Bayesian quantum frequency estimation in presence of collective dephasing

Katarzyna Macieszczak¹,², Martin Fraas³ and Rafał Demkowicz-Dobrzański²
¹ University of Nottingham, School of Mathematical Sciences, University Park, NG7 2RD Nottingham, UK
² Faculty of Physics, University of Warsaw, ul. Hoża 69, PL-00-681 Warszawa, Poland
³ Theoretische Physik, ETH Zürich, 8093 Zürich, Switzerland
E-mail: demko@fuw.edu.pl

Received 18 June 2014, revised 16 September 2014
Accepted for publication 17 September 2014
Published 31 October 2014

New Journal of Physics 16 (2014) 113002
doi:10.1088/1367-2630/16/11/113002

Abstract

We advocate a Bayesian approach to optimal quantum frequency estimation—an important issue for future quantum enhanced atomic clock operation. The approach provides a clear insight into the interplay between decoherence and the extent of prior knowledge in determining the optimal interrogation times and optimal estimation strategies. We propose a general framework capable of describing local oscillator noise as well as additional collective atomic dephasing effects. For a Gaussian noise, the average Bayesian cost can be expressed using the quantum Fisher information. Thus we establish a direct link between the two, often competing, approaches to quantum estimation theory.

Keywords: quantum metrology, frequency estimation, atomic clocks

1. Introduction

Modern atomic clocks allow for time keeping with instability better than $10^{-15}$ and are reaching towards new applications in geodesy [1] and tests of fundamental physics [2]. As of 2013 the best atomic clock has an instability of $10^{-18}$ after seven hours of averaging [3]. Crystal oscillators or stabilized lasers have excellent short-time frequency stability, but their frequency
ω tends to drift due to variations of temperature or stress. They need to be locked to a narrow atomic transition $\omega_0$ between two atomic levels $\mid 0 \rangle$, $\mid 1 \rangle$ in order to guarantee long-time stability (see figure 1). The resulting stability is limited by local oscillator (LO) noise and the atomic signal to noise ratio. A single estimation strategy has to be used periodically to determine the frequency offset $\omega(t) = \omega_{LO}(t) - \omega_0$. Knowledge of the LO noise and the value of $\omega$ from the previous feedback cycle provide a natural prior, making Bayesian analysis well suited for a study of the optimal estimation protocol.

In a typical Ramsey interferometric scheme, $N$ atoms interact with an external electromagnetic field evolving at frequency $\omega$, which for now is a time independent parameter. First, the atoms experience a $\pi/2$-pulse, which transforms the ground state $\mid 0 \rangle$ of each one into the $(\mid 0 \rangle + \mid 1 \rangle)/\sqrt{2}$ superposition. After evolving freely for a time $t$, they experience another $\pi/2$-pulse, subsequently being subjected to a measurement determining the number $n_1$ of atoms that made a transition to the excited state $\mid 1 \rangle$. In the case of perfect synchronization $\omega_{LO} = \omega_0$, all atoms end up in the $\mid 1 \rangle$ state, while in the presence of detuning, the average ratio reads $\langle n_1 \rangle/N = \cos^2[\omega t/2]$ [4]. The scheme is identical regardless of whether the $\omega_0$ transition is microwave (Cs-fountain clocks) or optical (ion traps, optical lattices). The procedure estimates the frequency difference $\omega$ on the basis of the number of excited atoms detected and performs a feedback on $\omega_{LO}$, locking it to the atomic frequency $\omega_0$. Mathematically, the scheme is equivalent to an $N$-photon Mach–Zehnder interferometric experiment with a relative optical phase delay $\phi = \omega t$ between the arms, which play the role of the two atomic levels [5]. Just as in optical interferometry, fluctuation in the number of detected atoms gives rise to the shot noise $1/\sqrt{N}$ scaling of the frequency estimation precision. Shot noise scaling is a direct consequence of lack of correlations among the atoms. If the atoms were prepared in a correlated quantum state, such as a spin-squeezed or a Greenberger–Horne–Zeilinger (GHZ) state, the $1/\sqrt{N}$ limit could be beaten and the $1/N$ Heisenberg bound approached, at least in idealized decoherence free scenarios [5–10]. While quantum enhanced sensing ideas have proved useful in a number of practical applications [11–14], it has been observed that in the presence of losses [15–20],

Figure 1. (a) Basic scheme of atomic clock operation. An LO coupled to an atomic reference generates an error signal on the frequency difference $\omega = \omega_{LO} - \omega_0$, which is used to keep the LO locked to the atomic transition. (b) Prior distribution of $\omega$ spreads while atoms are being interrogated due to LO instability. Estimation and a feedback correction procedure help to keep the $\omega$ distribution narrow.
dephasing [21–23], or other decoherence processes [24, 25] the maximum achievable quantum enhancement is limited. The ultimate goal of this type of research is to study the performance of optimal strategies in which probe states, measurements, and inferring strategies are optimized in a given sensing protocol [26, 27]. The resulting optimal sensing precision may then be regarded as fundamental, i.e., imposed by the laws of nature. Powerful tools are available for the effective identification of these limits [28–30]. A number of papers have studied the effectiveness of quantum strategies in frequency estimation [21, 23, 31, 32] or, more specifically, in atomic clock performance [33–37]. These approaches, however, lacked generality. Some ignored the role of prior frequency distribution or the presence of decoherence. Others studied a particular estimation scheme. The aim of this paper is to propose a simple, general, and effective Bayesian approach to the study of general quantum enhancement schemes in frequency estimation. We show that finding the optimal estimation scheme in our setting has equivalent complexity to optimizing quantum Fisher Information (QFI) [38, 39], and, in the case of Gaussian priors, we establish a direct link between these two approaches. We also argue that the Bayesian approach is more fundamental in the study of the limitations on the performance of quantum enhanced atomic clocks—a point of view shared by other authors [33, 36, 40, 41].

2. Cramér–Rao bound approach

In the case of a probe state \( \rho_0 \) with an encoded parameter \( \omega \) to be estimated, the standard quantum Cramér–Rao bound (CRB) states that, regardless of the measurements and unbiased estimators used, the estimation variance \( \Delta^2 \tilde{\omega} \) is lower bounded by

\[
\Delta^2 \tilde{\omega} \geq \frac{1}{F(\rho_0)}, \quad F(\rho_0) = \text{Tr} \left( \rho_0 L_\omega^2 \right), \quad \frac{1}{2} \{ L_\omega, \rho_0 \} = \frac{d\rho_0}{d\omega},
\]

where \( F \) is the QFI, \( \{ \cdot , \cdot \} \) is the anticommutator, and \( L_\omega \), implicitly defined above, is the symmetric logarithmic derivative of \( \rho_0 \). When estimating frequency, the probe state will typically be:

\[
\rho_0 = e^{-iH_\omega t} \Lambda_t(\rho) e^{iH_\omega t},
\]

where \( \rho \) is the interferometer input state, \( t \) is the interrogation time, \( H \) is the generator of the unitary evolution encoding \( \omega \), while \( \Lambda_t \) represents decoherence processes. In this case, since \( \frac{d\rho_0}{d\omega} = it [\rho_0, H], F(\rho_0) \) does not depend on \( \omega \) and reads:

\[
F(\rho_0) = F(\Lambda_t(\rho), Ht) = 2t^2 \sum_{ij} \frac{|\langle i|H|j \rangle|^2 (\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j},
\]

where \( \lambda_i, |i\rangle \) are the eigenvalues and the eigenvectors of \( \Lambda_t(\rho) \), respectively. The optimal strategy is obtained by maximizing \( F \), leading to the optimal \( \rho, t \) and the optimal projective measurement given by the eigenbasis of \( L_\omega \). Although applied fruitfully in many realistic metrological scenarios, including lossy interferometry [17, 19, 20, 28] and noisy frequency estimation [23, 28, 31, 32], the CRB approach suffers from a number of drawbacks. CRB is not saturable in general. It is only guaranteed to be saturable in special cases including Gaussian models, or in a repeated independent experiment framework (for a large number \( k \) of
experiments, it is possible to achieve $\Delta^2 \tilde{\omega} \approx \frac{1}{kF}$ [42]. Moreover, since the QFI is a local quantity—for a given $\omega$ it only depends on $\rho_0$ and its first derivative—it completely ignores any possible ambiguities in a reconstruction of the frequency value from a phase value that may arise for a sufficiently broad prior parameter distribution.

In frequency estimation, the interrogation time $t$ is a controllable parameter subject to optimization. If $t$ is large, even a very narrow prior distribution in $\omega$ may result in a phase $\phi_\omega = t$ distribution outside a local regime, or even broad enough on the $[0, 2\pi]$ interval for reconstruction ambiguities to become relevant. The validity of the local regime, where CRB based conclusions hold, cannot be a priori assumed. It depends on all the details of the estimation scheme, the LO noise, the interrogation time, and the initial state. In contrast, the Bayesian approach allows for full control of all the issues raised above, and any optimal scheme derived in this framework can be used as a universal benchmark. In certain parameter regimes, the Bayesian approach may produce results compatible with the QFI approach, but this is not the case in general. The Bayesian approach has been present in quantum estimation literature from the beginning [38]. Nevertheless, a strict Bayesian approach is often computationally challenging, and rigorous solutions are scarce and limited to decoherence-free scenarios [9, 43]. In this paper we show that the frequency estimation problem described above is efficiently solvable within the Bayesian framework, even in the presence of arbitrarily time-correlated collective dephasing processes.

3. Bayesian approach

Let $p_\omega$ be the prior distribution and $\rho_\omega$ the evolved probe state. Without loss of generality we assume $\int d\omega p_\omega \omega = 0$. The state $\rho_\omega$ is subject to a POVM measurement [42] $\{\Pi_x\}$, $\Pi_x \geq 0$, $\int dx \Pi_x = 1$, and the parameter $\omega$ is estimated on the basis of a measurement result $x$ using an estimator function $\tilde{\omega}_x$. For optimal performance, the average estimation variance

$$\Delta^2 \tilde{\omega} = \int d\omega dx p_\omega \text{Tr} \left( (\rho_\omega \Pi_x)(\omega - \tilde{\omega}_x)^2 \right)$$

should be minimized over $\rho_\omega$, $\{\Pi_x\}$, $\tilde{\omega}_x$, as well as the interrogation time $t$. It was shown [44] (see also [38, Chapter 8]) that the optimal measurement may be restricted to the class of standard projection von Neumann measurements $\Pi_x = |x\rangle \langle x|$, $\delta_{x,x'} = \delta_{x,x'}$, and the full information on the measurement-estimation strategy is contained in a single observable $L = \int \tilde{\omega}_x |x\rangle \langle x| dx$. Optimization of $L$ yields the minimal variance

$$\Delta^2 \tilde{\omega} = \Delta^2 \omega - \text{Tr} \left( \rho L^2 \right), \quad \frac{1}{2} \{L, \rho\} = \rho', \quad \text{equation (5)}$$

where $\Delta^2 \omega$ is the variance of the prior distribution $p_\omega$, and the optimal $L$ is implicitly given by equation (5), where $\rho = \int d\omega p_\omega \rho_\omega$, $\rho' = \int d\omega p_\omega \rho_\omega \omega$. A simple derivation of the above formula, together with a new elementary proof of its optimality, is given in appendix A. The optimal estimator $\tilde{\omega}(x)$ equals the mean of the updated frequency distribution $p_\omega \text{Tr} (\rho_\omega \Pi_x)$, which is the prior for the next experiment.

For a fixed initial state $\rho$, the optimization amounts to solving the anti-commutator equation (5) for $L$ and is equivalent in complexity to a computation of QFI for a mixed state. Full optimization requires further optimization of $\Delta^2 \tilde{\omega}$ over $\rho$ and $t$. Interestingly, we have found that an iterative algorithm analogous to the one proposed in [43] is very effective. We
begin with a random input state and iteratively find the optimal measurements and corresponding states. The procedure converges to optimal solutions and, in the cases we have studied, outperforms the brute force optimization of the QFI, allowing us to obtain a solution in the number of particle regimes where the brute-force optimization ceases to be practical on a standard PC ($n \gtrsim 50$). The convergence of the procedure has been analyzed in [45], and its efficiency has been additionally corroborated in the optimization of phase estimation schemes in the presence of local dephasing and loss [46]. See appendix B for details of the implementation of the algorithm.

The similarity between equations (5) and (1) becomes even more evident when considering a Gaussian prior distribution $\rho_{\Delta \omega} \propto -\omega \exp(-\omega^2/2\Delta^2 \omega)$. In this case

$$\int \rho_{\Delta \omega} \rho_{\Delta \omega}^\prime \propto (\rho_{\Delta \omega} \rho_{\Delta \omega}^\prime)_{\Delta \omega} = \rho_{\Delta \omega} \rho_{\Delta \omega}^\prime,$$

with $F(\bar{\rho}, Ht)$ defined in equation (3). Looking for the optimal probe states in the Bayesian approach is thus equivalent to maximizing the QFI for a spread-out state $\bar{\rho}$.

### 4. Decoherence-free frequency estimation

Let us first consider an idealized estimation model, in which $N$ two-level atoms are subject to unitary evolution $e^{-iHt}$ with the generator $H = \sum_{i=1}^{N} |i1\rangle\langle i|$, where $|i1\rangle\langle i|$ is the projector on the excited state of the $i$th particle. Without loss of optimality, we may assume that the input probe state is pure $\rho = |\psi\rangle \langle \psi|$ and is supported on the symmetric subspace [33]. Let $|n\rangle$ denote a symmetric state with $n$ atoms in $|1\rangle$, then $|\psi\rangle = \sum_{n=0}^{N} c_n |n\rangle$ and $H = \sum_{n} n |n\rangle \langle n|$. For a Gaussian prior, the averaged state $\bar{\rho}$ in the $|n\rangle$ basis reads:

$$\bar{\rho}_{nm} = \rho_{nm} e^{-(n-m)^2/2\Delta^2 \omega}/2,$$

where $\rho_{nm} = \langle n| \rho |m\rangle = c_n^* c_m$. The results of the numerical minimization of $\Delta^2 \tilde{\phi}$ with respect to $|\psi\rangle$ as a function of interrogation time for different atom numbers are presented in figure 2, where we introduced the optimal variance reduction factor as $R(\tau) = \frac{\Delta^2 \tilde{\phi}(\tau)|\sqrt{\rho_{\Delta \omega}}|}{\Delta^2 \omega}$, with $\tau$ being the dimensionless time parameter. Minimal estimation variances achievable with non-entangled states are presented for comparison. The variances which correspond to the optimal interrogation times are depicted in the inset as a function of $N$. For the optimal states, the curve slowly approaches the $1/N$ Heisenberg scaling, while for product states it is limited by $1/\sqrt{N}$.

For small times $\tau \ll 1/N$, the GHZ states $|\psi\rangle = (|0\rangle + |N\rangle)/\sqrt{2}$ minimize $R(\tau)$, as the effective prior phase distribution is narrow enough not to suffer from the characteristic GHZ $2\pi/N$ ambiguity. In this regime, the exact formula for the minimal variance reads $R(\tau) = 1 - N^2 \tau^2 \exp(-N^2 \tau^2)$. For times $\tau \gg 1/N$, the optimal states closely approach the ‘sine’ states introduced in [9, 47], while in the intermediate regime, the optimal states have a structure which interpolates between the GHZ and ‘sine’ states. Note also that there is no point in increasing the interrogation time too much, $\tau \gtrsim 1$, as inferring a frequency value becomes ambiguous, and thus the overall estimation variance increases. Thanks to equation (6) relating the Bayesian cost and the QFI, one can get a detailed insight into the structure of the optimal states by invoking the results presented in [48], where the C–R bound approach in presence of global dephasing has been pursued (see, e.g., figure 2(i) of [48]). One just needs to identify the
collective dephasing parameter $\Gamma_0$ from [48] with $\tau^2$ in our formulas. We should also add that instead of using the exact optimal states, which may be difficult to deal with in practice, we have checked numerically that the optimal performance may be closely approached using appropriately prepared one- or two-axis spin squeezed states, even though their structure differs significantly from that of the optimal states.

5. Frequency estimation under dephasing

We now consider a model with time dependent $\omega$ that reflects the relevant aspects of atomic clock operation as depicted in figure 1. The dominant decoherence effects are the LO noise [3, 34] and the collective dephasing of atoms [11, 23, 49]. Local dephasing of atoms [21] or loss [50] might also be relevant in certain atomic interferometry setups, but we neglect them here for clarity of presentation. Let $\omega_{\text{LO}}(t)$ be a stochastic process describing LO noise and $\omega(t) = \omega_{\text{LO}}(t) - \omega_0$ the corresponding detuning. $\omega_{\text{LO}}(0)$ represents the LO frequency, which stochastically drifted and was corrected during preceding feedback cycles. For a given realization of $\omega(t)$, the output state of the atoms after the interrogation time $t$ is $\rho_t = U_t \rho U_t^\dagger$, where $U_t = e^{-iH_0 \int_0^t dx (\omega(x) + \Omega(x))}$, and $\Omega(t)$ is the stochastic process representing the non-LO sources of the collective dephasing of the atoms. Measurement of $\rho_t$ yields a result $x$ with probability equal to $\text{Tr}(\rho_t \Pi_x)$. The estimator $\tilde{\omega}(x)$ is used to correct the LO frequency. The outcome frequency $\omega(t) - \tilde{\omega}(x)$ has a variance

![Figure 2. Frequency variance reduction for the optimal estimation strategies (solid) compared with optimized measurement strategies for uncorrelated states (dashed) and the GHZ states (dotted) as a function of evolution time for a different number of atoms $N$. Time $\tau$ is given in natural units of inverse prior frequency uncertainty. The inset depicts the dependence of estimation variance at the optimal time (dotted lines in main figure) as a function of $N$ for optimal (solid) and uncorrelated (dashed) states. The shaded areas serve as a guide to the eye to judge the character of the scaling of precision with $N$. Prior distribution is assumed to be Gaussian.](image-url)
\[ \Delta^2 \bar{\omega}(t) = \left\langle \int dx \, \text{Tr} \left( \rho \Pi_x \right) \left[ \omega(t) - \bar{\omega}(x) \right]^2 \right\rangle, \]

where \( \langle \cdot \rangle \) denotes averaging with respect to the stochastic processes \( \omega(t) \) and \( \Omega(t) \). Averaging over \( \omega(t) \) corresponds to the average over the prior in equation (4). The optimal estimation strategy yields the same formula for the minimal variance as in the decoherence-free case above, equation (5), but now with \( \Delta^2 \omega(t) = \langle \omega(t)^2 \rangle, \bar{\rho} = \langle \rho_i \rangle \) and \( \bar{\rho} = \langle \omega(t) \rho_i \rangle \).

When \( \Omega(t) \), \( \omega(t) \) are independent Gaussian processes with zero means, we obtain (see appendix C for the derivation):

\[ \bar{\rho}_{\text{hm}} = \rho_{\text{hm}} e^{-(n-m)^2(t)}[K^X_x(t)+K^X_2(t)]/2, \quad \bar{\rho}' = i\hbar \left[ \bar{\rho}, H \right] K^\omega_1(t), \]

where \( K^X_2(t) = \frac{1}{\omega} \int_0^t \int_0^t dt_1 dt_2 K^X(t_1, t_2), \quad K^X_1(t) = \frac{1}{\omega} \int_0^t dt_1 K^X(t_1, t), \quad \text{and} \quad K^X(t_1, t_2) = \langle X(t_1)X(t_2) \rangle \text{ is the two point correlation function of the process } X(t). \]

Again, making a connection to the QFI, we have

\[ \Delta^2 \bar{\omega}(t) = \Delta^2 \omega(t) - K^\omega_1(t)^2 F(\bar{\rho}, Ht). \]

Calculation of \( \bar{\rho} \) can be viewed as averaging \( \rho \) with respect to a Gaussian ‘effective prior’ with variance

\[ \Delta^2 \omega_K(t) = K^\omega_2(t) + K^\Omega_2(t). \]

It is clear that the optimal input state for Bayesian frequency estimation in the presence of decoherence is that which maximizes the QFI after being spread out with the ‘effective prior’. This implies that the solution in the presence of decoherence is in our case already contained in the solution of the decoherence-free case and therefore requires no additional numerical optimization. Using the variance reduction factor \( R(\tau) \) introduced for the decoherence-free case and depicted in figure 2, the formula for the minimal \( \Delta^2 \bar{\omega}(t) \) in the presence of decoherence reads:

\[ \Delta^2 \bar{\omega}(t) = \Delta^2 \omega(t) - \frac{K^\omega_1(t)^2}{\Delta^2 \omega_K(t)} \left[ 1 - R\left( t \sqrt{\Delta^2 \omega_K(t)} \right) \right]. \]

As an example, consider \( \omega(t) = (\omega_{\text{LO}}(0) - \omega_0)e^{-\gamma t} + f^\text{OU}(t) \). This is the Ornstein–Uhlenbeck (OU) process with the initial value \( \omega_{\text{LO}}(0) - \omega_0 \) and the correlation function of the zero-mean stochastic term \( f^\text{OU}(t) \), given by \( K^\text{OU}(t_1, t_2) = \alpha \left[ e^{-\gamma(t_1-t_2)} - e^{-\gamma(t_1+t_2)} \right] \). We assume that \( \omega_{\text{LO}}(0) \) is normally distributed with \( \langle \omega_{\text{LO}}(0) \rangle = \omega_0 \) and is independent from \( f^\text{OU}(t) \); hence:

\[ \Delta^2 \omega(t) = K^\omega(t, t) = \Delta^2 \omega(0) e^{2\gamma t} + \alpha \left( 1 - e^{-2\gamma t} \right), \]

which for times smaller than the OU correlation time \( t \ll \gamma^{-1} \) yields the diffusive character of frequency distribution broadening \( \Delta^2 \omega(t) \approx \Delta^2 \omega(0) + 2\gamma \alpha \left[ \alpha - \Delta^2 \omega(0) \right] \). The OU process is not perfect in representing noise in real LOs such as lasers, where \( 1/f \) noise for low frequencies and white noise for larger frequencies is typical [34, 51–53]. It may, however, well approximate the real noise in the low frequency regime, with simple analytic formulas for the relevant quantities:

\[ K^\omega_1(t) = \left[ \Delta^2 \omega(0)(1 - e^{-\gamma t}) + \alpha(1 - e^{-\gamma t})^2 \right]/(\gamma t), \quad K^\omega_2(t) = \left[ \Delta^2 \omega(0)(1 - \alpha)(e^{-\gamma t} - 1) + 2\alpha(\gamma t + e^{-\gamma t} - 1) \right]/(\gamma t)^2 \].

In order to take into account the white noise contribution present in real systems, we represent it in the atomic dephasing process \( \Omega \) with \( K^\Omega(t_1, t_2) = \beta \delta(t_1 - t_2), \quad K^\Omega_2(t) = \beta/\hbar t \). We exclude the white noise part from the LO noise, as this would yield infinite variances \( \Delta^2 \omega(t) \), but take it into account as a factor in atomic decoherence. This is consistent with the fact that high frequency white noise contributes to the
tails of atomic Lorenzian line shapes without significantly affecting their full-width at half maximum [54]. Using variance as a measure of the width of the line is justified, provided only low frequency noise is taken into account. Treating [53] as a guideline, we choose \( \alpha = 1 \text{ Hz}^2 \), \( \gamma = 0.2 \text{ Hz} \), \( \beta = 0.001 \text{ Hz} \) so that the OU power spectrum approximates reality for small frequencies \( \lesssim 10 \text{ Hz} \) and set \( \beta = 10^{-3} \text{ Hz} \) to represent the effects of high frequency white noise. The reduction of frequency variance as a function of the interrogation time and the number of atoms used is presented in figure 3. We took initial frequency variance \( \Delta^2 \omega(t) = 0.167 \text{ Hz}^2 \) so that the blue solid curve representing the optimal quantum strategy for \( N = 5 \) touches the \( \Delta^2 \omega(t)/\Delta^2 \omega(0) = 1 \) line. This indicates that a stationary operation, where the spread due to frequency noise is compensated by estimation feedback, is possible with optimally entangled states of \( N = 5 \) atoms. It is also clear that using product states with the same number of atoms (blue dashed) or a smaller number of entangled atoms does not permit us to keep frequency uncertainty at this level. For a given \( N \), varying \( t \) to minimize \( \Delta^2 \omega(0) \) while keeping the stationary condition fulfilled yields the optimal stationary strategy. The inset depicts the minimal stationary variances \( \Delta^2 \omega(0) \) as a function of \( N \) for the optimal (solid) and product (dashed) probe states. 

**Figure 3.** Frequency variance reduction for initial variance \( \Delta^2 \omega(0) = 0.167 \text{ Hz}^2 \) as a function of interrogation time for optimal (solid) and product (dashed) probe states for a different number of atoms \( N \), taking into account LO and atomic frequency noise with noise parameters \( \alpha = 1 \text{ Hz}^2 \), \( \gamma = 0.2 \text{ Hz} \), \( \beta = 0.001 \text{ Hz} \). Strategies in the gray region permit keeping the variance below \( \Delta^2 \omega(0) \) level. The inset depicts a log-log plot of the minimal stationary variances \( \Delta^2 \omega(0) \) as a function of \( N \) for the optimal (solid) and product (dashed) probe states.
6. Summary

To summarize, we have presented a consistent Bayesian approach to studying the potential quantum enhancement in frequency calibration. The tools presented are general enough to be able to deal with more sophisticated and promising atomic clock setups, including many simultaneously evolving atomic ensembles [52, 55]. However, a rigorous connection between our results and the performance of the actual atomic clock requires the study of the Allan variance, instead of instantaneous variance, as a figure of merit. The Allan variance takes into account the effects of the correlations between subsequent interrogation steps, i.e., $\omega_{LO}(0)$ and $\omega_{LO}(t) - \omega_{LO}(0)$, which is more adequate in the quantification of the performance of atomic clocks but at the same time makes the study of the optimal Bayesian strategies much more involved. We hope to address this problem in the near future. Actually, some work has already been done recently in this direction, using bounds on precision rather than considering optimal estimation strategies [40] or taking a numerical approach based on semi-definite programming [41].

Acknowledgements

We thank Wojciech Wasilewski, Jan Kołodyński, Marcin Jarzyna, Andrew Ludlow, and Mădălin Guță for fruitful discussions. This research was supported by the Polish NCBiR under the ERA-NET CHIST-ERA project QUASAR, the Foundation for Polish Science TEAM project, and the FP7 IP project SIQS co-financed by the Polish Ministry of Science and Higher Education. MF was supported by the Swiss National Science Foundation.

Appendix A. Optimal quantum Bayesian estimation strategy for the quadratic cost function

Here we prove that minimization of the average estimation variance $\Delta^2 \tilde{\omega}$, as given in equation (4), with respect to the choice of measurements $\Pi_x$ and estimators $\tilde{\omega}_x$, yields equation (5). Equation (4) can be rewritten as:

$$\Delta^2 \tilde{\omega} = \Delta^2 \omega + \text{Tr} (\hat{\rho} L_2) - 2 \text{Tr} (\hat{\rho}' L),$$

where $\hat{\rho} = \int d\omega \rho_{\omega} \rho_{\omega}$, $\hat{\rho}' = \int d\omega \rho_{\omega} \rho_{\omega} \omega$, $L = \int dx \tilde{\omega}_x \Pi_x$ and $L_2 = \int dx \tilde{\omega}_x^2 \Pi_x$. First we prove that, without loss of optimality, we can restrict the class of measurements to standard projective von Neumann measurements so that $x \in \{1, ..., d\}$ where $d$ is the dimension of the relevant Hilbert space, and $\Pi_x = |x\rangle \langle x|, \langle x \tilde{x}' \rangle = \delta_{x,x'}$.

Let $\Pi_x, \tilde{\omega}_x$ be a measurement (not necessarily projective) and the accompanying estimator. We write the eigendecomposition of the corresponding $L = \int dx \tilde{\omega}_x \Pi_x$ as:

$$L = \int \tilde{\omega}_y |y\rangle \langle y| \ dy = L^{\text{proj}},$$

where the eigenprojectors $|y\rangle \langle y|$ and the eigenvalues $\tilde{\omega}_y$ can now be interpreted as new measurement operators and estimators, respectively; we will refer to this strategy as the projective strategy. Below, we will prove that replacing the original strategy with the projective one can only reduce the estimation variance. Inspecting equation (A.1), we notice that information on measurement and estimation strategy enters the formula only through $L$ and $L_2$. 

...
operators. Had the projective strategy been used, we would have obtained:

\[ L_{2}^{\text{proj}} = \int \delta_{y}^{2} |y\rangle\langle y| \, dy = L^{2}. \]  

(A.3)

In order to show that we can only benefit by replacing the original strategy with the projective one, it is sufficient to prove that \( L_{2} - L^{2} \geq 0 \), as this implies \( \text{Tr} (\rho L_{2}) \geq \text{Tr} (\rho L^{2}) \), and, consequently, the resulting estimation variance of the projective strategy is not greater than the original one. More explicitly, we need to demonstrate that:

\[ \int dx \Pi_{x} \bar{\omega}_{x}^{2} - \left( \int dx \Pi_{x} \omega_{x} \right)^{2} \geq 0. \]  

(A.4)

The above inequality, however, is a special case of operator generalization of Jensen’s inequality [56] when applied to the operator convex function \( f(t) = t^{2} \). The operators \( \Pi_{x} \) now play the role of weights in the convex combination.

Having proved the optimality of the projective strategy, we can substitute \( L_{2} = L^{2} \) in equation (A.1) and, as result, face a simple problem of the minimization of:

\[ \Delta^{2} \bar{\omega} = \Delta^{2} \omega + \text{Tr} \left( \rho L^{2} - 2\bar{\rho}' L \right) \]  

(A.5)

with respect to a single hermitian matrix \( L \) and no additional constraints on optimization. Explicit differentiation with regard to matrix elements \( L_{ij} \) leads to:

\[ \sum_{k} \bar{\rho}_{ki} L_{jk} + \rho_{jk} L_{ki} - 2\bar{\rho}'_{ji} = 0, \]  

(A.6)

which in compact notation results in equation (5).

We note that the optimal estimator \( L \) can have both continuous or discrete spectra. The former is the case, for example, for Gaussian states, where mean position is the parameter to be estimated, in which case \( L \) is proportional to the position operator. We should also remark that since we have put no restriction on the allowed measurements, they could be the most general POVMs [42]. In particular, this includes the adaptive measurement strategies where measurements on part of the atoms depend on the results obtained from previous measurements on the other part of the atoms; see, e.g., discussion in [57]. This implies that such adaptive strategies are not advantageous for the present problem, as the simple von Neumann measurement has been proven optimal.

### Appendix B. Efficient iterative procedure for finding the optimal strategy numerically

Using equation (5) allows one to find the optimal estimation strategy for a given input probe state \( \rho \) and the corresponding evolved state \( \rho_{\omega} = \Lambda_{\omega} (\rho) \). Finding the optimal strategy, however, also requires the optimization of the input probe state. The brute force optimization of equation (5) over \( \rho \) proves to be highly inefficient. Here we propose a heuristic iterative procedure that has proved to be extremely efficient in searching for the optimal strategy. We know that for a particular input probe state, the corresponding optimal measurement and estimation strategy are given by the explicit formula in equation (5). The opposite are also true. Given a particular measurement and estimator, the corresponding optimal input probe state can easily be found, as demonstrated below.
We start by rewriting equation (A.1) as:

\[ \Delta^2 \tilde{\omega} = \Delta^2 \omega + \text{Tr} \left[ \int \text{d} \omega \Pi \langle \rho \rangle (L_2 - 2\omega L) \right]. \]  

(B.1)

Let us define a map \( \Lambda^* \) that is dual to \( \Lambda \), i.e., \( \text{Tr} [\Lambda (\rho) A] = \text{Tr} [\rho \Lambda^* (A)] \). Thus, an equivalent expression for \( \Delta^2 \tilde{\omega} \) reads:

\[ \Delta^2 \tilde{\omega} = \Delta^2 \omega + \text{Tr} \left[ \rho \int \text{d} \omega \Pi \Lambda^* (L_2 - 2\omega L) \right]. \]  

(B.2)

Therefore, the optimal input probe state is pure \( \rho = | \psi \rangle \langle \psi | \), where \( | \psi \rangle \langle \psi | \) should be chosen as the eigenvector of \( \int \text{d} \omega \Pi \Lambda^* (L_2 - 2\omega L) \) operator, corresponding to its most negative eigenvalue.

The heuristic procedure we advocate is now as follows. We start with a random input pure state \( \rho^{(0)} = | \psi^{(0)} \rangle \langle \psi^{(0)} | \) (or a state that one expects to be close to the optimal one, but preferably with some small random noise). Using equation (5) we find the corresponding optimal projective measurement strategy \( L^{(0)} \). We now calculate operator \( \int \text{d} \omega \Pi \Lambda^* (L^{(0)}_2 - 2\omega L^{(0)}) \) (where \( L^{(0)}_2 \) was replaced by \( L^{(0),2} \), since the measurement is projective) and find the eigenvector corresponding to its most negative eigenvalue. This way we obtain \( | \psi^{(1)} \rangle \). We repeat the procedure until we arrive at a satisfactory convergence. The results presented in figure 2 were obtained exactly in this way for \( \Lambda (\rho) = e^{-iHt} \rho e^{iHt} \) and a Gaussian prior \( \rho_0 \). This method can be applied to any problem involving optimization of the Fisher information with respect to input probe states and is described in detail in [45], where also the analysis of the convergence of the algorithm is given.

Appendix C. Minimal frequency estimation variance under LO fluctuations and atomic dephasing—equations (9), (10)

If we rewrite equation (8) analogously as in equation (A.1), the LO frequency variance immediately after the estimation procedure reads:

\[ \Delta^2 \tilde{\omega} (t) = \langle \omega (t)^2 \rangle + \text{Tr} (\tilde{\rho} L_2) - 2 \text{Tr} (L \tilde{\rho}'), \]  

(C.1)

where \( \tilde{\rho} = \langle \rho \rangle \), \( \tilde{\rho}' = \langle \omega (t) \rho \rangle \), \( L = \int \text{d} x \tilde{\omega} \Pi \), and \( L_2 = \int \text{d} x \tilde{\omega}^2 \Pi \). Optimization of the estimation strategy proceeds exactly as in the decoherence-free case and yields the minimal variance:

\[ \Delta^2 \tilde{\omega} (t) = \langle \omega (t)^2 \rangle - \text{Tr} (\tilde{\rho} L^2), \quad \frac{1}{2} \{L, \tilde{\rho}\} = \tilde{\rho}'. \]  

(C.2)

Matrix elements of \( \tilde{\rho} \) read:

\[ \tilde{\rho}_{mn} = \rho_{mn} e^{-i(n-m)} \int_0^t \text{d} s (\omega (s) + \Omega (s)) \]  

(C.3)

Since \( \omega (t) \) and \( \Omega (t) \) are independent Gaussian processes with zero means, the standard cumulant expansion method yields:

\[ \tilde{\rho}_{mn} = \rho_{mn} e^{-\frac{(n-m)^2}{2}} \int_0^t \text{d} s_1 \text{d} s_2 \langle (\omega (s_1) + \Omega (s_1)) (\omega (s_2) + \Omega (s_2)) \rangle = \rho_{mn} e^{-\frac{(n-m)^2}{2}} (K_{1}^2 + K_{2}^2). \]  

(C.4)

In order to apply the cumulant expansion method to the calculation of \( \tilde{\rho}' \), we write its matrix elements as
where $\delta (s - t)$ is the Dirac delta. This gives us:

$$\bar{\rho}'_{nm} = \frac{d}{d\xi} \left( e^{\int_{0}^{t} ds \, i(m-n)[\omega(s)+\Omega(s)]+\xi \omega(s)\delta(s-t)} \right) \bigg|_{\xi=0},$$  \hspace{1cm} (C.5)  

which also reads as:

$$\bar{\rho}' = it \left[ \bar{\rho}, H \right] K^{\omega}_{1}(t).$$  \hspace{1cm} (C.7)  

Taking into account the definition of the Fisher information, equations (1) and (3) and equation (C.2), we obtain the desired formula:

$$\Delta^{2}\omega(t) = \Delta^{2}\omega(t) - K^{\omega}_{1}(t)^{2} F(\bar{\rho}, Ht).$$  \hspace{1cm} (C.8)
[24] Shaji A and Caves C M 2007 Phys. Rev. A 76 032111  
[25] Sarovar M and Milburn G J 2006 J. Phys. A: Math. Gen. 39 8487  
[26] Giovannetti V, Lloyd S and Maccone L 2011 Nat. Photonics 5 222–9  
[27] Banaszek K, Demkowicz-Dobrzański R and Walmsley I A 2009 Nat. Photonics 3 673–6  
[28] Escher B M, de Matos Filho R L and Davidovich L 2011 Nat. Phys. 7 406–11  
[29] Demkowicz-Dobrzański R, Kołodyński J and Guță M 2012 Nat. Commun. 3 1063  
[30] Kołodyński J and Demkowicz-Dobrzański R 2013 New J. Phys. 15 073043  
[31] Chin A W, Huelga S F and Plenio M B 2012 Phys. Rev. Lett. 109 233601  
[32] Szankowski P, Chwedenczuk J and Trippenbach M 2012 arXiv:1212.2528  
[33] Bužek V, Derka R and Massar S 1999 Phys. Rev. Lett. 82 2207–10  
[34] André A, Sørensen A S and Lukin M D 2009 Science 325 1476–8  
[35] Mullan M and Knill E 2012 Quantum Inf. Comput. 12 553–74  
[36] Borregaard J and Sørensen A S 2013 Phys. Rev. Lett. 111 090801  
[37] Helstrom C W 1976 Quantum Detection and Estimation Theory (New York: Academic)  
[38] Braunstein S L and Caves C M 1994 Phys. Rev. Lett. 72 3439–43  
[39] Knysh S I, Chen E H and Durkin G A 2014 arXiv:1402.0495  
[40] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Hänsel W, Hennrich M and Blatt R 2011 Phys. Rev. Lett. 106 130506  
[41] Gross C, Zibold T, Nicklas E, Esteve J and Oberthaler M K 2010 Nature 464 1165–9  
[42] Numata K, Kemery A and Camp J 2004 Phys. Rev. Lett. 93 250602  
[43] Rosenband T and Leibrandt D R 2013 arXiv:1303.6357  
[44] Jiang Y Y, Ludlow A D, Lemke N D, Fox R W, Sherman J A, Ma L S and Oates C W 2011 Nat. Photonics 5 158–61  
[45] Domenico G D, Schilt S and Thomann P 2010 Appl. Opt. 49 4801–7  
[46] Hansen F and Pedersen G K 2003 Bull. London Math. Soc. 35 553–64  
[47] Kołodyński J and Demkowicz-Dobrzański R 2010 Phys. Rev. A 82 053804