tilde{\theta}$. In an analogous way, we can evolve the sub-normalized quantum state $\tilde{\rho}_{\omega_{1:N}}^{\tilde{\theta}}$ on $\omega_{1:N}$. The likelihood can then be written as

$$\ell(\tilde{\theta}) = \Pr[\omega_{1:N}|\tilde{\theta}] = \text{Tr}[\tilde{\rho}_{\omega_{1:N}}^{\tilde{\theta}}]. \quad (39)$$

Let us now assume that, in the estimation region, the loglikelihood function has a single maximum. Note that for well-behaved loglikelihood functions this is not a strong constraint since one can always restrict the estimation region. If the maximum lies in the interior of the region, then it can be detected by imposing $\partial_\theta \ell(\theta) = 0$. If the maximum lies on the border, it cannot be detected in this way. For this reason, it makes

sense to check the loglikelihood function on the border separately, for instance with the method of Ref. [26]. We now focus on the case of a maximum inside the region.

The maximization condition yields

$$\partial_{\tilde{\theta}} \text{Tr} [\tilde{\rho}_{\omega_{1:N}}^{\tilde{\theta}}] = 0 \iff \text{Tr} [\xi_{\omega_{1:N}}^{\tilde{\theta}}] = 0. \quad (40)$$

In this way, we mapped the problem of finding the $\tilde{\theta}$ that maximizes the likelihood to the one of finding the $\tilde{\theta}$ that makes the monitoring operator traceless. Remarkably, the latter is a numerically stable problem, because the evolution of the monitoring operator can be integrally computed from the normalized conditional state. For practical purposes, the problem can be further mapped onto finding the $\tilde{\theta}$ that minimizes the squared trace of the monitoring operator. To summarize:

$$\hat{\theta} = \underset{\tilde{\theta}}{\text{argmax}} \ell(\tilde{\theta}) = \underset{\tilde{\theta}}{\text{argmin}} \left(\text{Tr} [\xi_{\omega_{1:N}}^{\tilde{\theta}}] \right)^2. \quad (41)$$

We remark that the monitoring operator formalism for MLE estimation is fully general, regardless of the specific form of the \mathcal{M}_x superoperators. An application to the case of quantum jump trajectories is presented in Fig. 2 for the coupled qubit model, whose details are discussed in Sec. 5.3. In the asymptotic limit of a large final time t_f , MLE gives normally distributed estimates for the unknown parameter, and the variance of such distribution saturates the Cramér-Rao bound. Code for MLE estimation on quantum jump trajectories, using the monitoring operator formalism, is made available in Julia [60].

3.3 The Fisher-Gillespie algorithm

To compute the Fisher information according to Eq. (33), one has to sample from the distribution of all the possible sequences of superoperators \mathcal{M}_x , weighted by the appropriate probability of the trajectory. In the case of quantum jumps, $\mathcal{M}_x = \mathcal{J}_k e^{\mathcal{L}_0 \tau}$, so the time evolutions add an additional level of complexity as the superoperators are parameterized by a continuous variable, time, that can be arbitrarily subdivided.

To make this computation tractable, we introduce a new algorithm, which jointly exploits the monitoring operator formalism for the Fisher information, and the Gillespie approach [61, 62] to quantum jump trajectories, introduced in Ref. [44]. Here, we give a quick overview of the working principles of the algorithm.

To sample from the distribution of all possible jump trajectories given by Eq. (8), we assume that we start in a state ρ_{j-1} , immediately after a jump. We then sample a random waiting time T_j from the distribution

$$W(\tau|\rho_{j-1}) = \text{Tr} \left[\mathcal{J} e^{\mathcal{L}_0 \tau} \rho_{j-1} \right], \quad (42)$$

where $\mathcal{J} = \sum_{k \in \mathbb{M}} \mathcal{J}_k$. This gives the WTD for the next jump to occur, irrespective of the monitored channel it takes place in. We then update the state ρ_{j-1} to

$$\bar{\rho}_j = \frac{e^{\mathcal{L}_0 T_j} \rho_{j-1}}{\text{Tr}(e^{\mathcal{L}_0 T_j} \rho_{j-1})}. \quad (43)$$

Next we sample the jump channel k_j from the probabilities

$$p(k|\bar{\rho}_j) = \frac{\text{Tr}(\mathcal{J}_k \bar{\rho}_j)}{\sum_q \text{Tr}(\mathcal{J}_q \bar{\rho}_j)}, \quad (44)$$

and use the outcome k_j to update the state to

$$\rho_j = \frac{\mathcal{J}_{k_j} \bar{\rho}_j}{\text{Tr}(\mathcal{J}_{k_j} \bar{\rho}_j)}. \quad (45)$$

We now continue in this way. At each step, the time T_j is sampled from $W(\tau|\rho_{j-1})$ and the next jump k_j is sampled from $p(k|\bar{\rho}_j)$. Combining these steps, the state is updated to

$$\rho_j = \frac{\mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j} \rho_{j-1}}{\text{Tr}(\mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j} \rho_{j-1})}. \quad (46)$$

This gives a way to efficiently obtain the correct weight (8) for each trajectory. Now we have to consider the evolution of the monitoring operator on the same trajectory, according to Eq. (36), where $\mathcal{M}_{x_j} = \mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j}$,

$$\xi_j = \frac{\partial_{\theta} (\mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j}) \rho_{j-1} + \mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j} \xi_{j-1}}{\text{Tr}(\mathcal{J}_{k_j} e^{\mathcal{L}_0 T_j} \rho_{j-1})}. \quad (47)$$

This equation can be read in direct comparison with the results of Ref. [31]. In our case, however, the evolution is written in the Gillespie formalism, rather than the time-discretized picture. The FI is then obtained by repeating this process over sufficiently many trajectories to get the average (33).

Importantly, many of the numerically heavier objects involved do not depend on the specific trajectory and can be pre-computed. To do so, in

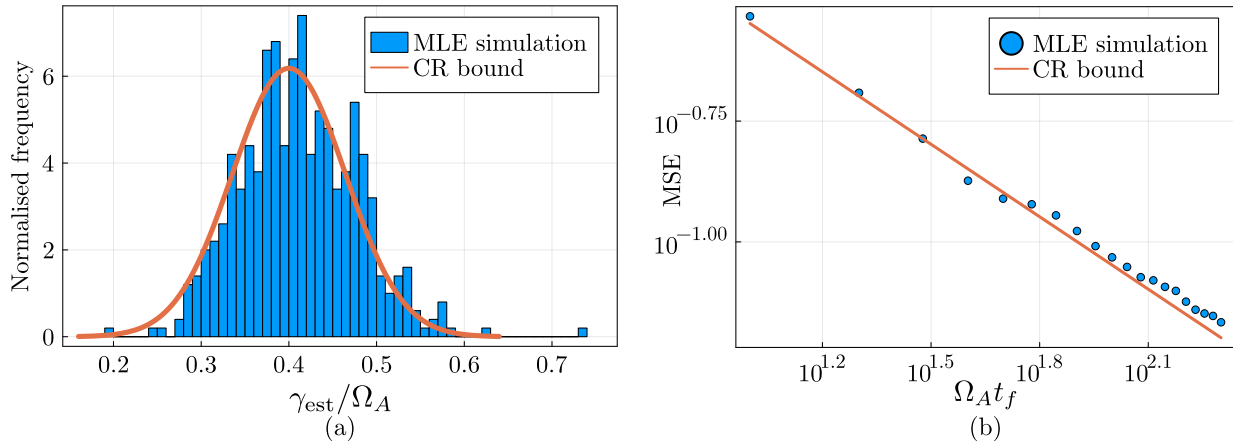


Figure 2: Estimation task results for the coupled qubit model. MLE estimation on 500 trajectories for a fixed final time 100 (a) yields normally distributed estimates; the variance is compatible with the prediction of the Cramér-Rao bound. (b): the variance of the MLE estimates is compared with the Cramér-Rao bound. As expected, perfect saturation is not reached for finite time, but the error on the estimates is comparable in order of magnitude with the bound. Parameters: $\Omega_A = \Omega_B = \bar{n} = 1$, $\gamma = 0.4$, $g = 0.01$, timestep $dt = 0.001$, γ increment for the derivatives $d\gamma = 0.0001$.

analogy with the Gillespie algorithm of Ref. [44], we define a list $\{T_\alpha\}$ of times, which has to (i) be sufficiently fine-grained to be able to resolve all relevant features of $W(\tau|\rho)$ and (ii) span times much larger than the average waiting times, in order to cover as faithfully as possible the structure of the WTD. At this point, we can pre-compute not only the superoperators $e^{\mathcal{L}_0\tau}$ required to evolve the state in Eq. (43), but also the superoperators $\partial_\theta \mathcal{J}_k$ for all the jump channels and $\partial_\theta e^{\mathcal{L}_0\tau}$, needed to evolve the monitoring operator ξ_t according to Eq. (47).

If the initial state is pure, all jump operators are one-dimensional projectors, and all jump channels are monitored, then the conditional state remains a pure state for the entire duration of the dynamics. It is therefore possible to write its evolution in terms of the effective Hamiltonian $H_e = H - i/2 \sum_k \mathcal{J}_k$, with significant savings in terms of memory and computational time.

A pseudocode description of the Gillespie-Fisher algorithm in this simpler case is given in Algorithm 1; an implementation in Julia can be found in Ref. [60]. The latter allows both for pure state evolution and for mixed states, making also possible to consider the partial monitoring of specific channels.

At the beginning of the dynamics, $\xi_0 = \partial_\theta \rho_0$ [Eq. (34)]; it follows that the monitoring operator is initially traceless (hence $F_0(\theta) = 0$), in accordance with the fact that no information is

available about θ at the beginning. However, the initial state may depend on θ ; e.g., one could be given the steady state of the system as the initial state, and the steady state ordinarily depends on the dynamical parameters. While this cannot play a role in the initial Fisher information, it may become relevant at later times, when the value of ξ is modified by the subsequent application of the jump and no-jump superoperators. In principle henceforth all that follows may depend upon the specific choice of the initial state ρ_0 . However, in cases wherein the information about the initial state is quickly lost during the time evolution, these effects eventually become vanishingly small.

Finally, we observe that the Fisher-Gillespie algorithm is, in effect, a Monte Carlo integration method for the Fisher information. By following the evolution, it allows to focus on the most relevant areas of the space of measurement records. Since the trajectories are generated independently of each other, the error in the evaluation of the Fisher information follows from the law of large numbers, scaling as $1/\sqrt{M}$, where M is the number of trajectories.

3.4 Fisher information rate

In this section, we discuss the Fisher information rate dF/dt of the measurement record of any quantum jump process. For this, we need to work

in the t_f -ensemble representation: the measurement record $\omega_{1:N}$ is equivalently represented by a sequence of discrete time steps of equal length dt . For each step t , a random variable X_t can take values in the set $\{0, 1, \dots, |\mathbb{M}|\}$, where the value 0 represents the absence of a jump in the step, and the other values index the possible jump channels.

Given the Fisher information in the measurement record up to time step T , $F_\theta(X_{1:T})$, one can update it to the next time step $T+1$ via the chain rule for the Fisher information [47, 54]

$$F_\theta(X_{1:T+1}) = F_\theta(X_{1:T}) + F_\theta(X_{T+1}|X_{1:T}), \quad (48)$$

where $F(X|Y)$ represents the conditional Fisher information for X given Y . The second term gives the variation of the Fisher information dF_T . By definition of the conditional Fisher information:

$$\begin{aligned} dF_T &= F(X_{T+1}|X_{1:T}) \\ &= \sum_{x_{1:T+1}} \frac{\Pr(x_{1:T}) (\partial_\theta \Pr(x_{T+1}|x_{1:T}))^2}{\Pr(x_{T+1}|x_{1:T})}. \end{aligned} \quad (49)$$

The update probability $\Pr(x_{T+1} = j|x_{1:T})$, for $j = 1, \dots, |\mathbb{M}|$, is given by

$$\Pr(x_{T+1} = j|x_{1:T}) = I_T^j dt, \quad (50)$$

and $\Pr(x_{T+1} = 0|x_{1:T}) = 1 - \sum_{j=1}^{|\mathbb{M}|} I_T^j dt$, where $I_T^j \equiv \text{Tr} [L_j \rho_T^{(c)} L_j^\dagger]$ is the stochastic current in channel j with corresponding jump operator L_j and $\rho_T^{(c)}$ the conditional normalized state at time T . Expanding the calculations (see Appendix G), we find the Fisher information rate to be

$$\frac{dF_T}{dt} = \sum_{k=1}^{|\mathbb{M}|} \mathbb{E} \left[\frac{1}{I_T^k} \left(\partial_\theta I_T^k \right)^2 \right], \quad (51)$$

where the expectation value is taken with respect to all the trajectories up to time step T . Hence: the Fisher information rate is only a function of the individual dependence on θ of the currents; the interplay among stochastic currents of different channels does not play a role at first order.

4 Information loss and compression

The measurement record $\omega_{1:N}$ in Eq. (1) represents all the information obtained from the quantum jump unraveling. The Fisher information

studied in the previous sections describes the best attainable precision, maximized over all estimator functions $\hat{\theta}(\omega_{1:N})$ that exploit the full data record. In practice, however, we may often not end up using the full data. For example, we might not have access to the jump symbols k_i , but only the time tags τ_i , or vice versa. Or we might want to use only a summary statistic, like the average waiting time. Discarding part of the record in this fashion constitutes a form of information compression. Since we are throwing away information, the corresponding Fisher information must be smaller; i.e., the estimation precision will be reduced. This can be formalized by the data-processing inequality for the Fisher information [47, 63, 64]: given any non-parameter-dependent map Φ acting on a trajectory $\omega_{1:N}$,

$$F_{\omega_{1:N}}(\theta) \geq F_{\Phi(\omega_{1:N})}(\theta). \quad (52)$$

This naturally establishes a trade-off between estimation precision and memory compression.

In this section we explore the loss of FI for different choices of data compression. It turns out that most expressions derived in Secs. 2 and 3 can be adapted to also encompass this possibility. The simplest case is that of partial monitoring; i.e., when one does not have access to all jump channels, or when they operate with finite efficiency. This case was already discussed around Eq. (5), and simply amounts to adding an efficiency η_k to each channel, or restricting the sum in Eq. (6) to a subset \mathbb{M} . One can think of the channel Φ in Eq. (52) as an operation that removes entries with $k_i \notin \mathbb{M}$ from $\omega_{1:N}$. Partial monitoring therefore modifies the main superoperators \mathcal{J}_k and \mathcal{L}_0 . But the general structure of the trajectory distributions remain unchanged, and hence all formulas from Secs. 2 and 3 still apply.

Another example of data compression is when we have access only to channel symbols k_i , but not time-tags τ_i , or vice-versa. Starting from the general trajectory probability (8) and marginalizing over all time-tags yields the symbol probability

$$\Pr(k_1, \dots, k_N) = \text{Tr} \{ M_{k_N} \cdots M_{k_1} \rho_0 \}, \quad (53)$$

where $M_k = -\mathcal{J}_k \mathcal{L}_0^{-1}$. This statistics was recently studied in detail in Ref. [50]. Renewal processes continues to be renewal in this case. Hence, the corresponding FI will be given exactly

by Eq. (22). For non-renewal processes, on the other hand, one can still apply the monitoring operator formalism in Eq. (36), which remains largely unchanged. All one has to do is replace \mathcal{M}_x with M_k .

If we instead marginalize (8) over the jump channels we obtain

$$\Pr(\tau_1, \dots, \tau_N) = \text{Tr} \{ \mathcal{J} e^{\mathcal{L}_0 \tau} \dots \mathcal{J} e^{\mathcal{L}_0 \tau} \rho_0 \}, \quad (54)$$

where $\mathcal{J} = \sum_{k \in \mathbb{M}} \mathcal{J}_k$. Notice how this result still depends on the choice of monitored channels \mathbb{M} , which influences both \mathcal{J} and \mathcal{L}_0 . Eq. (54) therefore assumes one does not know which specific channel clicked, but one knows it belongs to \mathbb{M} . In general, Eq. (54) will no longer be renewal, even if the original process was. The reason is because the renewal condition (10) is not, in general, satisfied for \mathcal{J} . Nonetheless, the monitoring operator formalism in Eq. (36) remains perfectly valid, provided one replaces $\mathcal{M}_x \rightarrow \mathcal{J} e^{\mathcal{L}_0 \tau}$.

So far, we have assumed compression schemes in which one has limited access to one aspect of the dataset or another, and which are specific to the form of the jump trajectory outcomes. Another important compression is to consider instead only summary statistics, analyzed in detail in Ref. [56]. In order to draw a relation, we here add an explanation of the effect of using such summary statistics for data originating from quantum jump trajectories. The broader case, independent of quantum jump dynamics, is discussed in Ref. [56].

Suppose the process has a single jump channel. Instead of considering estimators $\hat{\theta}(\tau_1, \dots, \tau_N)$ that are arbitrary functions of the entire dataset, one might consider estimators $\hat{\theta}(T)$ that depend only on the sample mean

$$T = \frac{\tau_1 + \dots + \tau_N}{N}. \quad (55)$$

This is another form of data compression and therefore must necessarily lead to a smaller Fisher information. To calculate this FI one must take into account the fact that the waiting times can be correlated with each other. As shown in [54, 56] the asymptotic FI, for large N , will be

$$F_T(\theta) \simeq N \left(\frac{\partial \mu}{\partial \theta} \right)^2 \frac{1}{\sigma^2 + 2 \sum_{i=1}^{N-1} C_i}, \quad (56)$$

where $\mu = \mathbb{E}[\tau_i]$ and $\sigma^2 = \text{Var}[\tau_i]$ are the mean and variance of the individual waiting times (assumed stationary), while $C_{i-j} = \text{cov}(\tau_i, \tau_j)$ is

the covariance between two waiting times. Note that C is represented with a single index because stationarity (time-translation invariance) is assumed. For renewal processes the waiting times become i.i.d. and $C_i = 0$. In this case, the WTD will be given by Eq. (12): $W(\tau) = \text{Tr} \{ \mathcal{J} e^{\mathcal{L}_0 \tau} \sigma \}$. And the mean and variance in Eq. (56) can be computed from the Laplace transform

$$\tilde{W}(s) = \int_0^\infty d\tau W(\tau) e^{-s\tau} = \text{Tr} \{ \mathcal{J} (s - \mathcal{L}_0)^{-1} \sigma \}, \quad (57)$$

as $\mu = -\tilde{W}'(0)$ and $\sigma^2 = \tilde{W}''(0) - \mu^2$. Conversely, for non-renewal processes $C_i \neq 0$ and the correlations will play a non-trivial role, since the second term in the denominator of Eq. (56) can be positive or negative, depending on the model. This is explored in more details in Refs. [54, 56].

5 Examples

In this section, we consider several examples to illustrate concepts explored throughout the paper. The examples are summarized in Table 2.

Example	Renewal	Multi-channel
Q. therm. 5.1	X	X
Res. fluor. 5.2	X	
Coupled q. 5.3		
Micromaser 5.4		X

Table 2: Summary of presented examples.

5.1 Qubit thermometry

We consider a qubit coupled to a thermal bath. The Hamiltonian of the isolated qubit system is

$$H = \frac{\omega}{2} \sigma_z + \frac{\Omega}{2} \sigma_x, \quad (58)$$

where ω is the detuning and Ω is the Rabi frequency. The coupling with the thermal bath is described by two jump operators, of the form

$$L_+ = \sqrt{\gamma \bar{n}} \sigma_+ \quad (59a)$$

$$L_- = \sqrt{\gamma(\bar{n} + 1)} \sigma_-, \quad (59b)$$

where γ is the emission rate and \bar{n} is the average occupation number of the relevant mode in the environment; the value of \bar{n} is connected to the temperature of the environment via the

Bose-Einstein distribution. We assume that both the jumps L_+ and L_- are observable, and we would like to estimate the parameter \bar{n} , of interest in thermometry applications. This

process is renewal.

The Fisher information about \bar{n} contained in the measurement record can be computed as in Eq. (17), and yields

$$F_{1:N}(\bar{n}) = \frac{N}{2\bar{n}(\bar{n}+1)} \left[\frac{(\bar{n}+1)^2 + \bar{n}^2}{\bar{n}(\bar{n}+1) + \frac{(\Omega/\gamma)^2}{2} \frac{1}{1+4\left(\frac{\omega/\gamma}{2\bar{n}+1}\right)^2}} + \frac{(\Omega/\gamma)^2}{\bar{n}(\bar{n}+1) \left(1 + 4\left(\frac{\omega/\gamma}{2\bar{n}+1}\right)^2\right) + \frac{(\Omega/\gamma)^2}{2}} \right]. \quad (60)$$

WTD and Fisher information are shown in Fig. 3.

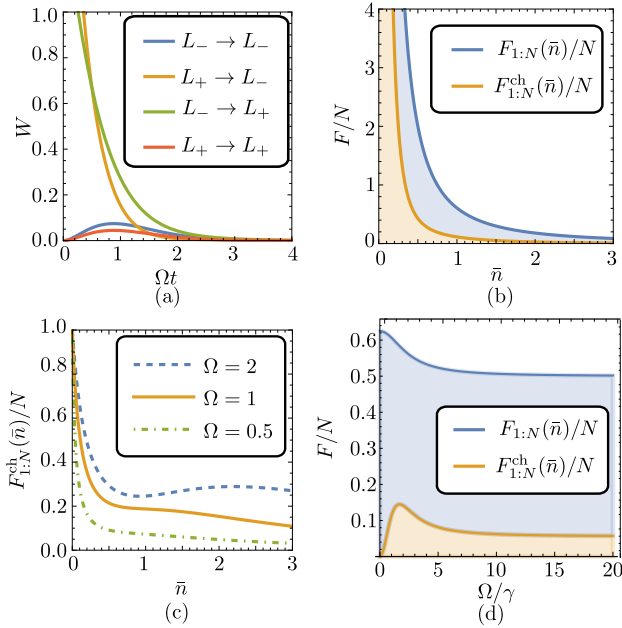


Figure 3: Thermometry with a qubit with observed raising and lowering operations. (a) The WTD for different observations. Parameters: $\omega = \Omega = \gamma = 1$ and $\bar{n} = 1.5$. (b) The Fisher information per observed jumps. Parameters: $\omega = \Omega = \gamma = 1$. The total rate (blue) is shown with the contribution from the channels F^{ch} alone (orange). (c) Proportion of total Fisher information in the channels, alone, for different Rabi oscillations Ω and $\omega = \gamma = 1$. (d) The Fisher information per observed jumps as a function of Ω , same parameters as in (b).

A special remark is in order for the limit $\Omega \rightarrow 0$ (see Fig. 3d). In this case, the jump operators L_+ and L_- satisfy the conditions for an exponential WTD discussed in Sec. D (the final states are orthogonal among each other, and the Hamiltonian commutes with the density matrix of the final states), as specified in the previous section. The

only non-zero WTDs are

$$W(\tau, L_+ | L_-) = \gamma \bar{n} e^{-\gamma \bar{n} \tau}; \quad (61a)$$

$$W(\tau, L_- | L_+) = \gamma(\bar{n} + 1) e^{-\gamma(\bar{n} + 1)\tau}. \quad (61b)$$

The Fisher information considerably simplifies, yielding

$$F_{1:N}(\bar{n}, \Omega = 0) = \frac{N}{2} \left[\frac{1}{\bar{n}} + \frac{1}{\bar{n} + 1} \right]. \quad (62)$$

If $\Omega \rightarrow 0$, the jump channels have to alternate between L_+ and L_- (no two jumps of the same kind can happen in a row). Therefore, no information about \bar{n} can be contained in the sequence of the jump channels, when one ignores the jump times: $F_{1:N}(\bar{n}) = F_{1:N}^{\text{times|ch}}(\bar{n})$.

The relative contribution of $F_{1:N}^{\text{ch}}$ to the total Fisher information is illustrated in Fig. 3b–c. One finds that, with all else equal, the sequence of observed channels dominates the Fisher information of the temperature for a cold system ($\bar{n} \rightarrow 0$) but is generally less important as temperature increases. Further, one observes that $F_{1:N}^{\text{ch}} \rightarrow 0$ as $\Omega \rightarrow 0$ (for finite temperature, $\bar{n} > 0$), reflecting the fact that $F_{1:N}(\bar{n}) = F_{1:N}^{\text{times|ch}}(\bar{n})$ in this case.

5.2 Resonant fluorescence

Consider again Eqs. (58) and (59), but assume $\omega = \bar{n} = 0$, so only emissions, $L = \sqrt{\Gamma}\sigma_-$, are observed.

The expression for the Fisher information w.r.t. the Rabi frequency Ω has been shown to be [27]

$$F(\Omega) = N \left(\frac{8}{\Gamma^2} + \frac{4}{\Omega^2} \right), \quad (63)$$

where N is the number of recorded jumps. The computation of the Fisher information of the

sample mean via Eq. (56) yields

$$F_T(\Omega) = \frac{N}{\Omega^2} \frac{4}{1 - 2(\Omega/\Gamma)^2 + 4(\Omega/\Gamma)^4}. \quad (64)$$

In Fig. 4, we represent the Fisher information for the entire measurement record, and for the sample mean for resonant fluorescence. We notice how, in general, the Fisher information of the sample mean is not comparable to the one of the full measurement record; while the former lower bounds the latter, the bound is clearly far from tight.

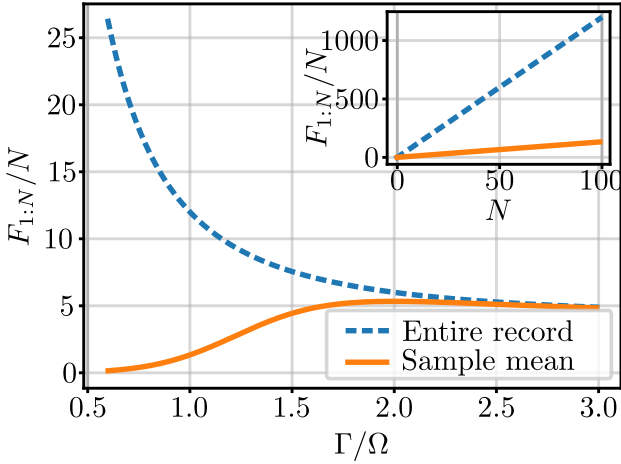


Figure 4: Fisher information of the entire measurement record, and of the sample mean for the resonant fluorescence process. In the main panel, the Fisher information rate for the Rabi frequency $\Omega = 1$ is represented as a function of the jump rate Γ ; in the inset plot, for fixed $\Gamma = \Omega = 1$, as a function of N , the number of recorded jumps.

5.3 Coupled qubits

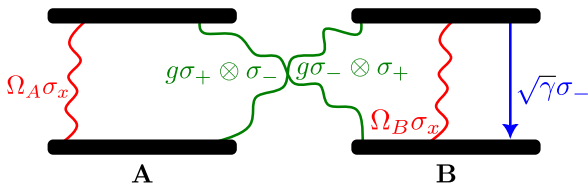


Figure 5: A depiction of the coupled-qubit model.

As a paradigmatic case of a non-renewal process, we consider two qubits, both undergoing Rabi oscillations, interacting with an exchange Hamiltonian, and such that one of them can “leak” excitations into a thermal bath. The

Hamiltonian of the system has the form

$$H = \Omega_A \sigma_x^A + \Omega_B \sigma_x^B + \omega_A \sigma_z^A + \omega_B \sigma_z^B + g \left(\sigma_+^A \sigma_-^B + \sigma_-^A \sigma_+^B \right), \quad (65)$$

where $\Omega_{A,B}$ are Rabi frequencies, $\omega_{A,B}$ are detuning terms, and g parametrizes the coupling strength. A tensor product with the identity is implied where not otherwise specified. There is only one jump operator, given by the emission from qubit B into the bath; the corresponding jump operator is

$$L = \sqrt{\gamma} \sigma_-^B. \quad (66)$$

The model is depicted in Fig. 5, and an example of a trajectory is given in 6a. It is clear that the model is non-renewal: while at each jump the state of qubit B is completely reset to the ground state, the state of qubit A is affected by the jump (which breaks all the entanglement between the qubits created by the Hamiltonian evolution), but not reset: there is quantum memory carried on at each jump.

We used the Fisher-Gillespie algorithm to compute the Fisher information of the measurement record for the estimation of γ in the coupled qubit model. The result is plotted in Fig. 6. It is apparent that the Fisher information scales asymptotically linearly with time; this is expected for any stochastic process whose correlations decay with time [54].

In Fig. 6c, we also highlighted one of the stochastic trajectories of $\text{Tr}[\xi_t]^2$. Remarkably, while the Fisher information (given by $\mathbb{E}[\text{Tr}[\xi_t]^2]$) always monotonically increases with time (up to small statistical fluctuations), on individual trajectories a decrease can be observed. In particular, the squared trace of the monitoring operator can be interpreted as an indicator of how atypical a trajectory is. Indeed, on a typical trajectory one would expect MLE to retrieve a very close estimate $\hat{\theta}$ to θ . By the MLE condition, $\text{Tr}[\xi_t]^2 = 0$. Hence, we expect $\text{Tr}[\xi_t]^2$ to be small for typical trajectories, and much larger for atypical trajectories.

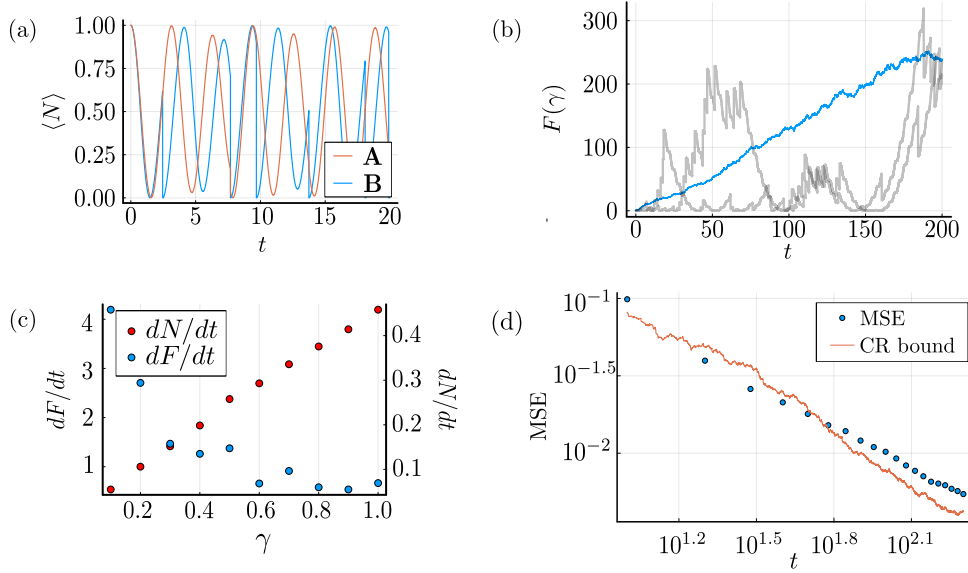


Figure 6: (a): a stochastic quantum trajectory for the two-qubit model, in terms of the expectation value of the number operator for each qubit. Note that, whereas qubit B undergoes a complete reset after each jump, qubit A , while affected by the jump, is not reset: this behavior clearly expresses the non-renewal nature of the process. To highlight this effect, the plot is obtained for an increased value $g = 0.1$, with respect to the parameters below. (b): the result of a Gillespie-Fisher algorithm calculation of the Fisher information for γ along the jump process. The average on the 2000 trajectories is represented by the red line. Some of the trajectories for $\text{Tr}[\xi_t]^2$ are represented in the background. (c): the asymptotic Fisher information rate for γ , as a function of γ itself, calculated via a linear fit of the Fisher information, represented along side the average jump rate. (d): mean squared error of MLE estimation for γ , compared with the numerical Cramér-Rao bound for 2000 trajectories. (Parameters: $\gamma = 0.4$, $\bar{n}_{th} = 1$, $\Omega_A = 1$, $\Omega_B = 1$, $g = 0.01$)

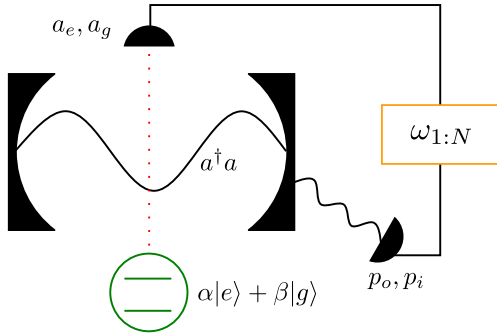


Figure 7: Illustration of the single-atom maser model. A beam of two-level atoms passes through a cavity where each atom interacts with the electromagnetic field. The energy of the atoms is measured when it exits the cavity. The cavity itself is coupled with the environment modeled as a thermal bath.

5.4 Single-atom maser

In this example, we apply the previously discussed concepts to the study of a model of relevant physical significance, the single-atom maser or micromaser [65], depicted in Fig. 7. While being developed in the last decades of the past century, the model is still the object of intensive study [66–68], and remains paradigmatic in the context of collision models. Physically, it consists of an optical cavity of high quality factor, across which a beam of highly excited Rydberg atoms passes. The typical energy of the excited states of those atoms with respect to their ground state is of the order of tens of GHz, i.e., in the microwave regime.

A beam of two-level atoms, all in the same initial state $|\psi\rangle = \alpha|e\rangle + \beta|g\rangle$, fly through a cavity. The beam is sparse enough that, on average, there is at most one atom inside the cavity. The atom interacts with the electromagnetic field inside the cavity via the Jaynes-Cummings Hamiltonian

$$H = \frac{g}{2} [a\sigma_+ + a^\dagger\sigma_-], \quad (67)$$

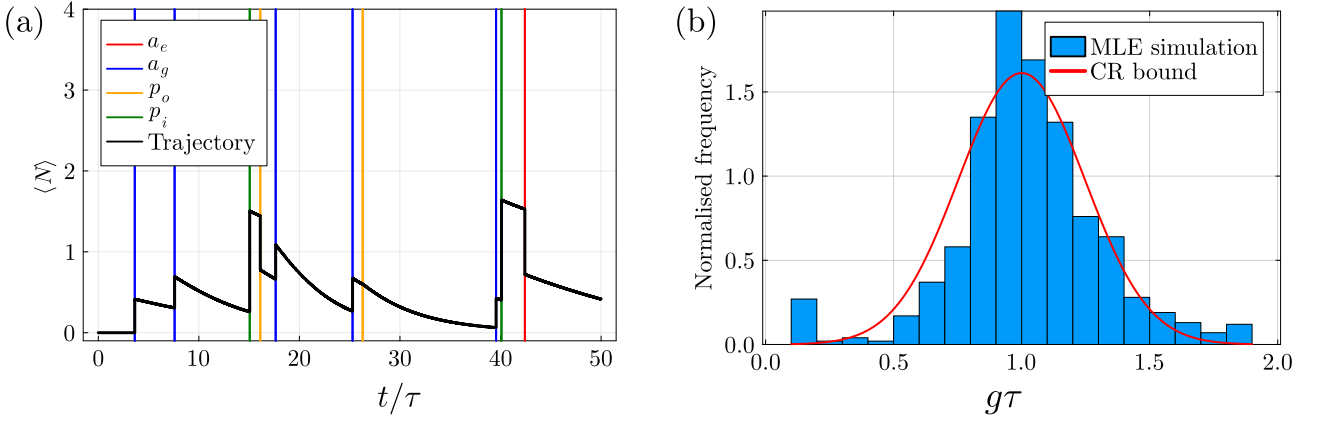


Figure 8: (a) Example of a trajectory for the number operator on the cavity, the different jump channels are highlighted. (b) Results of an MLE estimation task for g , compared with the Cramér-Rao bound obtained via the Fisher-Gillespie algorithm. Note that, formally, the convergence to a normal probability distribution is only guaranteed for an asymptotically large number of data points. Here, the convergence can already be observed for the final time considered for this plot. Parameters: $g = 1$, $\tau = 1$, $\theta = \pi/4$, $\bar{n}_{\text{th}} = 0.1$, $\gamma = 0.1$, e.m. field capped at the first 5 levels. Simulation obtained with 1000 trajectories, capped at final time 100τ .

where a is the annihilation operator for the single-mode field. As above, $\sigma_+ = |e\rangle\langle g|$, $\sigma_- = |g\rangle\langle e|$; g is the coupling constant. This Hamiltonian acts while the atom is inside the cavity, for time τ . Upon exiting the cavity, the atom is measured in the energy basis (projective measurement with $|e\rangle\langle e|$ and $|g\rangle\langle g|$).

The cavity itself is coupled to the environment, represented by a thermal bath; hence, photons leak from the cavity to the environment and enter the cavity from the environment.

In the analytical treatment of the micromaser model, one usually considers only the degrees of freedom of the field in the cavity, while the atomic ones are traced out. In this case, there are four possible incoherent jump channels, which we label as:

- a_e and a_g , corresponding to the atom exiting the cavity being detected respectively in the excited and the ground state;
- p_i and p_o , corresponding to a photon entering the cavity from the thermal bath or being emitted from the cavity to the thermal bath, respectively.

We can obtain the jump operators corresponding to the atomic measurements by tracing out the atomic d.o.f. from the unitary corresponding to the Jaynes-Cummings Hamiltonian [65]. Denot-

ing $\tilde{\mathbf{s}} = \sqrt{a^\dagger a}$ and $\mathbf{s} = \sqrt{a^\dagger a + 1}$ this yields:

$$L_{a_e} = \alpha \cos(g\tau\mathbf{s}) - i\beta \sin(g\tau\mathbf{s}) \mathbf{s}^{-1} a, \quad (68a)$$

$$L_{a_g} = \alpha a^\dagger \sin(g\tau\mathbf{s}) \mathbf{s}^{-1} + i\beta \cos(g\tau\tilde{\mathbf{s}}). \quad (68b)$$

The jump operators corresponding to the absorption by and emission from the cavity are given by

$$L_{p_i} = \sqrt{\gamma\bar{n}_{\text{th}}} a^\dagger, \quad (69a)$$

$$L_{p_o} = \sqrt{\gamma(\bar{n}_{\text{th}} + 1)} a, \quad (69b)$$

where \bar{n}_{th} is the average number of photons given by the Bose-Einstein distribution for the specific temperature of the bath, and γ is the emission rate.

This model is clearly non-renewal, as no jumps are able to reset the state to a fixed post-jump state. An example of a quantum trajectory for the model is shown in the right panel of Fig. 8.

The increased calculation efficiency given by the Gillespie-Fisher algorithm makes studying the Fisher information of a physical system such as the single-atom maser much more accessible. In particular, we are interested in the effect of the initial quantum state of the flying atoms on the Fisher information rate for the estimation of the coupling constant g . Let us define $\theta = \arccos(|\alpha|)$, so that $|\psi\rangle = \cos\theta|e\rangle + \sin\theta|g\rangle$, where we set the relative phase to zero. The behavior of the Fisher information rate for g , as a function of θ , is shown in Fig. 9.

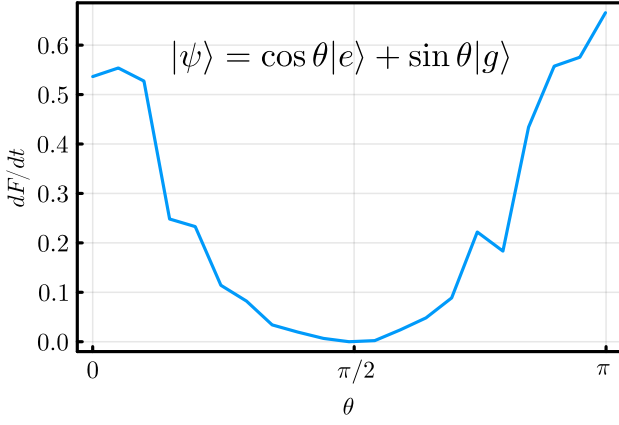


Figure 9: The Fisher information rate, as a function of θ , for the single-atom maser system. For each θ , the simulations were performed with $g = 1$, $\tau = 1$, $\gamma = 0.1$, $\bar{n}_{\text{th}} = 0.1$, the e.m. field limited to the first 5 levels, 100 trajectories, a time interval of $dt = 0.001$, and g variation for the derivatives $dg = 0.01$.

We observe that the highest Fisher information rate is obtained for $\theta \sim 0$ and $\theta \sim \pi$, corresponding to the incoming atoms being almost completely in the energy eigenstate $|e\rangle$. In contrast, the Fisher information rate appears to vanish for $\theta \sim \pi/2$, corresponding to the incoming atoms being almost completely in state $|g\rangle$, meaning that, in this regime, negligible information about g is contained in the measurement record. The state of the field is such that there is on average much less than one photon inside the cavity:

$$|f\rangle \sim (1 - \varepsilon) |0\rangle + \varepsilon |1\rangle, \quad (70)$$

with $\varepsilon \ll 1$. Applying the unitary evolution given by the Jaynes-Cummings Hamiltonian in Eq. (67) at first order, $U = \mathbb{I} - iH\tau$, one obtains a post-interaction state

$$|\Psi\rangle = (\mathbb{I} - iH\tau) |f\rangle \otimes |\psi\rangle. \quad (71)$$

If $|\psi\rangle = |g\rangle$, then $|\Psi\rangle = (1 - \varepsilon) |0\rangle \otimes |g\rangle + \varepsilon g/2 |0\rangle \otimes |e\rangle$, and we observe that the term bearing the g -dependence (hence the Fisher information) is suppressed in the $\varepsilon \rightarrow 0$ limit. On the contrary, if $|\psi\rangle = |e\rangle$, then $|\Psi\rangle \sim |0\rangle \otimes |e\rangle - ig\tau/2 |1\rangle \otimes |g\rangle$ in the $\varepsilon \rightarrow 0$ limit: the g -dependence is not suppressed. In other words, since there are almost always zero photons in the cavity field, some meaningful interaction can be obtained only if there is a significant overlap between the state of the incoming atom and $|e\rangle$, otherwise no interaction takes place, and the atom flies out of the cavity in the very same state in which it flew in. Since

the g -dependence is only picked up if the interaction happens, atoms starting in state $|g\rangle$ will not allow for any information about g to be collected.

6 Conclusions

In this work, we provided a complete framework for the estimation of parameters encoded in quantum jump processes, applicable in scenarios where an experimenter has access to jump times and multiple jump channels. First, we showed the asymptotic behavior of the Fisher information for multi-channel renewal processes in terms of the waiting time distribution, generalizing the results presented in Ref. [27], and discussing the partition of information between jump channels and jump times. We then discussed the more intricate case of non-renewal processes. Our treatment is based on the monitoring operator formalism, first introduced in Ref. [26], and computationally enhanced in Ref. [31]. Here, we introduced a new evolution method, the Gillespie-Fisher algorithm based on the quantum Gillespie method [44], which allows for an efficient calculation of the Fisher information on quantum trajectories in many cases of interest. As we showed, the Fisher information rate can be decomposed in terms of the stochastic currents, elucidating how it evolves in time. We showed how the monitoring operator can be employed to perform maximum likelihood estimation on quantum jump trajectories in a numerically stable way. To understand the behavior of the Fisher information under classical post-processing of the measurement record we provided an analysis based on the data-processing inequality. Finally, we presented a number of physical examples, of different levels of physical complexity, to showcase our findings.

These results provide a comprehensive set of tools for describing metrology of quantum jumps in any platform.

In recent years, metrological techniques based on the Fisher information have been extensively exploited in quantum thermodynamics, for example in the context of thermodynamic and kinetic uncertainty relations [69–71]. Our results, therefore, constitute the ideal toolbox for the computation of these inequalities in quantum jump processes.

We expect these tools to have a meaningful impact for developing new experiments — such as

developing novel quantum sensors — as well as theoretical work for understanding optimal methods for encoding information.

Acknowledgments

The authors thank Gerard Milburn for suggesting the single atom maser example, and Yoshihiko Hasegawa for insightful discussions about the waiting time distribution. MR thanks the ToCQS group at Trinity College Dublin for access to their computational resources, and the Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support. The research conducted in this publication was funded by the Irish Research Council under grant numbers IRCLA/2022/3922 and GOIPG/2022/2321. This publication was made possible through the support of Grant 62423 from the John Templeton Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation.

References

- [1] Harry L Van Trees. “Detection, estimation, and modulation theory, part i: detection, estimation, and linear modulation theory”. John Wiley & Sons. (2004).
- [2] Thomas M Cover and Joy A Thomas. “Information theory and statistics”. *Elements of information theory* **1**, 279–335 (1991).
- [3] Rafał Demkowicz-Dobrzański, Wojciech Górecki, and Mădălin Guță. “Multi-parameter estimation beyond quantum fisher information”. *J. Phys. A Math. Theor.* **53**, 363001 (2020).
- [4] Steven M Kay. “Fundamentals of statistical signal processing: estimation theory”. Prentice-Hall, Inc. (1993).
- [5] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Quantum metrology”. *Phys. Rev. Lett.* **96**, 010401 (2006).
- [6] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Advances in quantum metrology”. *Nat. Photon.* **5**, 222–229 (2011).
- [7] Matteo GA Paris. “Quantum estimation for quantum technology”. *Int. J. Quantum Inf.* **7**, 125–137 (2009).
- [8] Luca Pezze and Augusto Smerzi. “Quantum theory of phase estimation” (2014). [arXiv:1411.5164](https://arxiv.org/abs/1411.5164).
- [9] Luca Pezze, Augusto Smerzi, Markus K Oberthaler, Roman Schmied, and Philipp Treutlein. “Quantum metrology with non-classical states of atomic ensembles”. *Rev. Mod. Phys.* **90**, 035005 (2018).
- [10] Daniel Braun, Gerardo Adesso, Fabio Benatti, Roberto Floreanini, Ugo Marzolino, Morgan W Mitchell, and Stefano Pirandola. “Quantum-enhanced measurements without entanglement”. *Rev. Mod. Phys.* **90**, 035006 (2018).
- [11] Mankei Tsang, Howard M. Wiseman, and Carlton M. Caves. “Fundamental quantum limit to waveform estimation”. *Phys. Rev. Lett.* **106**, 090401 (2011).
- [12] Søren Gammelmark and Klaus Mølmer. “Fisher information and the quantum cramer-rao sensitivity limit of continuous measurements”. *Phys. Rev. Lett.* **112**, 170401 (2014).
- [13] Lewis A. Clark, Adam Stokes, and Almut Beige. “Quantum jump metrology”. *Phys. Rev. A* **99**, 022102 (2019).
- [14] Hendra I. Nurdin and Mădălin Guță. “Parameter estimation and system identification for continuously-observed quantum systems”. *Annu. Rev. Control* **54**, 295–304 (2022).
- [15] Daniel Burgarth, Vittorio Giovannetti, Airi N Kato, and Kazuya Yuasa. “Quantum estimation via sequential measurements”. *New J. Phys.* **17**, 113055 (2015).
- [16] Yaoling Yang, Victor Montenegro, and Abolfazl Bayat. “Extractable information capacity in sequential measurements metrology” (2023). [arXiv:2308.03370](https://arxiv.org/abs/2308.03370).
- [17] Antonella De Pasquale, Kazuya Yuasa, and Vittorio Giovannetti. “Estimating temperature via sequential measurements”. *Phys. Rev. A* **96**, 012316 (2017).
- [18] Mădălin Guță. “Fisher information and asymptotic normality in system identification for quantum Markov chains”. *Phys. Rev. A* **83**, 062324 (2011).
- [19] Merlijn van Horsen and Mădălin Guță. “Sanov and central limit theorems for output statistics of quantum Markov chains”. *J. Math. Phys.* **56**, 022109 (2015).

- [20] Stella Seah, Stefan Nimmrichter, Daniel Grimmer, Jader P Santos, Valerio Scarani, and Gabriel T Landi. “Collisional quantum thermometry”. *Phys. Rev. Lett.* **123**, 180602 (2019).
- [21] Eoin O’Connor, Bassano Vacchini, and Steve Campbell. “Stochastic collisional quantum thermometry”. *Entropy* **23**, 1634 (2021).
- [22] Alfred Godley and Madalin Guta. “Adaptive measurement filter: efficient strategy for optimal estimation of quantum markov chains”. *Quantum* **7**, 973 (2023).
- [23] Hideo Mabuchi. “Dynamical identification of open quantum systems”. *J. Opt., B Quantum Semiclass.* **8**, 1103 (1996).
- [24] Jay Gambetta and H. M. Wiseman. “State and dynamical parameter estimation for open quantum systems”. *Phys. Rev. A* **64**, 042105 (2001).
- [25] Luc Bouten, Madalin Guta, and Hans Maassen. “Stochastic Schrödinger equations”. *J. Phys. A: Math. Gen.* **37**, 3189 (2004).
- [26] Søren Gammelmark and Klaus Mølmer. “Bayesian parameter inference from continuously monitored quantum systems”. *Phys. Rev. A* **87**, 032115 (2013).
- [27] Alexander Holm Kiilerich and Klaus Mølmer. “Estimation of atomic interaction parameters by photon counting”. *Phys. Rev. A* **89**, 052110 (2014).
- [28] Dominic W. Berry, Mankei Tsang, Michael J. W. Hall, and Howard M. Wiseman. “Quantum Bell-Ziv-Zakai Bounds and Heisenberg Limits for Waveform Estimation”. *Phys. Rev. X* **5**, 031018 (2015).
- [29] Alexander Holm Kiilerich and Klaus Mølmer. “Bayesian parameter estimation by continuous homodyne detection”. *Phys. Rev. A* **94**, 032103 (2016).
- [30] Luis Cortez, Areeya Chantasri, Luis Pedro García-Pintos, Justin Dressel, and Andrew N. Jordan. “Rapid estimation of drifting parameters in continuously measured quantum systems”. *Phys. Rev. A* **95**, 012314 (2017).
- [31] Francesco Albarelli, Matteo AC Rossi, Dario Tamascelli, and Marco G Genoni. “Restoring heisenberg scaling in noisy quantum metrology by monitoring the environment”. *Quantum* **2**, 110 (2018).
- [32] Julia Amoros-Binefa and Jan Kołodyński. “Noisy atomic magnetometry in real time”. *New J. Phys.* **23**, 123030 (2021).
- [33] Julia Boeyens, Björn Annby-Andersson, Pharnam Bakhshinezhad, Géraldine Haack, Martí Perarnau-Llobet, Stefan Nimmrichter, Patrick P Potts, and Mohammad Mehboudi. “Probe thermometry with continuous measurements”. *New J. Phys.* **25**, 123009 (2023).
- [34] Enrico Rinaldi, Manuel González Lastre, Sergio García Herreros, Shahnawaz Ahmed, Maryam Khanahmadi, Franco Nori, and Carlos Sánchez Muñoz. “Parameter estimation by learning quantum correlations in continuous photon-counting data using neural networks” (2023). [arXiv:2310.02309v1](https://arxiv.org/abs/2310.02309v1).
- [35] M. B. Plenio and P. L. Knight. “The quantum-jump approach to dissipative dynamics in quantum optics”. *Rev. Mod. Phys.* **70**, 101–144 (1998).
- [36] Howard M Wiseman and Gerard J Milburn. “Quantum measurement and control”. Cambridge University Press. (2009).
- [37] A. Kubanek, M. Koch, C. Sames, A. Ourjoumtsev, P. W. H. Pinkse, K. Murr, and G. Rempe. “Photon-by-photon feedback control of a single-atom trajectory”. *Nature* **462**, 898–901 (2009).
- [38] R. Vijay, D. H. Slichter, and I. Siddiqi. “Observation of quantum jumps in a superconducting artificial atom”. *Phys. Rev. Lett.* **106**, 110502 (2011).
- [39] Karthik Siva, Gerwin Koolstra, John Steinmetz, William P Livingston, Debmalya Das, Larry Chen, John Mark Kreikebaum, Noah Stevenson, Christian Jünger, David I Santiago, et al. “Time-dependent hamiltonian reconstruction using continuous weak measurements” (2022). [arXiv:2211.07718](https://arxiv.org/abs/2211.07718).
- [40] Sébastien Gleyzes, Stefan Kuhr, Christine Guerlin, Julien Bernu, Samuel Deléglise, Ulrich Busk Hoff, Michel Brune, Jean-Michel Raimond, and Serge Haroche. “Quantum jumps of light recording the birth and death of a photon in a cavity”. *Nature* **446**, 297–300 (2007).
- [41] T. Rybarczyk, B. Peaudecerf, M. Penasa, S. Gerlich, B. Julsgaard, K. Mølmer, S. Gleyzes, M. Brune, J. M. Raimond, S. Haroche, and I. Dotsenko. “Forward-backward analysis of the photon-number

- evolution in a cavity”. *Phys. Rev. A* **91**, 062116 (2015).
- [42] Federico Carollo, Robert L. Jack, and Juan P. Garrahan. “Unraveling the large deviation statistics of markovian open quantum systems”. *Phys. Rev. Lett.* **122**, 130605 (2019).
- [43] Calum A Brown, Katarzyna Macieszczak, and Robert L Jack. “Unraveling metastable markovian open quantum systems”. *Phys. Rev. A* **109**, 022244 (2024).
- [44] Marco Radaelli, Gabriel T Landi, and Felix C Binder. “Gillespie algorithm for quantum jump trajectories”. *Phys. Rev. A* **110**, 062212 (2024).
- [45] Ronald A Fisher. “On the mathematical foundations of theoretical statistics”. *Philos. Trans. R. Soc.* **222**, 309–368 (1922).
- [46] Harald Cramér. “Mathematical methods of statistics”. Princeton University Press. (1946).
- [47] Pablo Zegers. “Fisher information properties”. *Entropy* **17**, 4918–4939 (2015).
- [48] B. R. Mollow. “Pure-state analysis of resonant light scattering: Radiative damping, saturation, and multiphoton effects”. *Phys. Rev. A* **12**, 1919–1943 (1975).
- [49] Heinz-Peter Breuer, Francesco Petruccione, et al. “The theory of open quantum systems”. Oxford University Press. (2002).
- [50] Gabriel T Landi. “Patterns in the jump-channel statistics of open quantum systems” (2023). [arXiv:2305.07957](https://arxiv.org/abs/2305.07957).
- [51] Juan P. Garrahan and Igor Lesanovsky. “Thermodynamics of quantum jump trajectories”. *Phys. Rev. Lett.* **104**, 160601 (2010).
- [52] Sarah E. Marzen and James P. Crutchfield. “Informational and causal architecture of discrete-time renewal processes”. *Entropy* **17**, 4891–4917 (2015).
- [53] Ryan G. James, John R. Mahoney, Christopher J. Ellison, and James P. Crutchfield. “Many roads to synchrony: Natural time scales and their algorithms”. *Phys. Rev. E* **89**, 042135 (2014).
- [54] Marco Radaelli, Gabriel T Landi, Kavan Modi, and Felix C Binder. “Fisher information of correlated stochastic processes”. *New J. Phys.* **25**, 053037 (2023).
- [55] Paul M. Riechers. “Ultimate limit on learning non-markovian behavior: Fisher information rate and excess information” (2023). [arXiv:2310.03968](https://arxiv.org/abs/2310.03968).
- [56] Joseph A. Smiga, Marco Radaelli, Felix C. Binder, and Gabriel T. Landi. “Stochastic metrology and the empirical distribution”. *Phys. Rev. Res.* **5**, 033150 (2023).
- [57] Aad W Van der Vaart. “Asymptotic statistics”. Volume 3. Cambridge University Press. (2000).
- [58] Bassano Vacchini. “Quantum renewal processes”. *Sci. Rep.* **10**, 5592 (2020).
- [59] K. Micadei, D. A. Rowlands, F. A. Pollock, L. C. Céleri, R. M. Serra, and K. Modi. “Coherent measurements in quantum metrology”. *New J. Phys.* **17**, 023057 (2015).
- [60] “Fisher-Gillespie algorithm in Julia” (2024). <https://github.com/marcoradaelli/WaitingTimeMetrology>.
- [61] Daniel T Gillespie. “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions”. *J. Comput. Phys.* **22**, 403–434 (1976).
- [62] Daniel T Gillespie. “Exact stochastic simulation of coupled chemical reactions”. *J. Phys. Chem.* **81**, 2340–2361 (1977).
- [63] Ram Zamir. “A proof of the fisher information inequality via a data processing argument”. *IEEE Trans. Inf. Theory* **44**, 1246–1250 (1998).
- [64] Christopher Ferrie. “Data-processing inequalities for quantum metrology”. *Phys. Rev. A* **90**, 014101 (2014).
- [65] Marlan O. Scully and M. Suhail Zubairy. “Theory of the micromaser”. Pages 383–401. Cambridge University Press. (1997).
- [66] Ufuk Korkmaz and Deniz Türkpençe. “Quantum collisional classifier driven by information reservoirs”. *Phys. Rev. A* **107**, 012432 (2023).
- [67] Chimdessa Gashu Feyisa and HH Jen. “A photonic engine fueled by quantum-correlated atoms”. *New J. Phys.* **26**, 033038 (2024).
- [68] A. B. Mikhalychev, S. V. Vlasenko, and S. Ya. Kilin. “Phase-space dynamics of a single-atom laser: Nonclassicality and bistability”. *Phys. Rev. A* **105**, 063723 (2022).
- [69] Yoshihiko Hasegawa and Tan Van Vu. “Uncertainty relations in stochastic processes: An information inequality approach”. *Phys. Rev. E* **99**, 062126 (2019).

- [70] Tan Van Vu and Keiji Saito. “Thermodynamics of precision in markovian open quantum dynamics”. *Phys. Rev. Lett.* **128**, 140602 (2022).
- [71] Saulo V Moreira, Marco Radaelli, Alessandro Candeloro, Felix C Binder, and Mark T Mitchison. “Precision bounds for multiple currents in open quantum systems”. *Phys. Rev. E* **111**, 064107 (2025).
- [72] A. Bhattacharyya. “On Some Analogues of the Amount of Information and Their Use in Statistical Estimation”. *Sankhyā: The Indian Journal of Statistics (1933-1960)* **8**, 1–14 (1946). url: <https://www.jstor.org/stable/25047921>.
- [73] R. G. Jarrett. “Bounds and Expansions for Fisher Information when the Moments are Known”. *Biometrika* **71**, 101–113 (1984).
- [74] Manuel S. Stein and Josef A. Nossek. “A Pessimistic Approximation for the Fisher Information Measure”. *IEEE Trans. Signal Process.* **65**, 386–396 (2017).
- [75] Alexander Ly, Maarten Marsman, Josine Verhagen, Raoul P.P.P. Grasman, and Eric-Jan Wagenmakers. “A tutorial on fisher information”. *J. Math. Psychol.* **80**, 40–55 (2017).

A Form of the jump operators for renewal processes

In this appendix, we prove that, in a renewal master equation in which all the jumps are monitored, the jump operators have to have a specific form. In particular, let L_k be a jump operator satisfying the renewal condition

$$\frac{L_k \rho L_k^\dagger}{\text{Tr}[L_k \rho L_k^\dagger]} = \sigma_k, \quad (72)$$

where σ_k does not depend on ρ . Then, we prove that $L_k = c_k |\mu_k\rangle \langle \nu_k|$, where $c_k = \sqrt{\gamma_k}$.

Let us perform a singular value decomposition of the operator L_k :

$$L_k = \sum_{\alpha} c_{k,\alpha} |\mu_{k,\alpha}\rangle \langle \nu_{k,\alpha}|, \quad (73)$$

where $\{c_{k,\alpha}\}$ are the singular values of L_k , and $\{|\mu_{k,\alpha}\rangle\}$ and $\{|\nu_{k,\alpha}\rangle\}$ are, respectively, the orthonormal bases of the left-singular and right-singular vectors.

The ratio defining the renewal condition can be written as

$$\frac{L_k \rho L_k^\dagger}{\text{Tr}[L_k \rho L_k^\dagger]} = \sum_{\alpha\beta} \frac{c_{k,\alpha} c_{k,\beta}^* |\mu_{k,\alpha}\rangle \langle \nu_{k,\alpha}| \rho |\nu_{k,\beta}\rangle \langle \mu_{k,\beta}|}{\sum_a |c_{k,a}|^2 \langle \nu_{k,\alpha}| \rho |\nu_{k,\alpha}\rangle} = \sigma_k. \quad (74)$$

We observe that, in order for σ_k not to depend on ρ , only one of the singular values $\{c_{k,\alpha}\}$ can be non-zero. It follows that the jump operator L_k can be written in the simpler form

$$L_k = c_k |\mu_k\rangle \langle \nu_k|. \quad (75)$$

By plugging this equation again in the definition of renewal jump, we obtain that $\sigma_k = |\mu_k\rangle \langle \nu_k|$. This implies that, in a renewal process in which all jumps are monitored, the destination state of every jump has to be a pure state.

B Fisher information for renewal processes

In this Appendix, we prove the expression for the Fisher information of renewal processes, Eq. (17). We first show an intermediate result on the dynamical activity of channels.

Lemma 1. *Let ρ be the steady state of a GKSL master equation, such that $\mathcal{L}\rho = 0$. Let $f_k = \text{Tr}[\mathcal{J}_k \rho]$ be the dynamical activity of channel k and define*

$$p_k = \frac{f_k}{\sum_j f_j}. \quad (76)$$

For renewal processes, p_k (or f_k) is the steady state of the Markov chain

$$\sum_{i \in \mathbb{M}} p(j|i) f_i = f_j, \quad (77)$$

where

$$p(j|i) = \int_0^\infty d\tau \text{Tr}[\mathcal{J}_j e^{\mathcal{L}_0 \tau} \sigma_i] = \int_0^\infty d\tau W(\tau, j|i) \quad (78)$$

is the probability of having a jump in channel j after a jump in channel i , with no other jumps in between, regardless of how much time elapses.

Proof. It suffices to prove the result for f_k . We assume that \mathcal{L}_0 is invertible; this condition is tantamount to the normalisability of the waiting time distribution, and means that a jump will always occur in the end with unit probability. Then, $p(j|i)$ can be rewritten as

$$p(j|i) = -\text{Tr}[\mathcal{J}_j \mathcal{L}_0^{-1} \sigma_i]. \quad (79)$$

By Eq. (10) one then has that

$$\begin{aligned}\sum_{i \in \mathbb{M}} p(j|i) f_i &= - \sum_{i \in \mathbb{M}} \text{Tr} \{ \mathcal{J}_j \mathcal{L}_0^{-1} \sigma_i \} \text{Tr} \{ \mathcal{J}_i \rho_{\text{ss}} \} \\ &= - \sum_{i \in \mathbb{M}} \text{Tr} \{ \mathcal{J}_j \mathcal{L}_0^{-1} \mathcal{J}_i \rho_{\text{ss}} \}.\end{aligned}$$

We now use the fact that $\sum_{i \in \mathbb{M}} \mathcal{J}_i = \mathcal{L} - \mathcal{L}_0$, and $\mathcal{L} \rho_{\text{ss}} = 0$. Hence, we arrive at $\sum_{i \in \mathbb{M}} p(j|i) f_i = \text{Tr} \{ \mathcal{J}_j \rho_{\text{ss}} \} = f_j$. \square

As defined, the measurement record probability distribution $\text{Pr}(\omega_{1:N})$ in Eq. (11) is not a stationary process. This is because it is conditioned on an arbitrary initial channel k_0 . We can make it stationary if we modify it as

$$\text{Pr}(\omega_{1:N}) = \sum_{k_0} W(\tau_N, k_N | k_{N-1}) \cdots W(\tau_1, k_1 | k_0) p_{k_0}, \quad (80)$$

where p_{k_0} is given in Eq. (76). To illustrate the stationarity, consider $\omega_2 = (k_1, \tau_1, k_2, \tau_2)$. Then

$$\begin{aligned}\sum_{k_1} \int_0^\infty d\tau_1 P(k_1, \tau_1, k_2, \tau_2) &= \sum_{k_1} \int_0^\infty d\tau_1 \sum_{k_0} W(\tau_2, k_2 | k_1) W(\tau_1, k_1 | k_0) p_{k_0} \\ &= \sum_{k_0, k_1} W(\tau_2, k_2 | k_1) p(k_1 | k_0) p_{k_0} \\ &= \sum_{k_1} W(\tau_2, k_2 | k_1) p_{k_1} = P(k_2, \tau_2).\end{aligned}$$

An important consequence of stationarity is the following lemma.

Lemma 2. *Let $g(\tau_i, k_{i-1}, k_i)$ denote a function depending only on two consecutive jumps, and the time τ_i between them. Then, any average over a N -steps trajectory simplifies to*

$$\begin{aligned}\mathbb{E}[g(\tau_i, k_{i-1}, k_i)] &= \sum_{k_0 \dots k_N \in \mathbb{M}} \int_0^\infty d\tau_1 \cdots \int_0^\infty d\tau_N \text{Pr}(k_0, (\tau_1, k_1), \dots, (\tau_N, k_N)) g(\tau_i, k_{i-1}, k_i) \\ &= \sum_{k_{i-1}, k_i \in \mathbb{M}} \int_0^\infty d\tau_i W(\tau_i, k_i | k_{i-1}) p_{k_{i-1}} g(\tau_i, k_{i-1}, k_i).\end{aligned} \quad (81)$$

We can now turn to the calculation of the Fisher information in a measurement record $\omega_{1:N}$ for a renewal process. Combining Eq. (3) and Eq. (80) we get

$$F_{1:N}(\theta) = -\mathbb{E} \left[\sum_{j=2}^N \frac{\partial^2}{\partial \theta^2} \log W(\tau_j, k_j | k_{j-1}) + \frac{\partial^2}{\partial \theta^2} \log \sum_{k_0} W(\tau_1, k_1 | k_0) p_{k_0} \right]. \quad (82)$$

The last term stems from our modification of $\text{Pr}(\omega_{1:N})$ in Eq. (80). Notice that this is independent of N , while the first term will be extensive in N . The second term, therefore, leads to a negligible contribution. As for the first term, we can use Eq. (81). This will give rise to N identical terms, leading us exactly to the expression in Eq. (17).

C MLE estimation for the qubit thermometry process

In this Appendix, we propose an estimation task for a renewal system, to illustrate the results of Sec. 2. We consider the qubit thermometer, whose Hamiltonian and jump operators are given in Sec. 5.1, and we aim to estimating the average occupation number \bar{n} .

For a renewal system, the log-likelihood function is easily expressed in terms of the WTD, as in Eq. (19); as expected, MLE estimation asymptotically saturates the Cramér-Rao bound (see Fig. 10).

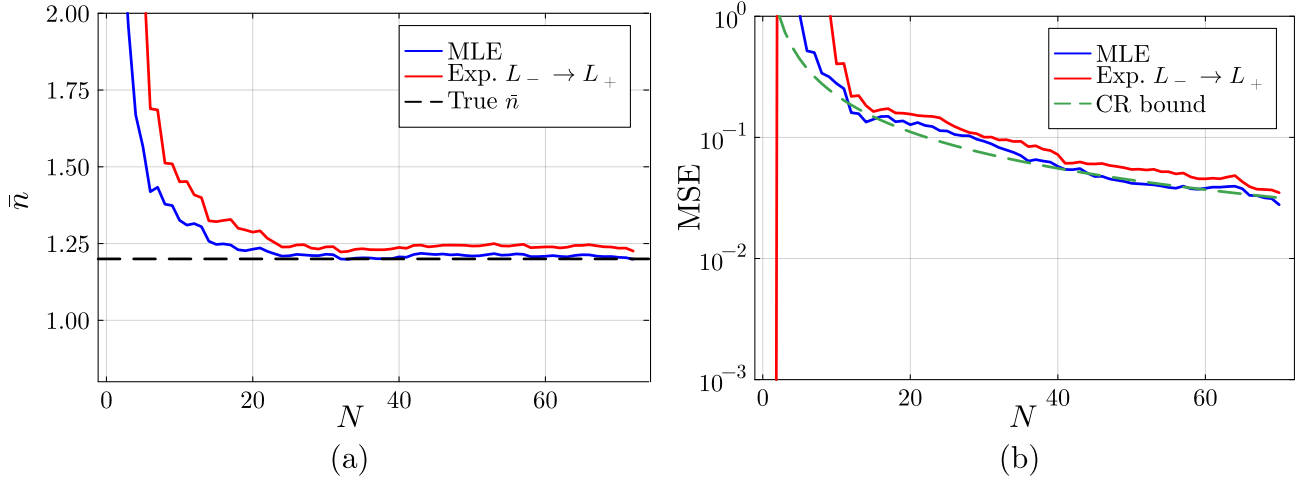


Figure 10: An estimation task performed on the qubit thermometer. On the left, the expectation value of the estimated \bar{n} as a function of the number of considered jumps; on the right, the mean squared error for the two estimation strategies. Parameters: $\omega = 1$, $\Omega = 0.1$, $\gamma = 0.02$, $\bar{n} = 1.2$; 100 trajectories.

We also consider a different estimator, given by the average waiting time between a L_- and a L_+ jumps. The expectation value of the waiting time between these two jumps is given by

$$\langle \tau \rangle_{-+} = \int_0^\infty d\tau \tau W(\tau | -, +) = - \frac{\text{Tr} [\mathcal{J}_+ (\mathcal{L}_0)^{-2} \mathcal{J}_- \rho_0]}{\text{Tr} [\mathcal{J}_+ (\mathcal{L}_0)^{-1} \mathcal{J}_- \rho_0]}, \quad (83)$$

and is, of course, a function of \bar{n} . Given a measurement record, one can compute the average waiting time between two jumps L_- , L_+ , then numerically invert $\langle \tau \rangle_{-+}$ as a function of \bar{n} to find the estimate. Clearly, the performances of the estimator based on the average waiting time are worse than the ones of the MLE, because the former does not have access to all the information contained in the shape of the waiting time distributions.

D Single exponential WTD and classical master equations

In general, carrying out the integral in Eq. (17) is difficult because the WTD is usually a sum of (possibly complex) exponentials in τ . A case that offers a significant simplification is when it is a single decaying exponential:

$$W(\tau, k|q) = p(k|q) b_{kq} e^{-b_{kq}\tau}, \quad (84)$$

for coefficients $p(k|q) > 0$ and $b_{kq} > 0$ that may both depend on θ . This, as we will discuss below, is the case of classical (Pauli) master equations describing incoherent dynamics. The coefficient $p(k|q)$ is defined to match Eq. (23), so that normalization implies $\sum_k p(k|q) = 1$ for all q . Carrying out the integral in Eq. (17) allows us to write it in the form of Eq. (21), with

$$F_{1:N}^{\text{ch}}(\theta) = N \sum_{k,q \in \mathbb{M}} p(k|q) p_q (\partial_\theta \ln p(k|q))^2, \quad (85a)$$

$$F_{1:N}^{\text{times|ch}}(\theta) = N \sum_{k,q \in \mathbb{M}} p(k|q) p_q (\partial_\theta \ln b_{kq})^2. \quad (85b)$$

The Fisher information contained in the channels is therefore the same as in the general renewal case of Eq. (22), but the second contribution now simplifies considerably.

We now discuss physical scenarios in which the single-exponential assumption of Eq. (84) will appear. Recall that for renewal processes $L_k = \sqrt{\gamma_k} |\mu_k\rangle \langle \nu_k|$, where $|\nu_k\rangle$ and $|\mu_k\rangle$ are generic (normalized) kets; i.e., they in principle do not need to be elements of the same basis. A WTD with a single exponential

will occur (i) $|\nu_k\rangle$ are orthogonal among each other and (ii) satisfy $[H, |\nu_k\rangle\langle\nu_k|] = 0$ for all k . The $|\mu_k\rangle$ can still be arbitrary. This therefore means that the post-jump states $\sigma_k = |\mu_k\rangle\langle\mu_k|$ can have arbitrary coherences, but the pre-jump states must be diagonal in the eigenbasis of H . Referring to Eq. (16), this implies that

$$\Psi(\tau, k|q) = \sqrt{\gamma_k} e^{-iH_{kk}\tau} e^{-\gamma_k\tau/2} \langle\nu_k|\mu_q\rangle, \quad (86)$$

where $H_{kk} = \langle\nu_k|H|\nu_k\rangle$. It then follows that

$$W(\tau, k|q) = \gamma_k e^{-\gamma_k\tau} |\langle\nu_k|\mu_q\rangle|^2, \quad (87)$$

which has exactly the form in Eq. (84), with transition probability between jumps $p(k|q) = |\langle\nu_k|\mu_q\rangle|^2$ and decay rate $b_{kq} = \gamma_k$. Interestingly, in this case b_{kq} only depends on the post-jump state, so that Eq. (26) simplifies to

$$W(\tau|k, q) \equiv W(\tau|k) = \gamma_k e^{-\gamma_k\tau}. \quad (88)$$

That is, the WTD between jumps in channels $q \rightarrow k$ depends only on the final jump state k .

In practice, the above scenario appears most often when dealing with incoherent systems described by classical (Pauli) master equations. Let $H|i\rangle = E_i|i\rangle$. Pauli equations occur when the jump operators have the form $L_{ji} = \sqrt{R_{ji}}|j\rangle\langle i|$, where R_{ji} is the transition rate to go from $|i\rangle \rightarrow |j\rangle$ (with $i \neq j$). Because the transitions occur between energy eigenstates, the long-time dynamics will be incoherent in the basis $|i\rangle$. In fact, starting from the original master Eq. (4), one may show that the probabilities $p_i = \langle i|\rho|i\rangle$ will satisfy the Pauli rate equation

$$\frac{dp_i}{dt} = \sum_j R_{ij}p_j - \Gamma_i p_i, \quad \Gamma_i = \sum_j R_{ji}. \quad (89)$$

In quantum master equations we label the jumps by the channels k appearing in the jump operators L_k . For classical master equations the index k is a composite index $k \equiv i \rightarrow j$, representing the initial and final state of that jump operator, $L_{ji} = \sqrt{R_{ji}}|j\rangle\langle i|$. If all channels are monitored, a jump $q = n \rightarrow m$ can be followed by a jump $k = i \rightarrow j$ if and only if $m \equiv i$. The corresponding WTD is readily found from Eq. (15) or [(87)], and reads

$$W(\tau, i \rightarrow j|n \rightarrow m) = \delta_{m,i} R_{ji} e^{-\Gamma_i\tau}. \quad (90)$$

This is clearly in the form of Eq. (84), so the results in Eq. (85) apply. However, some bookkeeping is required to translate the “ $q \rightarrow k$ ” notation to “ $n \rightarrow m = i \rightarrow j$.” The transition probability $p(k|q)$ in Eq. (23) becomes $p(i \rightarrow j|n \rightarrow m) = \delta_{m,i} \frac{R_{ji}}{\Gamma_i}$, while b_{kq} is replaced with $b_{i \rightarrow j, n \rightarrow m} = \Gamma_i$.

We also need $p_q \equiv p_{n \rightarrow m}$, which is the solution of Eq. (24):

$$\sum_n \frac{R_{ji}}{\Gamma_i} p_{n \rightarrow i} = p_{i \rightarrow j}. \quad (91)$$

As one may verify, the solution reads

$$p_{n \rightarrow m} = \frac{R_{mn} p_n^{\text{ss}}}{A}, \quad (92)$$

where $A = \sum_i \Gamma_i p_i^{\text{ss}}$ is the steady-state dynamical activity (number of jumps per unit time) and p_i^{ss} is the steady-state of the master equation (89); i.e., the solution of $\sum_j R_{ij} p_j^{\text{ss}} = \Gamma_i p_i^{\text{ss}}$. Physically, $p_{n \rightarrow m}$ is the probability of observing pairs of states (n, m) in the jump sequence. Plugging this in Eq. (85) then finally yields, after some simplifications (since $\Gamma_i = \sum_j R_{ji}$):

$$F_{1:N}(\theta) = \frac{N}{A} \sum_{i,j} p_i^{\text{ss}} \frac{(\partial_\theta R_{ji})^2}{R_{ji}}, \quad (93a)$$

$$F_{1:N}^{\text{ch}}(\theta) = \frac{N}{A} \sum_{i,j} R_{ji} p_i^{\text{ss}} (\partial_\theta \ln R_{ji}/\Gamma_i)^2, \quad (93b)$$

$$F_{1:N}^{\text{times|ch}}(\theta) = \frac{N}{A} \sum_i p_i^{\text{ss}} \frac{(\partial_\theta \Gamma_i)^2}{\Gamma_i}, \quad (93c)$$

and $F_{1:N}^{\text{ch}}(\theta) = F_{1:N}(\theta) - F_{1:N}^{\text{times|ch}}(\theta)$.

E Bound on Fisher information

Even when Fisher information is difficult to calculate, it can be easier to calculate some bounds on the Fisher information. For example, if the theory is described by a probability distribution of a real variable $P(x)dx$ with mean μ and variance σ^2 , then $F \geq \frac{(\partial_\theta \mu)^2}{\sigma^2}$. It is often the case that the mean and variance are relatively easy to calculate. In this appendix, a similar bound is derived in the case of WTDs of renewal processes.

The bound arises from the Cauchy-Schwarz inequality,

$$\int d\tau f^2(\tau)p(\tau) \times \int d\tau g^2(\tau)p(\tau) \leq \left[\int d\tau f(\tau)g(\tau)p(\tau) \right]^2, \quad (94)$$

where f and g are any real functions, while p is a non-negative function. For our purposes, the integral is evaluated over $[0, \infty)$, representing the time intervals. Observe that this inequality follows from the functional inner product, $\langle f, g \rangle = \int d\tau f(\tau)g(\tau)p(\tau)$. Here,

$$f(\tau) = \partial_\theta \log p(x), \quad (95a)$$

$$p(\tau) = W(\tau, k|q). \quad (95b)$$

Evaluating the terms in the Cauchy-Schwarz inequality, first,

$$\int d\tau f^2(\tau)p(\tau) = \int d\tau \frac{[\partial_\theta W(\tau, k|q)]^2}{W(\tau, k|q)}. \quad (96)$$

Then

$$\begin{aligned} \int d\tau f(\tau)g(\tau)p(\tau) &= \int d\tau g(\tau)\partial_\theta W(\tau, k|q) \\ &= \int d\tau \partial_\theta \{g(\tau)W(\tau, k|q)\} - d\tau [\partial_\theta g(\tau)]W(\tau, k|q) \\ &= \partial_\theta \{\mathbb{E}[g(\tau)|k, q] p(k|q)\} - \mathbb{E}[\partial_\theta g(\tau)|k, q] p(k|q). \end{aligned}$$

The above expression can be simplified by demanding that the (conditional) expectation of $g(\tau)$ vanish. For a function $\tilde{g}(\tau)$ that does not satisfy this condition, one can shift the function:

$$\tilde{g}(\tau) \rightarrow g(\tau) = \tilde{g}(\tau) - \mathbb{E}[\tilde{g}(\tau)|k, q]. \quad (97)$$

So

$$\int d\tau f(\tau)g(\tau)p(\tau) = -\mathbb{E}[\partial_\theta g(\tau)|k, q] p(k|q). \quad (98)$$

Finally,

$$\int d\tau g^2(\tau)p(\tau) = \text{Var}[g(\tau)|k, q] p(k|q), \quad (99)$$

where $\mathbb{E}[g^2(\tau)|k, q] = \text{Var}[g(\tau)|k, q]$ because $\mathbb{E}[g(\tau)|k, q] = 0$.

The resulting inequality Eq. (96) can be written as

$$\int d\tau \frac{[\partial_\theta W(\tau, k|q)]^2}{W(\tau, k|q)} \geq \frac{\mathbb{E}[\partial_\theta g(\tau)|k, q]^2}{\text{Var}[g(\tau)|k, q]} p(k|q). \quad (100)$$

Performing a weighted sum of the above inequality over k and q with weights p_q , yields the lower bound on Fisher information Eq. (17),

$$F_{1:N}(\theta)/N \geq \sum_{k,q} \frac{\mathbb{E}[\partial_\theta g(\tau)|k, q]^2}{\text{Var}[g(\tau)|k, q]} p(k|q) p_q, \quad (101)$$

where $g(\tau)$ is any function of τ with $\mathbb{E}[g(\tau)|k, q] = 0$.

One may also consider the Fisher information conditioned on the knowledge of the jumps. The derivation of a lower bound follows as above in Eq. (100) by considering k to be trivial (so, e.g., $p(k|q) \rightarrow 1$) and $q \rightarrow (k, q)$,

$$\int d\tau \frac{[\partial_\theta W(\tau|k, q)]^2}{W(\tau|k, q)} \geq \frac{\mathbb{E}[\partial_\theta g(\tau)|k, q]^2}{\text{Var}[g(\tau)|k, q]}.$$

Comparing to Eq. (25) and taking the weighted sum over k and q with weights $p(k|q)p_q$ yields a bound,

$$F_{1:N}^{\text{times|ch}}(\theta)/N \geq \sum_{k,q} \frac{\mathbb{E}[\partial_\theta g(\tau)|k, q]^2}{\text{Var}[g(\tau)|k, q]} p(k|q) p_q. \quad (102)$$

This is precisely the same as Eq. (101), except that the bound is on the conditional part of the Fisher information, alone. Because the Fisher information can be decomposed as $F = F^{\text{ch}} + F^{\text{times|ch}}$, we can thus improve the bound in Eq. (101) by adding the information contained in the observed channels,

$$F_{1:N}(\theta) \geq F_{1:N}^{\text{ch}}(\theta) + N \sum_{k,q} \frac{\mathbb{E}[\partial_\theta g(\tau)|k, q]^2}{\text{Var}[g(\tau)|k, q]} p(k|q) p_q, \quad (103)$$

The common form of Eq. (102) uses $g(\tau) = \tau - \mu_{kq}$ with $\mu_{kq} = \mathbb{E}[\tau|k, q]$. Similarly denoting the variance $\sigma_{kq}^2 = \text{Var}[\tau|k, q]$, a bound on Fisher information is obtained,

$$F_{1:N}^{\text{times|ch}}(\theta) \geq N \sum_{k,q} \frac{[\partial_\theta \mu_{kq}]^2}{\sigma_{kq}^2} p(k|q) p_q. \quad (104)$$

This bound roughly arises from considering the average delay between observations.

One may consider how the calculations change when rescaling $g(\tau)$:

$$\bar{g}(\tau) = h(\theta, k, q)g(\tau). \quad (105)$$

It remains the case that $\mathbb{E}[\bar{g}(\tau)|k, q] = h(\theta, k, q)\mathbb{E}[g(\tau)|k, q] = 0$. The variance is

$$\text{Var}[\bar{g}(\tau)|k, q] = h^2(\theta, k, q) \text{Var}[g(\tau)|k, q].$$

Then

$$\begin{aligned} \mathbb{E}[\partial_\theta \bar{g}(\tau)|k, q] &= h(\theta, k, q)\mathbb{E}[\partial_\theta g(\tau)|k, q] + (\partial_\theta h(\theta, k, q)) \mathbb{E}[g(\tau)|k, q] \\ &= h(\theta, k, q)\mathbb{E}[\partial_\theta g(\tau)|k, q]. \end{aligned}$$

The factor of $h^2(\theta, k, q)$ from this expectation value (after squaring) cancels the same factor from the variance. That is, rescaling $g(\tau)$ does not affect the bound on Fisher information. Considering a bound based on a weighted average instead of the average, as in Eq. (104), will then result in the same lower bound.

Though rescaling has no effect, the bound can be improved by considering the function $g(\tau)$ with higher order terms in τ . For example, if one considers quadratic functions in τ , a stronger bound is obtained in terms of the first four statistical moments [72–74].

F Monitoring operator for multi-parameter estimation

In this appendix, we present a generalization of the monitoring operator formalism to multi-parameter estimation. The parameters to be estimated are represented by the vector $\vec{\theta}$. The Fisher information matrix [47, 75] is defined as

$$\left[F(\vec{\theta}) \right]_{ij} = \mathbb{E} \left[\frac{1}{\text{Pr}(X)^2} \partial_{\theta_i} \text{Pr}(X) \partial_{\theta_j} \text{Pr}(X) \right]. \quad (106)$$

This expression can be rewritten in terms of the monitoring operator, introduced in Eq. (32), by defining parameter-specific operators $\xi_t^{(i)}$, $\xi_t^{(j)}$, with the important property

$$\text{Tr} \left[\xi_t^{(i)} \right] = \frac{1}{\text{Pr}(X)} \partial_{\theta_i} \text{Pr}(X). \quad (107)$$

The operators $\xi_t^{(i)}$ evolve in the same way as the operator ξ_t , but all the derivatives have to be taken with respect to θ_i , while the other parameters are assumed to be fixed. In this way, we finally get the expression for the Fisher information matrix

$$\left[F(\vec{\theta}) \right]_{ij} = \mathbb{E} \left[\text{Tr} \left[\xi_t^{(i)} \right] \text{Tr} \left[\xi_t^{(j)} \right] \right]. \quad (108)$$

In order to obtain the Fisher information matrix, one should then separately evolve monitoring operators $\xi_t^{(i)}$ for each component of the parameters vector, and combine them according to Eq. (108).

G Fisher information rate: average

In this appendix, we show the derivation of Eq. (51), yielding the average Fisher information rate for a generic measurement record. The Fisher information variation is given by

$$dF = F(X_{T+1}|X_{1:T}) = \mathbb{E}_{1:T} \left[\frac{1}{\text{Pr}(x_{T+1}|x_{1:T})} (\partial_{\theta} \text{Pr}(x_{T+1}|x_{1:T}))^2 \right], \quad (109)$$

where the expectation value is taken with respect to all the possible trajectories of T time steps, not including timestep $T + 1$. We now introduce the stochastic currents I_T^k , such that $\text{Pr}(k|x_{1:t}) = I_t^k dt$, for k labeling one of the $|\mathbb{M}|$ monitored jump channels:

$$dF = \mathbb{E} \left[\sum_{k=1}^{|\mathbb{M}|} \frac{1}{I_T^k dt} (\partial_{\theta} I_T^k dt)^2 + \frac{1}{1 - \sum_{k=1}^{|\mathbb{M}|} I_T^k dt} \left(\sum_{k=1}^{|\mathbb{M}|} \partial_{\theta} I_T^k dt \right)^2 \right]. \quad (110)$$

We note that the first term, which only contains single-current contributions, is of order dt in the $dt \rightarrow 0$ limit. On the contrary, the second term, which contains also some non-diagonal contributions mixing currents together, is of order dt^2 . To compute the rate, only terms of order dt should be taken into account; we obtain therefore

$$dF = \mathbb{E} \left[\sum_{k=1}^{|\mathbb{M}|} \frac{1}{I_T^k dt} (\partial_{\theta} I_T^k dt)^2 \right], \quad (111)$$

from which the Fisher information rate follows.

Algorithm 1: Gillespie evolution for the stochastic Fisher information

Input: Hamiltonian H_θ ; displaced Hamiltonians $H_{\theta+d\theta}$ and $H_{\theta-d\theta}$;
list of monitored jumps Kraus operators $[M_\theta]$;
displaced lists of monitored jumps Kraus operators $[M_{\theta+d\theta}]$ and $[M_{\theta-d\theta}]$;
derivation increment $d\theta$; initial state $|\psi_0\rangle$; final time t_f ; timestep dt ; number of trajectories;
Precompute $J = \sum_k M_k^\dagger M_k$;
Effective Hamiltonian $H_e = H - \frac{i}{2}J$; Displaced effective Hamiltonian $H_{e\pm} = H_{\theta\pm d\theta} - \frac{i}{2}J_\pm$;
Displaced jump operators $J_\pm = \sum_k M_{k,\theta\pm d\theta}^\dagger M_{k,\theta\pm d\theta}$;
Let $V[t]$ be the list of the no-jump evolution operators;
Let $Q_s[t]$ be the list of the non-state-dependent parts of the WTD;
Let $\dot{V}[t]$ be the list of the derivatives of the no-jump evolution operators with respect to θ ;
for all times $t < t_f$ **do**
 $V[t] = e^{-iH_e t}$;
 $Q_s[t] = V[t]^\dagger J V[t]$;
 $\dot{V}[t] = [\exp(-iH_{e+}t) - \exp(-iH_{e-}t)] / 2d\theta$;
end
Let \dot{M} be the list of the derivatives of the jump operators with respect to the parameter θ ;
Let $\Delta[t]$ be the list of all the derivatives of $M_k V[t]$ for all values of k ;
for all the jump operators in $[M]$ **do**
 $\dot{M}_k = (M_{k+} - M_{k-}) / 2d\theta$;
 Let Δ_k be the list of the derivatives of $M_k V[t]$;
 for $t < t_f$ **do**
 $\Delta_k[t] = \dot{M}_k V[t] + M_k \dot{V}[t]$;
 end
 $\Delta[k] = \Delta_k$;
end
for all trajectories **do**
 Fix initial state $|\psi\rangle = |\psi_0\rangle$, initial monitoring operator $\xi = 0$ and initial time $\tau = 0$;
 while $t < t_f$ **do**
 $\rho = |\psi\rangle\langle\psi|$;
 Let $W[t]$ be the list to contain the WTD;
 for every value Q in Q_s **do**
 $W[t] = \langle\psi|Q[t]|\psi\rangle$;
 end
 Sample a time T from the distribution $W[t]$;
 $\tau+ = T$;
 Update the state accordingly: $|\psi\rangle = V[T]|\psi\rangle$;
 Let $[p]$ be the list of the jump probabilities;
 for all the jump operators in $[M]$ **do**
 $p = \langle\psi|M_k^\dagger M_k|\psi\rangle$;
 end
 Sample a jump channel \bar{k} from the weights distribution $[p]$;
 Update and normalize the state $|\psi\rangle = M_{\bar{k}}|\psi\rangle / \|\psi\|$;
 Compute the evolution of the ξ operator.
 $\xi = \frac{1}{\|\psi\|^2} \left(M_{\bar{k}} V[T] \xi V[T]^\dagger M_{\bar{k}}^\dagger + \Delta[\bar{k}][T] \rho V[T]^\dagger M_{\bar{k}}^\dagger + M_{\bar{k}} V[T] \rho \Delta[\bar{k}][T]^\dagger \right)$;
 Compute the stochastic Fisher information $f = \text{Tr}[\xi]^2$;
 end
end
Output: $[(t_j, k_j, \rho_j, f_j)]$ vectors of measurement records, states and stochastic Fisher information values for all the trajectories, $V[t]$, $\dot{V}[t]$ often needed for further calculations, and the list of times at which V is known.
