Asymptotic inference in system identification for the atom maser

BY CĂTĂLIN CĂTANĂ, MERLIN VAN HORSSEN AND MĂDĂLIN GUȚĂ*

School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

System identification is closely related to control theory and plays an increasing role in quantum engineering. In the quantum set-up, system identification is usually equated to process tomography, i.e. estimating a channel by probing it repeatedly with different input states. However, for quantum dynamical systems such as quantum Markov processes, it is more natural to consider the estimation based on continuous measurements of the output, with a given input that may be stationary. We address this problem using asymptotic statistics tools, for the specific example of estimating the Rabi frequency of an atom maser. We compute the Fisher information of different measurement processes as well as the quantum Fisher information of the atom maser, and establish the local asymptotic normality of these statistical models. The statistical notions can be expressed in terms of spectral properties of certain deformed Markov generators, and the connection to large deviations is briefly discussed.

Keywords: atom maser; quantum Markov processes; system identification; Fisher information; asymptotic normality

1. Introduction

We are currently entering a new technological era [1] where quantum control is a becoming a key component of quantum engineering [2]. In the standard set-up of quantum filtering and control theory [3,4], the dynamics of the system and its environment, as well as the initial state of the system, are usually assumed to be known. In practice, however, these objects may depend on unknown parameters, and inaccurate models may compromise the control objective. Therefore, system identification [5], which lies at the intersection of control theory and statistics, is becoming an increasingly relevant topic for quantum engineering [6].

In this paper, we introduce probabilistic and statistical tools aimed at a better understanding of the measurement process, and at solving system identification problems in the set-up of quantum Markov processes. Although the mathematical techniques have a broader scope, we focus on the physically relevant model of the atom maser, which has been extensively investigated both theoretically [7,8] and experimentally [9,10] and found to exhibit a number of interesting dynamical phenomena. In the standard set-up of the atom maser, a beam of two-level atoms

*Author for correspondence (madalin.guta@nottingham.ac.uk).

One contribution of 15 to a Theo Murphy Meeting Issue ‘Principles and applications of quantum control engineering’.
prepared in the excited state passes through a cavity with which they are in
resonance, interact with the cavity field and are detected after exiting the cavity.
We consider the case where the atoms are measured in the standard basis, but
more general measurements can be analysed in the same framework.

The specific questions we want to address are how to estimate the strength
of the interaction between cavity and atoms, what is the accuracy of the
estimator and how it relates to the spectral properties of the Markov evolution.
In particular, we aim at investigating fundamental statistical concepts such as
Fisher information and asymptotic normality, in the context of multi-parameter
quantum system identification. These topics are well understood in the classical
context, and our aim is to adapt and extend such techniques to the quantum
set-up. In classical statistics, it is known that if we observe the first $n$ steps of
a Markov chain whose transition matrix depends on some unknown parameter
$\theta$, then $\theta$ can be estimated with an optimal asymptotic mean square error of
$\sqrt{nI(\theta)^{-1}}$, where $I(\theta)$ is the Fisher information (per sample) of the Markov chain.
Moreover, the error is asymptotically normal (Gaussian),

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I(\theta)^{-1}).$$

The key feature of our estimation problem is that the atom maser’s output
consists of atoms that are correlated with each other and with the cavity. Therefore, state and process tomography methods do not apply directly. In
particular, it is not clear what is the optimal measurement, what is the quantum
Fisher information of the output and how it compares with the Fisher information
of simple (counting) measurements. These questions were partly answered in
we extend the results to a continuous time set-up with an infinite-dimensional
system. For a better grasp of the statistical model, we consider several thought
and real experiments, and compute the Fisher information of the data collected
in these experiments. For example, we analyse the set-up where the cavity is
observed as it jumps between different Fock states when counting measurements
are performed on the output as well as in the temperature bath. We also consider
estimators that are based solely on the statistic given by the total number of
ground or excited state atoms detected in a time period. We show that the
quantum Fisher information of the closed system made up of atoms, a cavity and a
bath, depends strongly on the value of the interaction strength and is proportional
to the cavity mean photon number in the stationary state. Furthermore, we find
that monitoring all channels achieves the quantum Fisher information, while
excluding the detection of excited atoms leads to a drastic decrease of the Fisher
information in the neighbourhood of the first ‘transition point’. The asymptotic
regime relevant for statistics is that of central limit, i.e. moderate deviations
around the mean of order $n^{-1/2}$ as in (1.1). One of our main results is to establish
local asymptotic normality (LAN) for the atom counting process that implies the
central limit and provides the formula of the Fisher information. We also prove
a quantum version of this result showing that the quantum statistical model
of the atom maser can be approximated in a statistical sense by a displaced
coherent state model. The moderate-deviation regime analysed, as well as the
related regime of large deviations, are closely connected to the spectral properties
of certain deformations of the Lindblad operator. Some of these connections are
pointed out in this paper, but other questions such as the existence of dynamical phase transitions [12,13] and the quantum Perron–Frobenius theorem will be addressed elsewhere [14].

In §§2 and 3, we give brief overviews of the atom maser’s dynamics, and respectively of classical and quantum statistical concepts used in the paper. Section 4 contains the main results about Fisher information and asymptotic normality in different set-ups. We conclude with comments on future work.

2. The atom maser

The atom maser’s dynamics is based on the Jaynes–Cummings model of the atom–cavity Hamiltonian

$$H = H_{\text{free}} + H_{\text{int}} = \hbar \Omega a^* a + \hbar \omega \sigma^* \sigma - \hbar g(t) (\sigma a^* + a^* \sigma),$$  \hspace{1cm} (2.1)

where $a$ is the annihilation operator of the cavity mode, $\sigma$ is the lowering operator of the two-level atom, $\Omega$ and $\omega$ are the cavity frequency and the atom transition frequency that are assumed to be equal and $g$ is the coupling strength or Rabi frequency. In the standard experimental set-up, the atoms prepared in the excited state arrive as a Poisson process of a given intensity, and interact with the cavity for a fixed time. Additionally, the cavity is in contact with a thermal bath with mean photon number $n$. By coarse graining, the time evolution to ignore the very short time-scale changes in the cavity field, one arrives at the following master equation for the cavity state $\rho$:

$$\frac{d\rho}{dt} = \mathcal{L}(\rho),$$  \hspace{1cm} (2.2)

where $\mathcal{L}$ is the Lindblad generator,

$$\mathcal{L}(\rho) = \sum_{i=1}^{4} \left( L_i \rho L_i^* - \frac{1}{2} \{ L_i^* L_i, \rho \} \right);$$  \hspace{1cm} (2.3)

with operators

$$L_1 = \sqrt{N_{\text{ex}}} a^* \frac{\sin(\phi \sqrt{aa^*})}{\sqrt{aa^*}}, \quad L_2 = \sqrt{N_{\text{ex}}} \cos(\phi \sqrt{aa^*}),$$  \hspace{1cm} (2.4)

$$L_3 = \sqrt{n + 1} a \quad \text{and} \quad L_4 = \sqrt{n} a^*.$$  \hspace{1cm} (2.5)

Here, $N_{\text{ex}}$ is the effective pump rate (number of atoms per cavity lifetime), and the parameter $\phi$ (called the accumulated Rabi angle) is proportional to $g$. Later, we will consider that $\phi$ is an unknown parameter to be estimated.

The operators $L_i$ can be interpreted as jump operators for different measurement processes: detection of an output atom in the ground state or excited state ($L_1$ and $L_2$) and emission or absorption of a photon by the cavity ($L_3$ and $L_4$). In each case, the cavity makes a jump up or down on the ladder of Fock states. Because both the atom–cavity interaction (2.1) and the cavity–bath interaction leave the commutative algebra of number operators invariant, we can
Figure 1. The stationary state as a function of $\alpha := \phi \sqrt{N_{ex}}$, for $\nu = \sqrt{1.15}$, $N_{ex} = 100$. The white patches represent the photon number distribution and the black line is the expected photon number.

restrict our attention to this classical dynamical system, provided that the atoms are measured in the $\sigma_z$ basis. The cavity jumps are described by a birth–death (Markov) process with birth and death rates

$$q_{k,k+1} := N_{ex} \sin(\phi \sqrt{k+1})^2 + \nu(k+1), \quad k \geq 0,$$

and

$$q_{k,k-1} := (\nu + 1)k, \quad k \geq 1.$$

In §4, we will come back to the birth–death process, in the context of estimating $\phi$.

The Lindblad generator (2.3) has a unique stationary state (i.e. $\mathcal{L}(\rho_s) = 0$) that is diagonal in the Fock basis and has coefficients

$$\rho_s(n) = \rho_s(0) \prod_{k=1}^{n} \left( \frac{\nu}{\nu + 1} + \frac{N_{ex}}{\nu + 1} \frac{\sin^2(\phi \sqrt{k})}{k} \right).$$

This means that the cavity evolution is ergodic in the sense that in the long run, any initial state converges to the stationary state. In figure 1, we illustrate the dependence on $\alpha := \phi \sqrt{N_{ex}}$ of the stationary state. The notable features are the sharp change of the mean photon number at $\alpha \approx 1, 2\pi, 4\pi, \ldots$, and the fact that the stationary state is bistable at these points with the exception of the first one. The bistability is accompanied by a significant narrowing of the spectral gap of $\mathcal{L}$. This behaviour is typical of a metastable Markov process where several ‘phases’ are very weakly coupled to each other, and the process spends long periods in one phase before abruptly moving to another.

An alternative perspective to these phenomena is offered by the counting process of the outgoing atoms. Because the rate at which an atom exchanges an excitation with the cavity depends on the cavity state, the counting process carries information about the cavity dynamics and, in particular, about the
interaction parameter $\phi$. The recent papers [12,13] propose to analyse the stationary dynamics of the atom maser using the theories of large deviations and dynamical phase transitions. Instead of looking at the ‘phases’ of the stationary cavity state, the idea is to investigate the long time properties of measurement trajectories and identify their dynamical phases, i.e. ensembles of trajectories that have markedly different count rates in the long run, or ‘activities’. The large-deviation approach raises important questions related to the existence of dynamical phase transitions, which can be formulated in terms of the spectral properties of the ‘modified’ generator $L$, defined in §4, and in terms of the Perron–Frobenius theorem for infinite-dimensional quantum Markov processes [14]. In this work, we concentrate on the closely related, but distinct, regime of moderate deviations characterized by the central limit theorem, which is more relevant for statistical inference problems. For later purposes, we introduce a unitary dilation of the master evolution that is defined by the unitary solution of the following quantum stochastic differential equation:

$$dU(t) = \sum_{i=1}^{4} \left( L_i dA^*_{i,t} - L^*_i dA_{i,t} - \frac{1}{2} L^*_i L_i dt \right) U(t). \quad (2.8)$$

The pairs $(dA_{i,t}, dA^*_{i,t})$ represent the increments of the creation and annihilation operators of four independent bosonic input channels, which couple with the cavity through the operators $L_i$. The master evolution can be recovered as usual by tracing out the bosonic environment, which is initially in the vacuum state,

$$e^{tL}(\rho) = \text{Tr}_{\text{env}}(U(t)(\rho \otimes \mathcal{O})(\mathcal{O}) U(t)^*).$$

If $d\Gamma_{i,t}$ denote the increments of the four number operators of the input channels, then the counting operators of the output are

$$A_{i,t} := U(t)^*(1 \otimes \Gamma_{i,t}) U(t), \quad (2.9)$$

which provide the statistics of counting atoms in the ground state, excited state, emitted and absorbed photons.

3. Brief overview of classical and quantum statistics notions

For convenience, we recall here some basic notions of classical and quantum parametric statistics, which will be useful for interpreting the results of the next section.

(a) Estimation for independent identically distributed variables

A typical statistical problem is to estimate an unknown parameter $\theta = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$, given the data consisting of independent, identically distributed (i.i.d.) samples $X_1, \ldots, X_n$ from a distribution $\mathbb{P}_\theta$, which depends on $\theta$.

An estimator $\hat{\theta}_n := \hat{\theta}_n(X_1, \ldots, X_n)$ is called unbiased if $\mathbb{E}_\theta(\hat{\theta}_n) = \theta$ for all $\theta$. The Cramér–Rao inequality gives a lower bound to the covariance matrix and mean square error of any unbiased estimator

$$\text{Cov}(\hat{\theta}_n) = \mathbb{E}_\theta[(\hat{\theta}_n - \theta)^T(\hat{\theta}_n - \theta)] \geq \frac{1}{n} I(\theta)^{-1} \quad (3.1)$$
and
\[ \mathbb{E}_\theta[\|\hat{\theta}_n - \theta\|^2] \geq \frac{1}{n} \text{Tr}(I(\theta)^{-1}). \] (3.2)

The \( k \times k \) positive matrix \( I(\theta) \) is called the Fisher information matrix and can be computed in terms of the log-likelihood functions \( \ell_\theta := \log p_\theta \), where \( p_\theta \) is the probability density of \( \mathbb{P}_\theta \) with respect to some reference measure \( \mu \),

\[ I(\theta)_{i,j} = \mathbb{E}_\theta \left( \frac{\partial \ell_\theta}{\partial \theta_i} \frac{\partial \ell_\theta}{\partial \theta_j} \right) = \int p_\theta(x) \frac{\partial \log p_\theta}{\partial \theta_i} \frac{\partial \log p_\theta}{\partial \theta_j} \mu(dx). \]

The Cramér–Rao bound is, in general, not achievable for a given \( n \). However, what makes the Fisher information important is the fact that the bound is asymptotically achievable. Furthermore, asymptotically optimal estimators (or efficient estimators) are asymptotically normal in the sense that

\[ \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{L} N(0, I(\theta)^{-1}), \] (3.3)

where the right-hand side is a centred \( k \)-variate Gaussian distribution with covariance \( I(\theta)^{-1} \), and the convergence is in law for \( n \to \infty \). Under certain regularity conditions, the maximum-likelihood estimator

\[ \hat{\theta}_n := \arg \max_{\theta} \prod_i p_\theta(X_i) \]

is efficient. The asymptotic normality of efficient estimators can be seen as a consequence of the more fundamental theory of LAN, which states that the i.i.d. statistical model \( \mathbb{P}^n_\theta \) can be ‘linearized’ in a local neighbourhood of any point \( \theta_0 \) and approximated by a Gaussian model. Because the uncertainty in \( \theta \) is of the order \( n^{-1/2} \), we write \( \theta := \theta_0 + u/\sqrt{n} \), where \( u \) is a local parameter that is considered unknown, whereas \( \theta_0 \) is fixed and known. LAN can be expressed as the (local) convergence of statistical models [15]

\[ \{\mathbb{P}^n_{\theta_0 + u/\sqrt{n}} : u \in \mathbb{R}\} \to \{N(u, I(\theta_0)^{-1}) : u \in \mathbb{R}\}, \]

where the limit is approached as \( n \to \infty \), and consists of a single sample from the normal distribution with unknown mean \( u \) and known variance \( I(\theta_0)^{-1} \). In §4d,e, we will prove two versions of LAN, one for quantum states and one for a classical counting process.

(b) Quantum estimation with identical copies

Consider now the problem of estimating \( \theta \in \mathbb{R}^k \), given \( n \) identical and independent copies of a quantum state \( \rho_\theta \). The quantum Cramér–Rao bound [16–18] says that for any measurement on \( \rho_\theta^{\otimes n} \) (including joint ones) and any unbiased estimator \( \hat{\theta}_n \) constructed from the outcome of this measurement, the lower bound (3.1) holds with \( I(\theta) \) replaced by the quantum Fisher information matrix

\[ F(\theta)_{i,j} = \text{Tr}(\rho_\theta D_{\theta,i} \circ D_{\theta,j}), \]

where \( X \circ Y := \{X, Y\}/2 \) and \( D_{\theta,i} \) are the self-adjoint operators defined by

\[ \frac{\partial \rho_\theta}{\partial \theta_i} = D_{\theta,i} \circ \rho_\theta. \]
When $\theta$ is one dimensional, the quantum Cramér–Rao bound is asymptotically achievable by the following two-step adaptive procedure. First, a small proportion $\tilde{n} \ll n$ of the systems is measured in a ‘standard’ way, and a rough estimator $\theta_0$ is constructed; in the second step, one measures $D_{\theta_0}$ separately on each system to obtain results $D_1, \ldots, D_{n-\tilde{n}}$ and defines the efficient estimator

$$\hat{\theta}_n = \theta_0 + \frac{1}{(n - \tilde{n}) F(\theta_0)} \sum_i D_i.$$ 

However, for multi-dimensional parameters, the quantum Cramér–Rao bound is not achievable even asymptotically, due to the fact that the operators $D_{\theta,i}$ may not commute with each other and cannot be measured simultaneously. Moreover, unlike the classical case, there are several Cramér–Rao bounds based on different notions of ‘Fisher information’ [19]. In this case, it is more meaningful to search for asymptotically optimal estimators in the sense of optimizing the risk given by the mean square error (3.2). In Hayashi & Matsumoto [20], it has been shown that for qubits, the asymptotically optimal risk is given by the so-called Holevo bound [17]. For arbitrary dimensions, the achievability of the Holevo bound can be deduced from the theory of quantum LAN developed in Khan & Guţă [21], and a discussion on this can be found in Guta & Kahn [22].

(c) **Fisher information for classical Markov processes**

Often, the data we need to investigate are not a sequence of i.i.d. variables but a stochastic process, e.g. a Markov process. A theory of efficient estimators and (local) asymptotic normality can be developed along the lines of the i.i.d. set-up, provided that the process is ergodic. We will describe the basic ingredients of a continuous time Markov process and write its Fisher information.

Let $I = \{1, \ldots, m\}$ be a set of states, and let $Q = [q_{ij}]$ be an $m \times m$ matrix of transition rates, with $q_{ij} \geq 0$ for $i \neq j$ and diagonal elements $q_{ii} := -\sum_{j \neq i} q_{ij}$. The rate matrix is the generator of a continuous time Markov process, and the associated semigroup of transition operators is

$$P(t) = \exp(tQ).$$

A continuous time stochastic process $(X_t)_{t \geq 0}$ with state space $I$ is a Markov process with transition semigroup $P(t)$ if

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_0 = i_0, \ldots, X_n = i_n) = p(t_{n+1} - t_n)_{i_n i_{n+1}},$$

for all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq \cdots \leq t_{n+1}$, and all states $i_0, \ldots, i_{n+1}$ where $p(t)_{ij}$ are the matrix elements of $P(t)$.

Let us denote by $J_0, J_1, \ldots$ the times when the process jumps from one state to another, so that $J_0 = 0$ and $J_{n+1} = \inf\{t > J_n : X_t \neq X_{J_n}\}$. The time between two jumps is called ‘holding time’ and is defined by $S_i = J_i - J_{i-1}$.

A probability distribution $\pi = (\pi_1, \ldots, \pi_m)$ over $I$ is stationary for the Markov process $(X_t)_{t \geq 0}$ if it satisfies $\pi Q = 0$ or equivalently $\pi P(t) = \pi$ at all $t$. If the transition matrix is irreducible, then this distribution is unique and the process is called ergodic, in which case, any initial distribution $\mu$ converges to the stationary distribution

$$\lim_{t \to \infty} \mu P(t) = \pi.$$
Suppose now that we observe the ergodic Markov process $X_t$ for $t \in [0, T]$, and that the rate matrix depends smoothly on some unknown parameter $\theta$ (which for simplicity we consider one dimensional), so that $q_{ij} = q^\theta_{ij}$. The asymptotic theory says that ‘good’ estimators like maximum likelihood (under some regularity conditions) are asymptotically normal in the sense of (3.3), with Fisher information given by

$$I(\theta) := \sum_{i \neq j} \pi_i^\theta q_{ij}^\theta (V_{ij})^2,$$  \hspace{1cm} (3.4)

where

$$V_{ij} := \frac{d}{d\theta} \log q_{ij}^\theta$$

and $\pi^\theta$ is the stationary distribution at $\theta$.

4. Fisher information for the atom maser

In this section, we return to the atom maser and investigate the problem of estimating the interaction parameter $\phi$, based on outcomes of measurements performed on the outgoing atoms. State and process tomography are key enabling components in quantum engineering, and have become the focus of research at the intersection of quantum information theory and statistics. Our contribution is to go beyond the usual set-up of repeated measurements of identically prepared systems, or that of process tomography, and look at estimation in the quantum Markov set-up. The first step in this direction was made in Guță [11], which deals with asymptotics of system identification in a discrete time setting with finite-dimensional systems. Here, we extend these ideas to the atom maser, including the effect of the thermal bath. In the next subsections, we consider several thought experiments in which counting measurements are performed in the output bosonic channels determined by the unitary coupling (2.8). While some of these scenarios are not meant to have a practical relevance, the point is to analyse and compare the amount of Fisher information carried by the various stochastic processes associated to the atom maser, as illustrated in figures 2 and 3.

(a) Observing the cavity

Consider first the scenario where all four channels are monitored by means of counting measurements. As already discussed in §2, the conditional evolution of the cavity is described by the birth and death process consisting of jumps up and down the Fock ladder, with rates (2.6). Note that when an atom is detected in the excited state, the cavity state remains unchanged; so the corresponding rate $N_{ex} \cos(\phi \sqrt{i + 1})^2$ does not appear in the birth and death rates. Later, we will see that these atoms do carry Fisher information about the interaction parameter, even if they do not modify the state of the cavity.

Because the cavity dynamics is Markovian, we can use (3.4) and the expression of the stationary state (2.7) to compute the Fisher information of the stochastic
process determined by the cavity state

\[ I_{\text{cav}}(\phi) = \sum_{i=0}^{\infty} \rho_{i}^{\phi}(i) \left( \frac{q_{i+1}^{\phi}}{q_{i}^{\phi}} \right)^{2} q_{i+1}^{\phi}. \]

We stress that this information refers to an observer who alone has access to the cavity state, and cannot infer whether a jump up is due to exchanging an excitation with an atom or absorbing a photon from the bath. Moreover, the observer does not get any information about the atoms which pass through the cavity without exchanging an excitation.

The function \( I_{\text{cav}}(\phi) \) is plotted as the dashed-dotted line in figure 2.

(b) Observing the cavity and discriminating between jumps

In the next step, we assume that besides monitoring the cavity, we are also able to discriminate between the two events producing a jump up, which in effect is equivalent to monitoring the emission and absorption from the bath and the atoms exiting in the ground state, but not those in the excited state.

But how do we model probabilistically the additional piece of information? Let us fix a given trajectory of the cavity that has jumps up at times \( t_{1}, \ldots, t_{l} \) from the Fock states with photon numbers \( k_{1}, \ldots, k_{l} \). Conditional on this trajectory,
the events ‘jump at \( t_i \) is due to atom’ and its complement ‘jump at \( t_i \) is due to bath’ have probabilities

\[
p_a^i = \frac{r_{a}^{k_i}}{r_{a}^{k_i} + r_{b}^{k_i}} \quad \text{and} \quad p_b^i = \frac{r_{b}^{k_i}}{r_{a}^{k_i} + r_{b}^{k_i}},
\]

where \( r_{a}^{k} = N_{\text{ex}} \sin(\phi \sqrt{k + 1})^2 \) and \( r_{b}^{k} = \nu (k + 1) \) are the rates for atoms and bath jumps. This means that we can model the process by independently tossing a coin with probabilities \( p_a^i \) and \( p_b^i \) at each time \( t_i \). For each toss, the additional Fisher information is

\[
I_{t_i}(\phi) = \left( \frac{dp_a^i}{d\phi} \right)^2 \frac{1}{p_a^i (1 - p_a^i)},
\]

and the information for the whole trajectory is obtained by summing over \( i \). The (asymptotic) Fisher information of the process is obtained by taking the time and stochastic averaging over trajectories, in the large time limit. Because, in the long run, the system is in the stationary state, the average number of \( k \to k + 1 \) jumps per unit of time is \( \rho^\phi_a(k) q^\phi_{k,k+1} \); so the additional Fisher information provided by the jumps is

\[
I_{\text{up}}(\phi) = \sum_{k=0}^{\infty} \rho^\phi_a(k) q^\phi_{k,k+1} \left( \frac{dp_a^k}{d\phi} \right)^2 \frac{1}{p_a^k (1 - p_a^k)}.
\]

Therefore, the Fisher information gained by following the cavity and discriminating between jumps is

\[
I_{\text{cav+up}}(\phi) = I_{\text{cav}}(\phi) + I_{\text{up}}(\phi),
\]

which is plotted as the dashed line in figure 2.

(c) Observing all counting processes

The next step is to incorporate the information contained in the detection of excited atoms, to obtain the full classical Fisher information of all four counting measurements. We will consider again a fixed cavity trajectory and compute the additional (conditional) Fisher information provided by the counts of excited atoms. During each holding time period \( S_i = t_{i+1} - t_i \), the cavity is in the state \( k_i \), and the excited atoms are described by a Poisson process with rate

\[
r_{e}^{k_i} := N_{\text{ex}} \cos(\phi \sqrt{k_i + 1})^2.
\]

Moreover, the Poisson processes for different holding times are independent; so the conditional Fisher information is the sum of information for each Poisson process. Now, for a Poisson process, the total number of counts in a time interval is a sufficient statistic, and the times of arrival can be neglected. Thus, we need only to compute the Fisher information of a Poisson distribution with mean \( \lambda_i := r_{e}^{k_i} S_i \) and add up over \( i \). A short calculation shows that this is equal to

\[
J_i = \left( \frac{d\lambda_i}{d\phi} \right)^2 \frac{1}{\lambda_i} = \left( \frac{dr_{e}^{k_i}}{d\phi} \right)^2 \frac{S_i}{r_{e}^{k_i}}.
\]

As before, it remains to add over all holding times and take averages over trajectories, and average over a long period of time. This amounts to replacing
$S_t$ by the stationary distribution $\rho_s(k_i)$, which is the average time in the state $k_i$ per unit of time. The Fisher information is

$$I_{\text{exc}} = \sum_{k=0}^\infty \rho_s(k) \frac{1}{r^k_e} \left( \frac{dr^k_e}{d\phi} \right)^2.$$ 

We can now write down the total classical Fisher information of the four counting processes

$$I_{\text{tot}} := I_{\text{exc}} + I_{\text{up}} + I_{\text{cav}} = 4N_{\text{ex}} \sum_{k=0}^\infty \rho_s(k)(k+1),$$

where the last equality follows from a simple calculation based on the explicit expressions of the three terms. The total information $I_{\text{tot}}$ is plotted as the solid line in figure 2.

The last expression is surprisingly simple, and as we will see in §4d, it is equal to the quantum Fisher information of the atom maser output process, which is the maximum information extracted by any measurement!

\[(d)\] The quantum Fisher information of the atom maser

Up to this point, we considered the problem of estimating $\phi$ in several scenarios involving counting processes. We will now investigate the more general problem of estimating $\phi$ when arbitrary measurements are allowed. As discussed in §3b, the key statistical notions of Cramér–Rao bounds, Fisher information and asymptotic normality can be extended to i.i.d. quantum statistical models, and can be used to find asymptotically optimal measurement strategies for parameter estimation problems. In Guță [11], these notions were extended to the non-i.i.d. framework of a quantum Markov chain with finite-dimensional systems. Here, we extend these results further to the atom maser, which is a continuous time Markov process with an infinite-dimensional system. We refer to Guță [11] for more details on the physical and statistical interpretation of the results.

Let $|\chi\rangle$ be the initial state of the cavity and $|\Omega\rangle$ the joint vacuum state of the bosonic fields. The joint (pure) state of the cavity and the four Bosonic channels at time $t$ is

$$|\psi(\phi(t))\rangle = U_\phi(t) (|\chi\rangle \otimes |\Omega\rangle),$$

where $U_\phi(t)$ is the unitary solution of the quantum stochastic differential equation (2.8). We emphasize that both the unitary and the state depend on the parameter $\phi$ and we would like to know what is the ultimate precision limit for the estimation of $\phi$ assuming that arbitrary measurements are available.

As argued in §3a, for asymptotics it suffices to understand the statistical model in a local neighbourhood of a given point, whose size is of the order of the statistical uncertainty, in this case $t^{-1/2}$. For this, we write $\phi = \phi_0 + u/\sqrt{t}$ and focus on the structure of the quantum statistical model with parameter $u \in \mathbb{R}$,

$$|\psi(u, t)\rangle := |\psi_{\phi_0 + u/\sqrt{t}}(t)\rangle.$$

Our main result is to show that this quantum model is asymptotically Gaussian, in the sense that this family of vectors converges to a family of coherent states of a continuous variable system, similar to results obtained in recent studies [21, 23–25]

Phil. Trans. R. Soc. A (2012)
Asymptotic inference for the atom maser

for identical copies of quantum states, and in Guţă [11] for quantum Markov chains. More precisely,

$$\lim_{t \to \infty} \langle \psi(u, t) | \psi(v, t) \rangle = \langle \sqrt{2F} v | \sqrt{2F} u \rangle = e^{-\frac{(u-v)^2}{8F}}, \quad (4.1)$$

where $|\sqrt{2F} u \rangle$ denotes a coherent state of a one-mode continuous variable system, with displacement $\sqrt{2F} u$ along one axis, and $F = F(\phi_0)$ is a constant that plays the role of quantum Fisher information (per unit of time). The meaning of this result is that for large times, the state of the atom maser and environment is approximately Gaussian when seen from the perspective of parameter estimation, and by performing an appropriate measurement, we can extract the maximum amount of information $F$. At the end of the following calculation, we will find that $F = I_{\text{tot}}$; so the counting measurement is in fact optimal! Recall however that the counting measurement involves the detection of emitted and absorbed photons that is experimentally unrealistic. However, the result is relevant, as it puts an upper bound on any Fisher information that can be extracted by measurements on the output. To prove (4.1), we express the inner product in terms of a (non completely positive) semigroup on the cavity space, by tracing over the atoms and bath,

$$\langle \psi(u, t) | \psi(v, t) \rangle = \langle \phi | e^{tL_{u,v}}(1) | \phi \rangle, \quad (4.2)$$

where the generator $L_{u,v}$ is

$$L_{u,v}(X) = \sum_{i=1}^{4} \left( L_{i}^{u,v} X L_{i}^{u} - \frac{1}{2} L_{i}^{u,v} X L_{i}^{u,v} - \frac{1}{2} X L_{i}^{u,v} L_{i}^{u,v} \right),$$

and $L_{i}^{u} = L_{i}(\phi_0 + u/\sqrt{t})$ are the operators appearing in the Lindblad generator (2.3), where we emphasized the dependence on the local parameter. The proof of (4.1) uses a second-order perturbation result of Davies [26], which will be discussed in more detail in §4e. Here, we give the final result that says that the quantum Fisher information is proportional to the mean energy of the cavity in the stationary state, and is equal to the classical Fisher information $I_{\text{tot}}$ for the counting measurement

$$F = 4N_{\text{ex}} \sum_{k=0}^{\infty} \rho_{s}(k)(k + 1).$$

(e) Counting ground or excited state atoms

We now consider the scenario in which the estimation is based on the total number of ground-state atoms $A_{1,t}$ defined in (2.9), ignoring their arrival times. A similar argument can be applied to the excited state atoms. The generating function of $A_t$ can be computed from the unitary dilation (2.8), which gives

$$\mathbb{E}(\exp(sA_{1,t})) = \text{Tr}(\rho_0 e^{sL_s}(1)), \quad (4.3)$$

where $\rho_0$ is the initial state of the cavity and $L_s$ is the modified generator

$$L_s(\rho) = e^{sL_1} \rho L_1^{*} - \frac{1}{2} \{L_1^{*} L_1, \rho\} + \sum_{j \neq 1} \left( L_j \rho L_j^{*} - \frac{1}{2} \{L_j^{*} L_j, \rho\} \right). \quad (4.4)$$
Note that $\mathcal{L}$ is the generator of a completely positive but not trace preserving semigroup. We will analyse the moderate deviations of $A_{1,t}$ and show that it satisfies the central limit theorem. In what concerns the estimation of $\phi$, we find an explicit expression of the Fisher information and establish asymptotic normality. The latter means that

$$\hat{A}_{1,t} := \frac{1}{\sqrt{t}} (A_{1,t} - E_{\phi_0}(A_{1,t})) \xrightarrow{\mathcal{L}} N(\mu u, V), \quad (4.5)$$

where the convergence holds as $t \to \infty$, with a fixed local parameter, i.e $\phi = \phi_0 + u/\sqrt{t}$. In particular, for $u = 0$, we recover the central limit theorem for $A_{1,t}$. From (4.5), we find that the estimator

$$\hat{\phi}_t := \phi_0 + \frac{1}{\sqrt{t}} \frac{\hat{A}_{1,t}}{\mu}$$

is efficient (as well as the maximum-likelihood estimator), in the sense that its (rescaled) asymptotic variance $t \text{Var}(\hat{\phi}_t)$ is equal to the inverse of the Fisher information of the total counts of ground-state atoms

$$I_{\text{gr}} = \frac{\mu^2}{V}.$$

In the rest of the section, we describe the main ideas involved in proving (4.5) and give the expressions of $\mu$ and $V$. We first rewrite (4.5) in terms of the moment generating functions

$$\varphi(s, t) := \mathbb{E} \left[ \exp \left( i \frac{s}{\sqrt{t}} (A_{1,t} - E_{\phi_0}(A_{1,t})) \right) \right] \to \exp \left( i \mu u s - \frac{1}{2} s^2 V \right). \quad (4.6)$$

Using (4.4) and (4.3) with $s$ replaced by $s/\sqrt{t}$, the left-hand side can be written as

$$\varphi(s, t) = \text{Tr}[\rho_0 e^{t \mathcal{L}(s/\sqrt{t}, u/\sqrt{t})(1)}], \quad \text{with} \quad \mathcal{L} \left( \frac{s}{\sqrt{t}}, \frac{u}{\sqrt{t}} \right) = \mathcal{L}_{s/\sqrt{t}} - \frac{1}{\sqrt{t}} E_{\phi_0} \left( \frac{A_{1,t}}{t} \right),$$

where $\rho_0$ is the initial state of the cavity. The generator can be expanded in $t^{-1/2}$

$$\mathcal{L} \left( \frac{s}{\sqrt{t}}, \frac{u}{\sqrt{t}} \right) = \mathcal{L}^{(0)} + \frac{1}{\sqrt{t}} \mathcal{L}^{(1)} + \frac{1}{t} \mathcal{L}^{(2)} + O(t^{-3/2}),$$

and by applying the second-order perturbation theorem 5.13 of Davies [26], we get

$$\lim_{t \to \infty} \varphi(s, t) = \exp(\text{Tr}(\rho_0 \mathcal{L}^{(2)}(1)) - \text{Tr}(\rho_0 \mathcal{L}^{(1)} \circ \tilde{\mathcal{L}} \circ \mathcal{L}^{(1)}(1))),$$

where $\tilde{\mathcal{L}}$ is effectively the inverse of the restriction of $\mathcal{L}^{(0)}$ to the subspace of operators $X$ such that $\text{Tr}(\rho_0 X) = 0$, which contains $\mathcal{L}^{(1)}(1)$. From the power expansion, it can be seen that the the expression inside the last exponential

*Phil. Trans. R. Soc. A* (2012)
is quadratic in $u, s$, which provides the formulae for $\mu$ and $V$ in (4.6). The method outlined earlier is very general and can be applied to virtually any ergodic quantum Markov process. However, in numerical computations, we found that the fact $L^{(0)}$ has a small spectral gap for certain values of $\phi_0$ may pose some difficulties in computing the inverse $\tilde{L}$. An alternative method that we do not detail here is based on large-deviation theory and shows that

$$
\mu = \left. \frac{d r(s)}{d s} \right|_{s=0} \quad \text{and} \quad V = \left. \frac{d^2 r(s)}{d s^2} \right|_{s=0},
$$

where $r(s)$ is the dominant eigenvalue of $L_s$. Moreover, the coefficient $\mu$ can be computed in a more direct way as $\mu = d \text{Tr}(\rho^\phi N)/d \phi$ because for large times

$$
E_\phi \left( \frac{A_{1,t}}{t} \right) = \text{Tr}(\rho^\phi N) - v = \sum_{k=0}^\infty \rho^\phi(k)k - v,
$$

which follows from an energy-conservation argument in the stationary state.

A similar argument can be made for the total counts of the excited state atoms. The Fisher information for the ground and excited state atoms is represented in figure 3. The Fisher information for the two counting processes taken together is represented by the solid line. We note that the counts Fisher informations are comparable to those of the previous scenarios (figure 2) in the region $0 \leq \alpha \leq 4$, but significantly smaller in the bistability regions. Also, they are equal to zero at $\phi \approx 0.16$, owing to the fact that the derivative with $\phi$ of the mean atom number is zero at this point.
We have investigated the problem of estimating the Rabi frequency of the atom maser in the framework of asymptotic statistics. The Fisher information of several classical counting processes was computed, together with the quantum Fisher information, which is the upper bound of the classical information obtained from an arbitrary measurement. The latter was found to be equal to $4N_{ex}(N+1)s$, and is attained by the joint counting process of ground and excited atoms plus emitted and absorbed photons. However, in the region of the first transition point, we find that the Fisher information for both ground and excited total atom counts are equal to zero, while the quantum Fisher information is maximum. Even counting photons plus ground-state atoms while ignoring the excited atoms, does not give a significant amount of information. It would be interesting to see whether estimation precision at this point can be improved by taking into account the full atom count trajectories. Although maximum likelihood can be applied to these processes, perhaps in conjunction with Bayesian estimation and state filtering methods, this may be rather expensive in terms of computational time. An alternative is to use other estimation methods that are not likelihood based, e.g. approximate Bayesian computation methods. Another future direction is to explore the relation between the moderate-deviation regime (which we have analysed here) and the large-deviation regime (which is relevant for the study of dynamical phase transition) [12,13]. Ultimately, the goal is to design measurements that optimize the statistical performance of the estimation, in the spirit of Wiseman’s adaptive phase estimation protocol [27] and to explore the connections with control theory, e.g. in the frame of adaptive control. Two papers detailing the proofs of the asymptotic normality results in a general Markov set-up and the large-deviations perspective [14] are in preparation.

M.G. acknowledges the support of the EPSRC fellowship EP/E052290/1.

References


