

Report on Progress

Bayesian frequency metrology with optimal Ramsey interferometry in optical atomic clocks

T Kielinski¹  and K Hammerer^{1,2,3,*} 

¹ Institute for Theoretical Physics, Leibniz University Hannover, Appelstrasse 2, 30167 Hannover, Germany

² Institute for Theoretical Physics, University Innsbruck, Technikerstrasse 21a, 6020 Innsbruck, Austria

³ Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Technikerstrasse 21a, 6020 Innsbruck, Austria

E-mail: klemens.hammerer@itp.uni-hannover.de

Received 12 March 2025, revised 1 September 2025

Accepted for publication 19 November 2025

Published 10 December 2025

Corresponding editor: Dr Maciej Lewenstein



CrossMark

Abstract

Frequency metrology is a cornerstone of modern precision measurements and optical atomic clocks have emerged as one of the most precise measurement devices. In this progress report, we explore various Ramsey interrogation schemes tailored to optical atomic clocks primarily limited by laser noise. To incorporate frequency fluctuations directly into the theoretical model, we consider a Bayesian framework. In this context, we review fundamental bounds arising in Bayesian estimation theory, which serve as a benchmark throughout this work. We investigate the trade-off between entanglement-enhanced sensitivity and robustness against laser noise in order to identify optimal initial states, measurement schemes and estimation strategies. Beside standard protocols based on coherent spin states, spin-squeezed states and Greenberger–Horne–Zeilinger states, we consider variational Ramsey protocols implemented via low-depth quantum circuits based on one-axis twisting operations to approach optimal stability. In particular, we review known and identify new optimal Ramsey interrogation schemes for a variety of scenarios, including different experimental platforms, ensemble sizes and regimes characterized by a wide range of interrogation durations and dead times. Hence, this work establishes a comprehensive theoretical framework for optimizing Ramsey interrogation schemes, providing guidance for the development of next-generation optical atomic clocks.

Keywords: quantum metrology, frequency metrology, Ramsey interferometry, optical atomic clocks, atomic clocks, quantum clocks, Bayesian estimation

* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Contents

1. Introduction	2	6. Outlook	26
2. Bayesian frequency metrology	5	Data availability statement	27
2.1. Motivation	5	Acknowledgments	27
2.2. Bayesian phase estimation	6	Appendix	27
2.3. Bounds	7	A.1. Derivations of Bayesian bounds	27
2.4. Estimators	9	A.1.1. BCRB	27
2.5. ADEV	10	A.1.2. BQCRB	28
2.6. Interrogation time and prior width	11	A.1.3. OQI	29
3. Quantum frequency metrology	11	A.2. Estimators	31
3.1. Standard protocols	11	A.2.1. Linear estimator	31
3.2. Variational classes	13	A.2.2. Optimal Bayesian estimator	31
3.3. Optimal protocols	16	A.3. ADEV	31
4. Application in the full feedback loop of an atomic clock	17	A.4. Sensitivity of standard protocols	32
4.1. Limitation due to fringe hops	17	A.4.1. CSSs	32
4.2. Clock stability	18	A.4.2. SSSs	33
4.3. Comparison of linear and optimal Bayesian estimation	19	A.4.3. GHZ states	33
5. Dead time	20	A.5. Numerical methods	34
5.1. Dead time in Bayesian frequency metrology	21	A.5.1. Optimization	34
5.2. General results	21	A.5.2. Ramsey signals	34
5.3. Setup specific dead times	23	A.5.3. Clock simulation	35
		A.5.4. Iterative prior width	36
		A.5.5. Prior width with dead time	37
		References	37

Acronym	Definition	Reference
ADEV	Allan deviation	Section 2.5
BMSE	Bayesian mean squared error	Section 2.2
BCRB	Bayesian Cramér–Rao bound	Section 2.3
BQCRB	Bayesian quantum Cramér–Rao bound	Section 2.3
CRB	Cramér–Rao bound	Section 2.3
CSSs	coherent spin states	Section 3.1
CTL	coherence time limit	Section 3.1
GHZ	Greenberger–Horne–Zeilinger	Section 3.1
HL	Heisenberg limit	Section 3.1
π HL	π -corrected Heisenberg limit	Section 2.3
LO	local oscillator	Section 2.1
MSE	mean squared error	Section 2.2
OAT	one-axis-twisting	Section 3.1
OQI	optimal quantum interferometer	Section 2.3
POI	phase operator based interferometer	Section 2.3
QFI	quantum Fisher information	Section 2.3
QPN	quantum projection noise	Section 2.4
SLD	symmetric logarithmic derivative	Section 2.4
SQL	standard quantum limit	Section 2.4
SSSs	spin-squeezed states	Section 3.1

1. Introduction

Frequency metrology constitutes a fundamental pillar in modern precision measurements, driving advancements across a wide range of scientific and technological fields [1–6]. At the forefront of this discipline are optical atomic clocks, which exploit narrow-linewidth atomic transitions in the optical domain [7, 8]. This new generation of clocks was spurred by technological advances over the past decades, including breakthroughs in laser technology [7], the invention of the

optical frequency comb [9, 10], and the development of highly controllable platforms such as ion-traps [11–13], tweezer-arrays [14–17], and optical lattices [18–21]. Today, state-of-the-art optical atomic clocks are among the most precise measurement devices ever built, achieving stabilities on the order of 10^{-18} and below [13, 22–32]. To illustrate this incredible precision, such clocks would gain or lose less than a second over the age of the Universe. They have surpassed traditional microwave-based Caesium atomic clocks, which long served as the standard for timekeeping, paving the way for the

redefinition of the SI second [7, 33]. This unprecedented stability renders optical clocks indispensable tools for a broad spectrum of applications. In research, they are instrumental in probing fundamental physics, from testing general relativity through gravitational redshift measurements [31, 33–37] to exploring variations in fundamental constants [38, 39] and searching for new physics beyond the standard model [40–42]. In technology, optical atomic clocks foster potential applications ranging from enhancing global navigation satellite systems [43, 44] and synchronizing large-scale networks [45] to supporting precision geodesy [46–49].

QPN is the most fundamental process limiting clock stability, arising from the stochastic nature of quantum measurements and the discrete outcomes inherent in finite-size ensembles [50, 51]. For separable states of many atoms, namely uncorrelated or classically-correlated states, the SQL imposes a fundamental bound on QPN [1, 7, 8, 52]. However, stability beyond this classical limit can be achieved by introducing entanglement within the atomic ensemble [1, 8, 52]. Three decades ago, Wineland *et al* proposed in seminal works [53, 54] to entangle cold ions via their common coupling to collective modes of motion to suppress projection noise in frequency spectroscopy, thereby overcoming the SQL and enhancing atomic clock stability. With the momentous advancements in optical atomic clocks and programmable quantum processors since then, this vision now encounters new opportunities and challenges. In recent years, entanglement on optical clock transitions has been demonstrated in various setups, including the generation of spin squeezing in trapped ions [55] and in neutral atoms mediated by cavities [27, 56] or Rydberg interactions [57]. Recently, also maximally entangled GHZ states and cascades thereof have been realized in optical clocks based on tweezer-arrays [58, 59]. In an ideal scenario, these GHZ states saturate the HL, representing the ultimate bound on projection noise and yielding a quadratic improvement over the SQL [1].

However, in realistic scenarios, decoherence processes and external noise degrade the coherence of the quantum system, impairing the stability and preventing the achievement of the HL [60–63]. While entanglement promises to overcome the SQL and thereby improving clock stability, the detrimental effects of decoherence are particularly pronounced in entangled states, since they are highly susceptible to the loss of coherence. In particular, Huelga *et al* have demonstrated that GHZ protocols suffer significantly from dephasing associated with random phase changes caused by stray fields or laser noise, ultimately showing no improvement over the SQL [64]. In contrast, GHZ-like states offer optimal stability for clocks with small ensembles that are primarily limited by spontaneous decay [65]. As a consequence, incorporating decoherence effects and external noise is crucial for identifying optimal interrogation protocols in frequency metrology.

Unlike magnetic field fluctuations or laser noise, the finite lifetime of qubits in the excited state represents a fundamental limit rather than an external noise source, a limitation examined in detail in [65]. State-of-the-art clock lasers achieve coherence times of several seconds [66], entering the regime of

the excited-state lifetime of various clock candidates, such as Sr^+ -ions (0.4 s) or Ca^+ -ions (1.1 s). With further technological improvements in the short-term laser stabilization, coherence times will potentially approach lifetimes of further clock species as Yb-atoms (15.9 s) or Al^+ -ions (20.7 s). Nevertheless, the excited-state lifetimes of several clock candidates remain far beyond the regime of laser coherence times, ranging from minutes (Sr-atoms) to years (Yb⁺-ions) [22]. Moreover, the impressive level of laser coherence is often degraded during propagation from the cavity to the location of the qubits. Consequently, a significant number of setups are currently, and will remain in the future, limited by laser noise.

Naively, frequency fluctuations and the associated laser noise could be regarded as a purely technical problem. However, stabilizing the laser is precisely the central objective of an atomic clock time standard, or equivalently a frequency standard, making frequency fluctuations the primary measurand [6]. Disregarding laser noise as a mere technical issue would thus contradict the fundamental concept of atomic clocks. In principle, one might also ask how laser noise can impose a limiting factor although, by definition, it is the measurand, the quantity to be stabilized. To be precise, only the component of laser noise that cannot be corrected through interrogation of the atomic reference ultimately limits clock stability. Since Ramsey protocols have a finite range within which they can unambiguously interpret frequency fluctuations, errors arise when laser noise exceeds this range, fundamentally constraining stability. Additionally, dead time in clock operation leads to undetected aliased frequency deviations, further degrading performance. Investigating the impact of frequency fluctuations is therefore essential for advancing next-generation clocks. To address this challenge, various approaches have been developed to account for frequency noise and to determine optimal interrogation schemes for specific experimental setups [67–70]. In this endeavor, a particularly promising framework is the application of Bayesian estimation theory to frequency metrology—denoted as Bayesian frequency metrology—which incorporates laser noise directly into the theoretical model and leverages prior knowledge of frequency fluctuations for estimation [71–74].

As a consequence, determining optimal interrogation protocols involves a trade-off between achieving entanglement-enhanced sensitivity, which enables surpassing the SQL, and ensuring robustness against noise. In recent years, operationally motivated echo protocols and variational quantum circuits have attracted significant interest, as they allow for a diverse range of interrogation schemes [73–86]. In particular, these approaches have the potential to generate a high degree of entanglement while maintaining resilience to noise [73–78]. OAT [87] interactions serve as a versatile tool for implementing such protocols as they give rise to a variety of entangled states, ranging from SSSs to GHZ states, and facilitate variational classes of generalized Ramsey protocols [73–78]. Furthermore, OAT interactions are accessible in several setups as in ion-traps via Mølmer–Sørensen gates [79, 88, 89], in tweezer-arrays via Rydberg interactions [57, 90] or Bose–Einstein condensates via elastic collisions [91–94].

This work presents a progress report on frequency metrology tailored to optical atomic clocks employing Ramsey interrogation schemes primarily limited by laser noise. Specifically, we focus on single-ensemble clocks in which the atomic reference is periodically interrogated using the same protocol in each clock cycle. The objective is to outline potential advancements and challenges across various Ramsey interrogation schemes, effectively providing a theoretical guide for clock operation on different experimental platforms. In particular, we systematically examine a broad range of ensemble sizes and regimes defined by interrogation duration and dead time. To incorporate frequency fluctuations into the theoretical model, we employ a Bayesian framework for single-ensemble clocks, where the atomic reference is periodically interrogated using the same protocol in each clock cycle, while more general schemes are addressed in the outlook. To establish a theoretical foundation, we review Bayesian estimation theory and the corresponding fundamental bounds on clock stability. Additionally, we incorporate previous findings on clocks limited by laser noise, such as those in [67, 68], within the Bayesian framework and extend them in certain regimes. Building on pioneering work in [73, 74, 76] on variational quantum circuits, we identify optimal Ramsey schemes for various experimental platforms. It should be noted that in certain setups, Rabi interrogation protocols may offer higher stabilities as they are less sensitive to high-frequency laser noise [28].

In the following, we provide a brief overview of each section and outline the primary results:

- **Section 2:** We introduce the fundamental principles of atomic clocks and establish the connection between frequency metrology and phase estimation theory. By highlighting the impact of LO noise on clock stability, we motivate the Bayesian approach as an effective framework for atomic clocks primarily limited by frequency fluctuations of the LO. To start with, the framework of Bayesian phase estimation theory is introduced and a hierarchy of lower bounds on the estimation uncertainty is collected, drawing an analogy to the local (frequentist) approach. In particular, the ultimate lower bound is derived, denoted as the OQI, which represents the primary benchmark in this work. Additionally, the linear estimation strategy is discussed and the optimal Bayesian estimator is determined. By introducing the ADEV, we explicitly connect Bayesian phase estimation theory to frequency metrology and establish a relation between interrogation time and prior knowledge of the phase. Furthermore, we discuss general trade-offs in the context of frequency metrology. Hence, this section introduces the fundamental concepts and establishes the theoretical framework of this work.
- **Section 3:** This section aims to saturate the ultimate lower bound imposed by the OQI. Initially, the standard protocols, utilizing CSSs, SSSs and GHZ states, are compared to the OQI. While GHZ states saturate the OQI at short interrogation times and SSS perform close to it at intermediate durations, substantial potential for enhancement remains across a broad range of interrogation times, particularly at long durations. To address this, especially considering small ensemble sizes characteristic of ion-traps and tweezer-arrays, we introduce generalized Ramsey protocols based on variational quantum circuits and identify optimal interrogation schemes. We demonstrate that in this regime, even low-depth quantum circuits suffice to approximate the OQI, which is crucial for maintaining reasonable operational complexity and thus enabling near-term experimental implementation. While the required circuit depth to achieve OQI stability increases with N , the performance gain diminishes with complexity, leading to a trade-off between reduced instability and increased experimental overhead, further motivating a focus on low circuit-complexity approaches.
- **Section 4:** To validate theoretical predictions on clock stability, we perform Monte Carlo simulations of the full feedback loop in an atomic clock, cf figure 1, from which we can infer its long-term stability as quantified by the ADEV. In this context, fringe hops emerge as a significant limitation. In particular, for small ensembles ($N \lesssim 20$), characteristic of ion-traps, fringe hops impose a stricter constraint on clock stability than the CTL of the LO. As a consequence, for long interrogation times, variational protocols provide marginal to no advantage over SSS, while GHZ states remain optimal at short interrogation times. In contrast, for ensemble sizes in the regime of tweezer-arrays ($N \gtrsim 20$), fringe hops and the CTL impose comparable limitations on clock stability at long interrogation times. Consequently, variational Ramsey protocols provide a substantial improvement over SSS. Nevertheless, the variation in stability across different clock runs, due to the stochastic nature of atomic clocks, and the relative reduction in enhancement with increasing circuit depth, further supports the focus on low-depth quantum circuits.
- **Section 5:** Incorporating dead time into atomic clock operation within the framework of Bayesian frequency metrology, we investigate the trade-off between QPN, the CTL, and dead time effects. While clock stability for short dead times or small ensembles closely resembles the dead time-free scenario, dead time effects become increasingly significant with growing ensemble size or dead time, ultimately limiting clock performance. Following a general analysis, we examine specific examples with state-of-the-art parameters relevant to different experimental platforms, such as ion-traps, tweezer-arrays and lattice clocks. While GHZ states and SSS remain optimal for ion-traps utilizing only a few ions, the potential gain from variational quantum circuits in tweezer-arrays with several tens of atoms is substantially diminished. Specifically, SSS perform close to the OQI across a wide range of interrogation times, whereas variational quantum circuits offer an enhancement only at long interrogation times. However, this improvement is significantly reduced compared to the dead time-free case. Additionally, in the presence of dead time, fringe hops remain the dominant limitation in this regime, whereas in the dead time-free case, they constrain clock stability only at the same

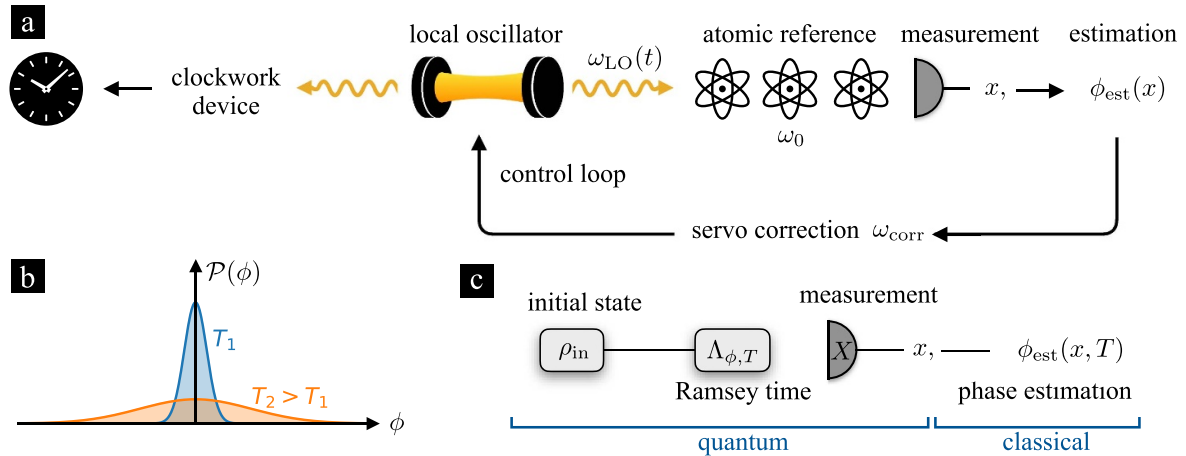


Figure 1. (a) Basic principle of an atomic clock: a local oscillator (LO) with fluctuating frequency $\omega_{\text{LO}}(t)$ is stabilized in a control loop to an atomic transition ω_0 . During the free evolution time T , the probe state accumulates a phase ϕ arising from the frequency deviation. Based on the measurement outcome x , the phase is estimated by ϕ_{est} and the LO frequency is corrected according to ω_{corr} by the servo. (b) Broadening of the prior distribution with increasing interrogation time. (c) Generalized Ramsey interferometry: the phase ϕ is encoded during the interrogation time T onto the initial state ρ_{in} via $\Lambda_{\phi, T}$. Based on the measurement outcome x of the observable X , an estimation ϕ_{est} of the phase is conducted.

level as the CTL. As a consequence, SSS emerge as the preferred choice due to their robustness and practicality. For lattice clocks with hundreds or thousands of atoms, dead time effects strictly constrain clock stability and thus CSS suffice to approximate the OQI.

Although this work is tailored to frequency metrology in atomic clocks, the developed tools and techniques extend beyond this specific application. In particular, they are broadly applicable to general frequency metrology and Ramsey interferometry, including implementations in atom interferometry and magnetometry.

2. Bayesian frequency metrology

2.1. Motivation

A clock, at its core, consists of two essential components: a frequency standard, a device which generates a continuous and consistent frequency signal, and a mechanism that counts the oscillations over time. While the clockwork device essentially translates the frequency signal into measurable time intervals, the frequency standard represents the true heart of a clock [6]. Frequency standards are commonly classified as either active or passive, depending on their operational principle. Active frequency standards generate their own oscillation at a given frequency, as the hydrogen maser or the Helium–Neon laser, where stimulated emission results in a highly coherent signal. Conversely, passive frequency standards require an external source to stimulate their oscillation. While active frequency standards typically excel in short-time stability, passive frequency standards often achieve superior long-term stability and accuracy, because the frequency can be precisely monitored and corrected against the reference response over time.

Consequently, passive frequency standards are commonly preferred for clocks [4–6].

The concept of a passive frequency standard can be illustrated by imagining two pendulums. The first pendulum is our primary, noisy pendulum, whose fluctuating frequency we aim to stabilize. The second pendulum serves as an (almost) ideal reference, though it does not oscillate on its own. Hence, the task of a passive frequency standard is to periodically adjust the primary pendulum’s frequency to match that of the reference pendulum by repeatedly measuring the frequency difference between the two [6]. However, each measurement introduces some noise into the system. Hence, it is desirable to extend the interrogation time as long as possible, thereby reducing the relative impact of this measurement noise and ultimately enhancing stability. If the interrogation time is extended too far, however, we risk missing a ‘tick’ of the reference, leading to synchronization errors that may accumulate over repeated measurements. Therefore, while longer interrogation times improve stability, there is an optimal duration beyond which stability is compromised [5–7].

In (passive) atomic clocks (cf figure 1(a)), the LO, representing the primary pendulum, generates an inherently noisy frequency signal $\omega_{\text{LO}}(t)$ that varies over time t . The LO is stabilized to an atomic transition frequency ω_0 , acting as the reference pendulum, through repeated interrogations of the atomic ensemble according to a particular Ramsey interferometry scheme. During the Ramsey time T , the atoms accumulate a phase $\phi = \int_t^{t+T} dt' \omega(t')$, which effectively reflects the average of the frequency deviation $\omega(t) = \omega_0 - \omega_{\text{LO}}(t)$ over the interrogation period. At the end of each interrogation sequence, a measurement with outcome x is performed, from which an estimate $\phi_{\text{est}}(x)$ of the monitored phase ϕ is derived. The control cycle is completed by the servo that applies feedback to correct the LO frequency by ω_{corr} , based

on the phase estimate $\phi_{\text{est}}(x)$, resulting in a stabilized LO signal. Consequently, frequency metrology is directly connected to phase estimation theory [1, 4–7].

LO noise plays a critical role in limiting the clock stability, even though it may seem counterintuitive given that its frequency is constantly measured and corrected based on feedback from the atomic reference. The phase fluctuations, arising from the LO's frequency noise, can be modeled by a distribution $\mathcal{P}(\phi)$, depending on the particular noise profile, by interpreting the phase ϕ as a random variable. As the interrogation time T increases, the LO noise grows, causing the distribution to broaden as illustrated in figure 1(b). Typically, the phase can only be estimated unambiguously within a limited range. Hence, if the invertible domain of the main fringe of the signal is exceeded, estimation errors are accumulated and ultimately limit the stability. In the worst case, the feedback loop passes to an adjacent Ramsey fringe, resulting in the clock running systematically wrong and severely degrading the clock stability. Consequently, frequency metrology features a trade-off: longer interrogation times enhance stability, but LO frequency fluctuations introduce limitations. An appropriate framework to investigate different interrogation schemes respecting this trade-off is represented by Bayesian estimation. This approach essentially bridges between high sensitivity at the transition frequency and large dynamic range, based on the regime of frequency deviations the LO is likely to generate [67–69].

Building on the motivation outlined above, it becomes evident that phase estimation is a fundamental aspect of frequency metrology, particularly in the context of atomic clocks. In this section, we begin by introducing general phase estimation theory within the Bayesian framework. By reviewing the literature, we establish a hierarchy of lower bounds on the estimation error, drawing an analogy to the local (frequentist) approach, and analyze different estimation strategies. Finally, we link phase estimation to frequency metrology in the context of atomic clocks through the ADEV and discuss emerging trade-offs and their implications for the clock stability.

2.2. Bayesian phase estimation

In interferometry, the objective is to estimate an unknown parameter ϕ as precise and accurate as possible. In generalized Ramsey spectroscopy (cf figure 1(c)), the phase ϕ is encoded onto the initial probe state ρ_{in} during the free evolution time T (Ramsey dark time) via a completely-positive trace-preserving map $\Lambda_{\phi,T}$. Additionally, this quantum channel $\Lambda_{\phi,T}$ may also include decoherence processes such as dephasing or spontaneous decay, with their impact depending on T . After the free evolution, an appropriately chosen observable X is measured. The measurement is described by a positive operator-valued measure (POVM) $\{\Pi_x\}$, where $\Pi_x \geq 0$ and $\sum_x \Pi_x = \mathbb{1}$, and x denotes the measurement outcome. Finally, an estimate $\phi_{\text{est}}(x)$ of the parameter ϕ is performed, based on the measurement outcome x (of X). In the context of an atomic clock, the phase $\phi = \omega T$ originates from the average frequency difference $\omega = \frac{1}{T} \int_0^T dt' [\omega_0 - \omega_{\text{LO}}(t')]$ between the atomic transition ω_0 and the LO ω_{LO} , while X typically corresponds to a

spin measurement. Due to the inherent indeterministic nature of quantum measurements, the outcomes x are random and occur with conditional probability

$$P(x|\phi) = \text{Tr}(\Pi_x \Lambda_{\phi,T}[\rho_{\text{in}}]). \quad (1)$$

Consequently, the estimator ϕ_{est} likewise is a random variable, as it depends on the measurement outcome x . Unlike state preparation, free evolution, and measurement, which are governed by quantum mechanics, the estimation process involves classical post-processing of measurement data and is thus addressed within the framework of classical phase estimation theory.

In local (or frequentist) phase estimation, it is typically assumed that the phase ϕ is tightly centered around a fixed working point ϕ_0 , such that $(\phi - \phi_0)^2 \ll 1$, and that the estimator is locally unbiased. Furthermore, probabilities are defined as the infinite-sample limit of an event. However, these assumptions are often not valid in the context of optical atomic clocks. When the finite coherence time of the laser becomes the dominant limitation on clock stability, fluctuations in the accumulated phase during the free interrogation time become relevant and in principle can take arbitrary values $-\infty < \phi < \infty$. Additionally, these fluctuations require phase estimation based on single measurements to ensure unambiguous determination of ϕ , as the phase may change significantly between measurements of consecutive cycles, potentially preventing a unique estimation or assignment. This constraint makes asymptotic estimation, i.e. the collection and averaging of large amounts of data, impossible.

To quantify the impact of these phase fluctuations and the resulting limitations to the phase estimation uncertainty, a commonly used cost function is the average MSE, defined as

$$(\Delta\phi)^2 = \int_{-\infty}^{+\infty} d\phi \mathcal{P}(\phi) \sum_x P(x|\phi) [\phi - \phi_{\text{est}}(x)]^2. \quad (2)$$

It corresponds to the MSE of the estimated phase $\phi_{\text{est}}(x)$ with respect to the true phase value ϕ —the typical cost function of local phase estimation—averaged over the phase distribution $\mathcal{P}(\phi)$, which characterizes the phase fluctuations in different clock cycles. The average MSE reflects a global approach—extending local (frequentist) phase estimation—by incorporating all possible values of ϕ , which additionally makes unbiasedness redundant. Moreover, this approach is well-suited for arbitrary signals and estimation strategies, as it assesses the overall performance by averaging over the entire phase distribution, eliminating the need for specific assumptions about the signal structure or the estimation method. In general, for a proper estimation strategy, further information about the phase is gained through the measurement. Consequently, the average MSE is smaller than the variance $(\delta\phi)^2$ of the prior phase distribution $\mathcal{P}(\phi)$ and thus, ultimately is bounded by $0 \leq (\Delta\phi)^2 \leq (\delta\phi)^2$. In the limit of narrow phase distributions, where $\mathcal{P}(\phi)$ approximates a delta distribution centered at the optimal working point ϕ_0 , the average MSE reduces to the MSE. Due to its global averaging, it is always

lower bounded by the MSE evaluated at the optimal working point ϕ_0 , where the MSE attains its maximum.

As argued above, local (frequentist) phase estimation is poorly suited in the regime where phase fluctuations around the optimal working point impose a substantial limitation. Although the accumulated phase is fixed for each interrogation sequence, it varies over different clock cycles and thus can be treated as a random variable, with the prior distribution $\mathcal{P}(\phi)$ reflecting the knowledge on the phase prior to any measurement. Consequently, Bayesian estimation theory represents the more appropriate framework. In Bayesian phase estimation, the posterior knowledge of ϕ , represented by the posterior distribution $P(\phi|x)$ and from which the estimator is ultimately determined, is updated according to Bayes theorem

$$P(\phi|x) = \frac{\mathcal{P}(\phi)P(x|\phi)}{P(x)} \quad (3)$$

based on the statistical model $P(x|\phi)$ and the prior distribution $\mathcal{P}(\phi)$. The marginal likelihood $P(x) = \int d\phi \mathcal{P}(\phi)P(x|\phi)$ represents the probability of observing outcome x , averaged over all possible values of ϕ . Thus, it essentially provides a normalization of the posterior distribution. The interplay between prior information and measurement data already becomes evident at this stage. If $\mathcal{P}(\phi)$ varies slowly compared to $P(x|\phi)$, for example in the case of a flat prior or in the asymptotic limit of large ensembles, it has minimal influence on the posterior knowledge, and the statistical model primarily governs the inference strategy. Conversely, if the prior is sharply peaked, prior information dominates the estimation process and significantly shapes the posterior distribution.

Using Bayes theorem, equation (2) can be expressed in terms of the posterior distribution according to

$$(\Delta\phi)^2 = \sum_x P(x) \int_{-\infty}^{+\infty} d\phi P(\phi|x) [\phi - \phi_{\text{est}}(x)]^2. \quad (4)$$

In the remainder of this work, we estimate the accumulated phase within the framework of Bayesian estimation theory and quantify the phase estimation uncertainty based on this cost function. Furthermore, we review corresponding Bayesian bounds, to which we compare the performance of the investigated Ramsey protocols. As a consequence, we denote equation (4)—and thus likewise equation (2) due to its equivalence—as the BMSE throughout this work, following the literature such as [73–76, 95, 96].

For the primary investigations in this work, we assume a unitary phase evolution through the quantum channel

$$\Lambda_{\phi,T}[\rho_{\text{in}}] = \mathcal{R}_z(\phi) \rho_{\text{in}} \mathcal{R}_z^\dagger(\phi) \quad (5)$$

with rotation $\mathcal{R}_z(\phi) = e^{-i\phi S_z}$, where $S_{x,y,z}$ denote the collective spin operators of N two level systems. Consequently, the quantum channel, and thus the statistical model $P(x|\phi)$, is 2π -periodic with respect to the phase, i.e. $\Lambda_{\phi,T} = \Lambda_{\phi+2\pi,T}$. In this case, it is common to use a periodic cost function. However, in the context of atomic clocks, we explicitly adopt a global definition of the phase spanning $-\infty < \phi < \infty$,

since $\phi + 2\pi k$ (with $k \in \mathbb{Z}$) originates from a different frequency deviation ω than ϕ , and thus has a distinct physical interpretation. This distinction proves particularly useful to quantify the CTL of the LO (cf section 2.3) and to discuss fringe hops within this framework (cf section 4). Furthermore, we assume a Gaussian prior distribution

$$\mathcal{P}(\phi) = \frac{1}{\sqrt{2\pi(\delta\phi)^2}} \exp\left(-\frac{\phi^2}{2(\delta\phi)^2}\right) \quad (6)$$

with zero mean and width $\delta\phi$, which is a reasonable approximation for the full feedback loop of an atomic clock [67].

2.3. Bounds

The goal of Bayesian estimation is to minimize the cost function, the BMSE. For a given prior distribution $\mathcal{P}(\phi)$, there are three degrees of freedom to optimize: the initial state ρ_{in} , the measurement $\{\Pi_x\}$, and the estimation strategy $\phi_{\text{est}}(x)$. Based on these degrees of freedom and building on [70, 97–99], we collect a hierarchy of lower bounds for the BMSE (see appendix A.1 for detailed proofs), analogous to the local estimation approach. The discussion in this section remains general, allowing for arbitrary prior distributions $\mathcal{P}(\phi)$ and quantum channels $\Lambda_{\phi,T}$. Specific assumptions and asymptotic results will be explicitly noted.

BCRB—For a given initial state ρ_{in} and measurement $\{\Pi_x\}$, the BCRB $(\Delta\phi_{\text{BCRB}})^2$ represents a lower bound on the BMSE $(\Delta\phi)^2$ and thus, implicates an optimization over all possible estimators ϕ_{est} . Assuming standard regularity conditions $\sum_x \partial_\phi P(x|\phi) = 0$ and vanishing of the prior at the boundaries $\lim_{\phi \rightarrow \pm\infty} \mathcal{P}(\phi) = 0$, the BCRB results from the van Trees inequality [97] and reads [100]

$$(\Delta\phi)^2 \geq (\Delta\phi_{\text{BCRB}})^2 = \min_{\phi_{\text{est}}} (\Delta\phi)^2 = \frac{1}{\overline{\mathcal{F}} + \mathcal{I}} \quad (7)$$

(proof in appendix A.1.1). Here, the measurement contribution is represented by the Fisher information averaged over the prior distribution

$$\overline{\mathcal{F}} = \int d\phi \mathcal{P}(\phi) \mathcal{F}(\phi) = \int d\phi \mathcal{P}(\phi) \sum_x \frac{1}{P(x|\phi)} \left(\frac{dP(x|\phi)}{d\phi} \right)^2 \quad (8)$$

and

$$\mathcal{I} = \int d\phi \frac{1}{\mathcal{P}(\phi)} \left(\frac{d\mathcal{P}(\phi)}{d\phi} \right)^2 \quad (9)$$

denotes the information contained in the prior knowledge, given by the Fisher information of the prior distribution. While $\mathcal{I} \geq 0$, the average Fisher information $\overline{\mathcal{F}}$ is upper bounded by its maximal value $\mathcal{F}_{\text{max}} = \mathcal{F}(\phi_0)$ achieved at the optimal working point ϕ_0 , i.e. $\overline{\mathcal{F}} \leq \mathcal{F}_{\text{max}}$. Hence, from equation (7) it is evident that the BCRB in turn is lower bounded by the CRB, the corresponding bound in local phase estimation,

$(\Delta\phi_{\text{BCRB}})^2 \geq (\Delta\phi_{\text{CRB}})^2 = 1/\mathcal{F}_{\text{max}}$. In contrast to the local approach, the optimal estimation strategy in the Bayesian framework can be derived explicitly as we will show in section 2.4.

For a Gaussian prior distribution, the prior information simplifies to $\mathcal{I} = (\delta\phi)^{-2}$. Moreover, while $\overline{\mathcal{F}}$ typically increases with the ensemble size, the prior information \mathcal{I} is independent of N . Consequently, in the asymptotic limit of large N , the prior knowledge primarily contributes in the averaging of the Fisher information and we obtain $(\Delta\phi_{\text{BCRB}})^2 \simeq \overline{\mathcal{F}}^{-1}$.

BQCRB—The BQCRB extends the classical BCRB by including the optimization over all measurements $\{\Pi_x\}$. For a given initial state ρ_{in} , the BQCRB

$$(\Delta\phi_{\text{BQCRB}})^2 = \min_{\{\Pi_x\}} (\Delta\phi_{\text{BCRB}})^2 = \min_{\{\Pi_x\}, \phi_{\text{est}}} (\Delta\phi)^2 \quad (10)$$

provides a lower bound on the BCRB and thus establishes the hierarchy

$$(\Delta\phi)^2 \geq (\Delta\phi_{\text{BCRB}})^2 \geq (\Delta\phi_{\text{BQCRB}})^2. \quad (11)$$

Naively, one might suggest to simply replace the average Fisher information $\overline{\mathcal{F}}$ in equation (7) by the average QFI $\overline{\mathcal{F}}_Q = \int d\phi \mathcal{F}_Q(\Lambda_{\phi, T}[\rho_{\text{in}}])$. However, in general, the optimal measurement depends on ϕ and thus, this approach would effectively correspond to averaging over a set of measurements, each optimized for a particular phase value ϕ . By restricting the measurements, without loss of optimality, to the class of projection-valued measures (PVMs) $\Pi_x = |x\rangle\langle x|$, with orthonormal eigenstates $|x\rangle$, $\langle x|x'\rangle = \delta_{x,x'}$, of the observable X with eigenvalue x , the BQCRB can be expressed as [72]

$$(\Delta\phi_{\text{BQCRB}})^2 = (\delta\phi)^2 - \text{Tr}(\overline{\rho}L^2) \quad (12)$$

(proof in appendix A.1.2). Here, the double minimization over the measurement $\{\Pi_x\}$ and estimator ϕ_{est} is combined in a single quantity $L = \sum_x \Pi_x \phi_{\text{est}}(x)$. The optimal L is determined by the implicit equation

$$\overline{\rho}' = \frac{1}{2}(\overline{\rho}L + L\overline{\rho}) \quad (13)$$

where $\overline{\rho} = \int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}[\rho_{\text{in}}]$ denotes the average state and $\overline{\rho}' = \int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}[\rho_{\text{in}}] \phi$. The corresponding optimal measurement basis and estimator are given by the eigenbasis and eigenvalues of the operator L , respectively.

Interestingly, equations (12) and (13) have a similar structure as the QFI in local phase estimation. Indeed, assuming a unitary phase evolution according to equation (5) and a Gaussian prior distribution as defined in equation (6), the BQCRB can be related to the QFI $\mathcal{F}_Q[\overline{\rho}]$ of the average state $\overline{\rho}$ by [99]

$$(\Delta\phi_{\text{BQCRB}})^2 = (\delta\phi)^2 \left[1 - (\delta\phi)^2 \mathcal{F}_Q[\overline{\rho}] \right] \quad (14)$$

(derivation in appendix A.1.2). In this case, the optimal measurement corresponds to the SLD of the QFI approach associated with $\mathcal{F}_Q[\overline{\rho}]$, and the optimal Bayesian estimator can be

determined explicitly (cf section 2.4.). Evaluating the BQCRB thus becomes computationally equivalent to calculating the QFI of the average state $\overline{\rho}$. Indeed, the relation between the BMSE and the QFI in equation (14) is an interesting mathematical coincidence that is worth pointing out. Although this relation is only valid in the special case of single-parameter estimation and Gaussian prior distributions, it connects two conceptually different problems: On the one hand, global phase estimation, where the parameter range is characterized by the prior distribution, and on the other hand, local phase estimation based on a probe state that has been averaged over the prior distribution.

OQI—The OQI represents the ultimate lower bound of the BMSE, completing the hierarchy

$$(\Delta\phi)^2 \geq (\Delta\phi_{\text{BCRB}})^2 \geq (\Delta\phi_{\text{BQCRB}})^2 \geq (\Delta\phi_{\text{OQI}})^2. \quad (15)$$

The OQI simultaneously optimizes over all three degrees of freedom: the initial state ρ_{in} , measurement $\{\Pi_x\}$ and estimator ϕ_{est} .

$$\begin{aligned} (\Delta\phi_{\text{OQI}})^2 &= \min_{\rho_{\text{in}}} (\Delta\phi_{\text{BQCRB}})^2 = \min_{\rho_{\text{in}}, \{\Pi_x\}} (\Delta\phi_{\text{BCRB}})^2 \\ &= \min_{\rho_{\text{in}}, \{\Pi_x\}, \phi_{\text{est}}} (\Delta\phi)^2. \end{aligned} \quad (16)$$

Unfortunately, no general expressions for the OQI sensitivity for arbitrary ensemble sizes are available, instead, they require complex optimization procedures. An algorithm presented in [72, 98] iteratively optimizes the initial probe state and the measurement (cf appendix A.1.3) and enables an efficient computation for ensembles with up to several hundreds of particles. However, numerical optimization becomes challenging as the ensemble size $N \gg 1$ increases and approaches the asymptotic limit.

Considering a 2π -periodic quantum channel with respect to the phase ϕ (cf equation (5)), the OQI allows for unambiguous phase estimation within the range $[-\pi, +\pi]$. However, for sufficiently broad prior distributions, the phase ϕ may exceed this invertible regime and an estimation error of $(2\pi)^2$ is accumulated, associated with transitions between adjacent Ramsey fringes. Although the Bayesian framework naturally accounts for this crossover, it is nevertheless instructive to examine their contribution separately. For a Gaussian prior distribution, the estimation error associated with these events can be modeled by (cf appendix A.1.3)

$$(\Delta\phi_{\text{OQI}}^{\text{CTL}})^2 = 4\pi^2 \left[1 - \text{erf} \left(\frac{\pi}{\sqrt{2}\delta\phi} \right) \right] \quad (17)$$

where $\text{erf}(z)$ denotes the error function. In the context of an atomic clock, in this regime of long interrogation times, the coherence time of the LO will become relevant and ultimately limits the clock stability. Consequently, we will denote equation (17) as the CTL of the OQI.

In the asymptotic limit ($N \gg 1$), assuming unitary phase evolution as described by equation (5) and restricting to the invertible range $[-\pi, +\pi]$, it has been shown for arbitrary prior

distributions [95, 99, 101, 102] that the ultimate lower bound is given by

$$(\Delta\phi_{\pi\text{HL}})^2 = \frac{\pi^2}{N^2}. \quad (18)$$

In the absence of decoherence, this asymptotic limit reflects Heisenberg scaling with an additional factor of π , and is therefore referred to as the πHL . Intuitively, the πHL can be interpreted as the maximal estimation error associated with estimating a phase within $[-\pi, +\pi]$ using $N + 1$ evenly spaced measurement outcomes. Additionally taking into account the estimation error outside of the invertible range, as modeled by equation (17), the overall asymptotic estimation error for the OQI reads

$$\begin{aligned} (\Delta\phi_{\text{OQI}}^{\text{asym}})^2 &= (\Delta\phi_{\pi\text{HL}})^2 + (\Delta\phi_{\text{OQI}}^{\text{CTL}})^2 \\ &= \frac{\pi^2}{N^2} + 4\pi^2 \left[1 - \text{erf} \left(\frac{\pi}{\sqrt{2}\delta\phi} \right) \right]. \end{aligned} \quad (19)$$

This result combines the fundamental limit set by the πHL with the contributions from phase estimation errors associated with transitions between Ramsey fringes, offering a comprehensive characterization of the OQI's performance in the asymptotic regime. Notably, this bound can be saturated asymptotically by the POI [70, 73, 101–107] (cf appendix A.1.3).

2.4. Estimators

Based on a statistical model $P(x|\phi)$, defined by an initial state ρ_{in} , free evolution $\Lambda_{\phi,T}$ and measurement $\{\Pi_x\}$, various estimation strategies can be applied. In this work, we focus on two such strategies: the linear estimator and the optimal Bayesian estimator. The linear estimator is renowned for its simplicity and is both theoretically and experimentally commonly used and well understood. It often arises naturally in a local approach (i.e. for a narrow prior distribution and linear error propagation), where it is the standard choice. In contrast, the optimal Bayesian estimator, as the name suggests, achieves the best possible performance in Bayesian phase estimation.

Linear estimator—The linear estimator

$$\phi_{\text{est}}^{\text{linear}}(x) = a \cdot x, \quad (20)$$

with scaling factor $a \in \mathbb{R}$, originates from local phase estimation. In this context, assuming an unbiased estimator and small deviations from the optimal working point ϕ_0 , the signal can be approximated linearly. In this case, the estimation error is typically characterized by QPN [1]

$$\Delta\phi_{\text{QPN}} = \frac{\xi}{\sqrt{N}} = \frac{\Delta X(\phi_0)}{|\partial_\phi \langle X(\phi) \rangle|_{\phi=\phi_0}}, \quad (21)$$

where ξ denotes the Wineland squeezing parameter [53, 54]. This local result is obtained in the limit of narrow prior distributions ($\delta\phi \rightarrow 0$) around the optimal working point ϕ_0 and by

choosing the particular scaling factor $a = (\partial_\phi \langle X(\phi) \rangle|_{\phi=\phi_0})^{-1}$, corresponding to the inverse slope of the signal at ϕ_0 .

In the Bayesian framework, however, this approach is poorly suited. First, the assumption of narrow prior distributions fails for realistic fluctuations of the phase, as discussed above. Second, the prior information explicitly influences the cost function and thus, the scaling factor a has to depend on the prior distribution. For an arbitrary prior distribution with zero mean $\int d\phi \mathcal{P}(\phi)\phi = 0$ and variance $(\delta\phi)^2 = \int d\phi \mathcal{P}(\phi)\phi^2$, the optimal scaling factor and corresponding BMSE are given by (cf appendix A.2.1)

$$a = \frac{\int d\phi \mathcal{P}(\phi)\phi \langle X(\phi) \rangle}{\int d\phi \mathcal{P}(\phi)\langle X^2(\phi) \rangle} \quad (22)$$

$$(\Delta\phi)^2 = (\delta\phi)^2 - \frac{[\int d\phi \mathcal{P}(\phi)\phi \langle X(\phi) \rangle]^2}{\int d\phi \mathcal{P}(\phi)\langle X^2(\phi) \rangle}. \quad (23)$$

As in the local approach, the linear estimator and its estimation error depend only on the first and second moments of the observable X , which typically are easier to evaluate than the full statistical model $P(x|\phi)$. This simplicity makes the linear estimator a practical choice for phase estimation. Nevertheless, despite its advantages and reliable performance in several situations, the linear estimation strategy in general does not saturate the BCRB and thus is not optimal.

Optimal Bayesian estimator—In contrast to local phase estimation, where the CRB can in general only be approximated in the infinite-sample limit using the maximum-likelihood estimator [1], the optimal estimator in Bayesian phase estimation can be derived explicitly [70]

$$\phi_{\text{est}}^{\text{opt}}(x) = \int d\phi \mathcal{P}(\phi|x)\phi \quad (24)$$

(derivation in appendix A.2.2), saturating the BCRB with single shot measurements. This estimator corresponds to the average phase with respect to the posterior distribution $P(\phi|x)$, which can be expressed in terms of the statistical model $P(x|\phi)$ and prior distribution $\mathcal{P}(\phi)$ according to Bayes theorem equation (3). As a consequence, the optimal Bayesian estimator can be highly non-linear. The associated BMSE is given by

$$(\Delta\phi)^2 = (\delta\phi)^2 - \sum_x \frac{[\int d\phi \mathcal{P}(\phi)P(x|\phi)\phi]^2}{P(x)}. \quad (25)$$

Although this resembles the structure of equation (23), the BMSE for the optimal Bayesian estimator explicitly depends on the statistical model, rather than merely on the first and second moments of the observable. Additionally, for the optimal Bayesian estimator, equation (4) reduces to the average posterior variance. Since the optimal Bayesian estimator saturates the BCRB, and thus minimizes the BMSE with respect to all estimation strategies, it is commonly referred to as the minimal MSE estimator. However, we continue to use the term ‘optimal Bayesian estimator’ throughout this work for consistency and clarity.

2.5. ADEV

The long-term stability of an atomic clock is quantified by the ADEV $\sigma_y(\tau)$ [108, 109] (cf appendix A.3 for a brief introductory overview), characterizing the fluctuations of fractional frequency deviations $y(t) = \omega(t)/\omega_0$ averaged over $\tau \gg T_C = T + T_D$. Here, the total cycle duration T_C accounts for the interrogation time T and any potential dead time T_D , arising from preparation steps and application of the feedback. In local frequency metrology, assuming short interrogation times leading to narrow prior distributions, the ADEV is well approximated by [1, 7]

$$\sigma_y(\tau) = \frac{1}{\omega_0} \frac{\Delta\phi_{\text{QPN}}}{T} \sqrt{\frac{T_C}{\tau}} = \frac{1}{\omega_0} \frac{\xi}{\sqrt{N}} \frac{1}{T} \sqrt{\frac{T_C}{\tau}}. \quad (26)$$

In this context, clock stability is determined by QPN $\Delta\phi_{\text{QPN}} = \xi/\sqrt{N}$, characterizing the uncertainty associated with the measurement process.

However, the BMSE in Bayesian frequency metrology—which leverages Bayesian phase estimation strategies tailored to frequency metrology and directly includes the frequency fluctuations into the theoretical model—incorporates both measurement uncertainty and prior knowledge, preventing a straightforward substitution of $\Delta\phi_{\text{QPN}}$ by $\Delta\phi$. To isolate the measurement contribution from the prior knowledge \mathcal{I} , we introduce the effective measurement uncertainty motivated by the BCRB in equation (7) and following [67, 73]

$$\Delta\phi_M = \left(\frac{1}{(\Delta\phi)^2} - \mathcal{I} \right)^{-1/2} = \left(\frac{1}{(\Delta\phi)^2} - \frac{1}{(\delta\phi)^2} \right)^{-1/2} \quad (27)$$

where $\mathcal{I} = (\delta\phi)^{-2}$ for a Gaussian distribution. Hence, $\Delta\phi_M$ quantifies the quality of the measurement process in a single interrogation cycle. According to the discussion of the BCRB in section 2.3, the effective measurement uncertainty is lower bounded by the average Fisher information $(\Delta\phi_M)^2 \geq 1/\overline{\mathcal{F}}$ and thus, a connection to the local approach can be established, yielding

$$(\Delta\phi_M)^2 \geq \frac{1}{\overline{\mathcal{F}}} \geq (\Delta\phi_{\text{CRB}})^2 = \frac{1}{\mathcal{F}_{\text{max}}} = \frac{1}{\mathcal{F}(\phi_0)}. \quad (28)$$

As a consequence, the clock stability in local frequency metrology, quantified by equation (26), emerges in the limit of narrow prior distributions ($\delta\phi \ll 1$) or equivalently short interrogation times ($T \ll 1$).

With the effective measurement uncertainty, the ADEV in the framework of Bayesian frequency metrology is given by

$$\sigma_y(\tau) = \frac{1}{\omega_0} \frac{\Delta\phi_M}{T} \sqrt{\frac{T_C}{\tau}}. \quad (29)$$

Consequently, the three key quantities for quantifying the sensitivity and ultimately the clock stability in Bayesian frequency metrology are the prior width $\delta\phi$, the BMSE $\Delta\phi$, and

the effective measurement uncertainty $\Delta\phi_M$. In the following, we will examine their relation in a qualitative discussion.

Due to noise in the LO, the phase diffusion grows with Ramsey dark time. Thus, the prior width $\delta\phi$ of the relative phase will be monotonically increasing with the interrogation time T (cf figure 1(b)). At first glance, one might assume that $\delta\phi$ is solely determined by the characteristics of the free running LO frequency $\omega_{\text{LO}}(t)$. However, it depends even more strongly on the noise of the stabilized frequency $\omega(t)$ and therefore on the details of interrogation, estimation and feedback. Although, in general, the prior phase distribution can vary over different clock cycles, it becomes stationary if the feedback loop stabilizes the LO reliably to the atomic reference. In this case, the residual noise can be considered to be white, to a good approximation, and thus can be modeled by a normal distribution characterized by the spread $\delta\phi$ (cf equation (6)).

For a given finite $\delta\phi$, the interrogation protocol and estimation strategy can be optimized to minimize the estimation error $\Delta\phi$. At the same time, the effective measurement uncertainty, equation (27), and thus also the ADEV, equation (29), are minimized. Consequently, $\Delta\phi$ will ultimately determine the stabilized frequency noise, which in turn affects $\delta\phi$. Therefore, in order to reflect the closed feedback loop of the atomic clock, $\Delta\phi$ has to be optimized iteratively for suitably chosen $\delta\phi$, as detailed in section 2.6.

The average error in phase estimation $\Delta\phi$ depends on the prior width $\delta\phi$ as well as the particular interrogation sequence and estimation strategy. As discussed before, $\Delta\phi \leq \delta\phi$ and thus the estimation error $\Delta\phi$ is reduced compared to the prior width $\delta\phi$, since a proper Ramsey protocol increases the information about the phase. Here, equality $\Delta\phi = \delta\phi$ corresponds a worst case scenario in which the effective measurement variance diverges $\Delta\phi_M \rightarrow \infty$. This case represents an ineffective interrogation scheme, where the information gained through measurement and estimation fails to improve the characterization of residual noise. Conversely, a hypothetical perfect phase estimation (precluded by quantum mechanics due to its intrinsic indeterminism) would result in a vanishing estimation error $\Delta\phi \rightarrow 0$. Likewise, the effective measurement uncertainty would also vanish $\Delta\phi_M \rightarrow 0$, since this scenario implies a perfect measurement.

The form of the ADEV in equation (29) suggests that the stability can be improved by increasing the interrogation time T . However, this is only true as long as the CTL of the LO remains negligible and QPN $\Delta\phi_{\text{QPN}} = \frac{\xi}{\sqrt{N}}$ of the measurement dominates the effective measurement uncertainty. In general, three regimes can be distinguished based on the relation between the prior width $\delta\phi$ and the QPN $\Delta\phi_{\text{QPN}}$:

- (i) Considering small prior widths $\delta\phi \ll 1$, the measurement and estimation protocol cannot significantly improve the knowledge of the phase distribution, since $\Delta\phi_{\text{QPN}} \gg \delta\phi$ and thus $\Delta\phi \simeq \delta\phi$. In this case, $\Delta\phi_M \simeq \Delta\phi_{\text{QPN}}$ and the local form of the ADEV, equation (26), is reproduced.
- (ii) With increasing interrogation time T , the prior width surpasses QPN $\delta\phi > \Delta\phi_{\text{QPN}}$. Nevertheless, in this regime,

the information gain on the phase distribution resulting from the measurement and estimation strategy leads to $\Delta\phi \ll \delta\phi$ and thus $\Delta\phi_M < \delta\phi$. Hence, the optimal working point of the atomic clock is located in this region.

- (iii) At long interrogation times, the coherence time of the LO will become relevant and ultimately limits the clock stability. Here, the phase noise exceeds the domain of the measurement scheme where an unambiguous estimation is possible, giving $\Delta\phi_M \gg \delta\phi \simeq \Delta\phi \gg \Delta\phi_{\text{QPN}}$.

Consequently, the ADEV features a trade-off between increased stability achieved through long interrogation times and the limitations imposed by the coherence time of the LO, which are characterized by the CTL. Fortunately, as previously discussed, this trade-off is inherently addressed within the framework of Bayesian frequency metrology. In terms of clock stability, this trade-off gives rise to an optimal interrogation time T_{\min} at which the minimal ADEV σ_{\min} is achieved. While σ_{\min} accounts for the bias in phase estimation for phases beyond the invertible range of the Ramsey sequence (via the CTL), it nevertheless remains restricted to a single clock cycle and neglects cumulative effects that might arise in a full feedback loop. The most prominent of these effects are fringe hops, which are discussed in detail in section 4 as they can only appear within the full feedback loop.

2.6. Interrogation time and prior width

In the previous section, we linked the clock stability at interrogation time T , characterized by the ADEV, with Bayesian phase estimation with prior width $\delta\phi$, described by the BMSE. Furthermore, we qualitatively discussed that the prior width increases with the interrogation time. To complete the connection between Bayesian phase estimation and frequency metrology, this section aims to establish an explicit relation between $\delta\phi$ and T . This relation serves as a bridge between the frequency fluctuations of the laser in an experiment and a theoretical representation in terms of a Gaussian prior distribution with a specific width. Establishing this connection is essential for accurately modeling experiments and ensuring the applicability of theoretical predictions to realistic scenarios.

State-of-the-art clock lasers are characterized by the spectral noise density $S_y(f) = \sum_{\alpha} h_{\alpha} f^{-(1+\alpha)}$, which can be modeled by a power law [7, 109–111], where $\alpha = -1, 0, 1$ corresponds to white, flicker and random walk frequency noise, respectively. Accordingly, the Allan variance of the free-running LO can be expressed as $\sigma_{y,\text{LO}}^2(\tau) = \tilde{h}_{\alpha} \tau^{\alpha}$. To compare different LOs, a single timescale is defined characterizing the stability. Following [67], we define the laser coherence time Z implicitly by

$$\sigma_{y,\text{LO}}(Z_C) 2\pi\nu_0 Z = 1 \text{ rad.} \quad (30)$$

Here, $\sigma_{y,\text{LO}}(Z_C)$ denotes the ADEV of the LO averaging over a single cycle duration $Z_C = Z + T_D$ with dead time T_D .

In [67], it was demonstrated that the prior width of the full feedback loop can be approximated by the power law

$$(\delta\phi)^2 \simeq \chi(\alpha) \left(\frac{T}{Z}\right)^{2+\alpha} \quad (31)$$

depending solely on the ratio of interrogation time T and coherence time of the LO Z , and the numerically determined factor $\chi(\alpha) = 1, 1.7, 2$ for $\alpha = -1, 0, 1$. This approximation was derived in the limit of large ensembles and long interrogation times using the conventional Ramsey protocol in the framework of local phase estimation, and successfully applied in [68, 73]. However, in the full feedback loop of an atomic clock, the prior width $\delta\phi$ and estimation error $\Delta\phi$ mutually influence each other. Therefore, $\delta\phi$ has to be adjusted iteratively to account for the closed feedback loop dynamics, as motivated in the previous section and detailed in appendix A.5.4. This iterative procedure is employed in section 4, where realistic Monte Carlo simulations of the full feedback loop of an atomic clock are performed. Nevertheless, equation (31) remains a convenient approximation for general investigations and is thus adopted in section 3.

In the following, motivated by state-of-the-art clock lasers [66], we assume a LO predominantly limited by flicker frequency noise. Additionally, we neglect systematic shifts in the atomic transition frequency ω_0 . Moreover, we will assume the atomic excited-state lifetime t_{exc} to be substantially longer than the clock cycle duration T_C , such that $t_{\text{exc}} \gg T_C$.

3. Quantum frequency metrology

In this section, we aim to saturate the OQI in the context of atomic clocks. We begin by analyzing standard protocols and compare their performance to the OQI. Afterwards, we introduce variational classes of quantum circuits and investigate the associated optimal Ramsey protocols.

3.1. Standard protocols

To start with, we examine the effective measurement uncertainty and corresponding clock stability of standard Ramsey protocols (see appendix A.4 for detailed derivations). Specifically, we focus on CSSs, SSSs, and GHZ states, as well as the ultimate lower bound defined by the OQI. For all Ramsey schemes, the dependence of clock stability on the interrogation time T reflects the three regimes discussed in section 2.5, representing a trade-off between enhanced stability achieved through longer interrogation times and the CTL. These distinct regimes are depicted in figure 2(a), which illustrates the generic dependence of stability on interrogation time. Furthermore, the scaling of the minimal ADEV σ_{\min} with ensemble size N at the optimal interrogation time T_{\min} is presented in figure 2(b). These figures are based on the work of Kaubrügger *et al* [73] and are adapted here within the framework defined above. In particular, emphasis is placed on comparing the linear and optimal Bayesian estimation

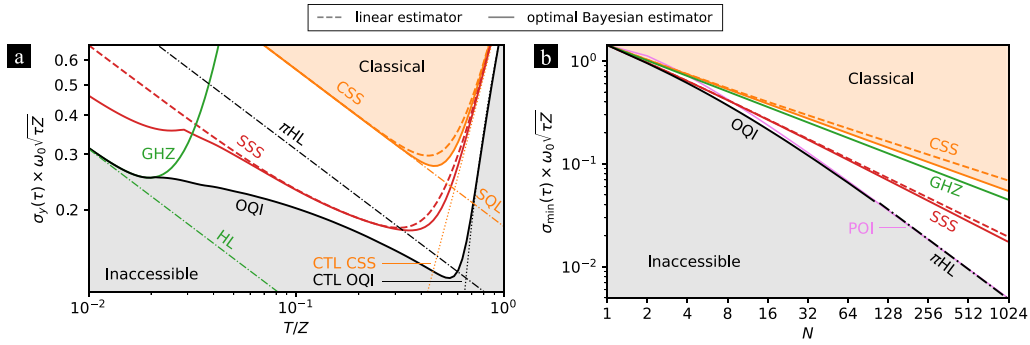


Figure 2. (a) Generic scaling of the dimensionless Allan deviation $\sigma_y(\tau) \times \omega_0 \sqrt{\tau Z}$ with the interrogation time T for the example of $N = 32$, rescaled by the averaging time τ , laser coherence time Z and transition frequency ω_0 . Stabilities for the CSS (orange), SSS (red) and GHZ (green) protocols are compared to the performance of the OQI (black). For CSS and SSS, both the linear (dashed) and optimal Bayesian estimator (solid) are depicted. The gray shaded area represents the inaccessible stability region set by the OQI limit, while the orange shaded area indicates achievable stabilities using uncorrelated atoms. Dotted lines correspond to the CTL for OQI (black) and for CSS and SSS with the linear estimator (orange). Additionally, benchmarks such as the SQL (green), HL (black) and π HL (black) are included as dashed-dotted lines. (b) Scaling of the dimensionless minimal Allan deviation $\sigma_{\min} \times \omega_0 \sqrt{\tau Z}$ with the ensemble size N . In addition to the standard protocols, the POI performance (violet) is presented. For the OQI and POI, numerical optimization is performed for $N \leq 100$, while the asymptotic behavior, represented by the π HL (black dashed-dotted), is shown for $N > 100$.

strategies across different protocols, with performance benchmarked against the OQI. To enable comparability between various setups, the achievable ADEVs $\sigma_y(\tau)$ are rescaled with respect to the atomic transition frequency ω_0 , the total averaging time τ and the laser coherence time Z . This rescaling ensures that the results are transferable to specific experimental parameters.

CSSs—The conventional clock protocol employs Ramsey interferometry with CSSs [112–114] as initial states, a collective projective spin measurement and a linear estimation strategy. In this scenario, the effective measurement uncertainty can be evaluated analytically [67] as

$$(\Delta\phi_M^{\text{CSS}})^2 = \frac{\cosh((\delta\phi)^2)}{N} + \sinh((\delta\phi)^2) - (\delta\phi)^2. \quad (32)$$

For short interrogation times $T/Z \ll 1$, leading to narrow prior widths $\delta\phi \ll 1$, the conventional SQL $\Delta\phi_{\text{SQL}} = 1/\sqrt{N}$ is recovered. Conversely, for long interrogation times $T/Z \sim 1$, frequency fluctuations of the LO dominate and the first term in equation (32) becomes negligible. This regime defines the CTL for CSS with a linear estimator,

$$(\Delta\phi_{\text{CTL}}^{\text{CSS}})^2 = \sinh((\delta\phi)^2) - (\delta\phi)^2. \quad (33)$$

Hence, the stability reflects a trade-off between these two regimes, as illustrated in figure 2(a), determining the minimal ADEV σ_{\min} . As the ensemble size N increases, the first term in equation (32) decreases, leading to shorter optimal interrogation times T_{\min} to achieve σ_{\min} .

For the optimal Bayesian estimator, an explicit evaluation of the conditional probabilities $P(x|\phi)$ is required, as discussed in section 2.4. Although $P(x|\phi)$ can be determined analytically for the CSS, the integrals in equation (25) generally have to be evaluated numerically. For short interrogation times, the narrow prior phase distribution allows for a good approximation by linearizing the signal. Thus, the optimal Bayesian

estimator reproduces the linear estimator in this regime. In contrast, for interrogation times in the region of the minimal ADEV, higher-order contributions of the sinusoidal signal become relevant and the curvature of the signal has to be considered. In this case, the optimal Bayesian estimator approximates the *arcsin* estimator, which directly inverts the signal and thus allows to estimate the phase unambiguously in the range $[-\pi/2, +\pi/2]$. This results in an extended dynamic range compared to the linear estimator, which cannot account for any non-linearity of the signal and thus exhibits a higher minimal instability. As a consequence, the optimal Bayesian estimator improves the scaling of $\Delta\phi_M$ with the ensemble size N to $\mathcal{O}(N^{-0.47})$, compared to $\mathcal{O}(N^{-0.42})$ for the linear estimator, as shown in figure 2(b). While the choice of estimator has limited impact for small ensembles, the stability gain from the optimal Bayesian estimator becomes significant for large ensembles $N \gg 1$. Importantly, this improvement arises solely from classical post-processing of the measurement outcomes, while the quantum circuit remains unchanged. Nevertheless, the CTL prevents both estimation strategies from achieving the SQL of $1/\sqrt{N}$.

SSSs—Extending the conventional Ramsey protocol with a single OAT interaction [87] for state preparation, various entangled states can be generated. Here, OAT interactions are denoted by $\mathcal{T}_{\mathbf{k}}(\mu) = \exp(-i\frac{\mu}{2}S_{\mathbf{k}}^2)$ with twisting strength μ around axis \mathbf{k} , where $S_{\mathbf{k}} = k_1S_x + k_2S_y + k_3S_z$ is the spin projection along direction \mathbf{k} . In particular, for small twisting strengths μ , OAT generates SSSs by shearing the initial CSS around the twisting axis, characterized by a squeezing parameter $\xi < 1$. Using the linear estimator, the effective measurement uncertainty is given by

$$(\Delta\phi_M^{\text{SSS}})^2 = \frac{\langle S_y^2 \rangle}{\langle S_x \rangle^2} \cosh((\delta\phi)^2) + \frac{\langle S_x^2 \rangle}{\langle S_x \rangle^2} \sinh((\delta\phi)^2) - (\delta\phi)^2, \quad (34)$$

with expectation values provided in appendix A.4.2. SSS show enhanced stability compared to CSS due to reduced fluctuations in the measured spin observable. However, the gain comes at the cost of smaller dynamic range, as the minimal ADEV is achieved at shorter interrogation times compared to CSS (cf figure 2(a)). This is a direct consequence of SSS sharing the same CTL as CSS, since $\langle S_y^2 \rangle / \langle S_x \rangle^2 \ll 1$ and $\langle S_x^2 \rangle / \langle S_x \rangle^2 \simeq 1$ for large prior widths $\delta\phi$ and optimal twisting strength. Similar to the conventional Ramsey protocol, SSS with the optimal Bayesian estimator achieve a slightly extended dynamic range at long interrogation times. For large ensembles $N \gg 1$, the asymptotic scaling of the effective measurement uncertainty with the optimal Bayesian estimator approximates $\mathcal{O}(N^{-2/3})$, depicted in figure 2(b), reflecting the scaling observed in local phase estimation [1, 87]. In contrast, the linear estimator exhibits a scaling $\mathcal{O}(N^{-0.63})$. Furthermore, the optimal Bayesian estimation strategy offers a remarkable advantage at short interrogation times, where $\delta\phi \lesssim 1/N$, as shown in figure 2(a). In this regime, the estimator becomes highly non-linear, allowing for substantially stronger twisting strengths μ , resulting in stronger squeezing and enhanced stability. However, as $\delta\phi$ approaches $1/N$, the scaling of the ADEV with the interrogation time T stagnates and converges towards the stability achieved with the linear estimator.

GHZ states—The maximally entangled GHZ state $|\text{GHZ}\rangle = [|\downarrow\rangle^{\otimes N} + |\uparrow\rangle^{\otimes N}] / \sqrt{2}$ [115] represents an equal superposition of the collective ground and excited states. The corresponding Ramsey sequence, initially proposed by Bollinger *et al* [116], is referred to as the GHZ protocol. During the free evolution time, the accumulated phase is amplified by a factor of N due to the maximal entanglement of the GHZ state. Subsequently, the parity Π is measured, resulting in a binary outcome ± 1 that indicates whether the number of atoms in the ground state is even or odd. Since a binary outcome inevitably results in a linear estimator, both estimation strategies coincide and result in the effective measurement uncertainty

$$(\Delta\phi_M^{\text{GHZ}})^2 = \frac{e^{N^2(\delta\phi)^2}}{N^2} - (\delta\phi)^2. \quad (35)$$

However, the optimal Bayesian estimator allows to avoid a parity measurement and perform a conventional projective spin measurement instead. In this case, the optimal Bayesian estimator effectively maps even and odd numbers of atoms in the ground state to the parity ± 1 , thereby mimicking a parity measurement and achieving the same sensitivity (cf appendix A.4.3). This strategy was essentially employed in a different framework in [117]. Both measurement and estimation strategies are optimal, since equation (35) aligns with the BQCRB. For short interrogation times, where $\delta\phi \lesssim 1/N$, the GHZ protocol achieves the conventional HL $\Delta\phi_{\text{HL}} = 1/N$, as illustrated in figure 2(a), which corresponds to the OQI in a decoherence-free local phase estimation scenario. However, the sensitivity of the GHZ protocol decreases N -times faster than that of CSS as the prior width increases. For a parity

measurement, this is attributed to the N -times increased oscillation frequency of the sinusoidal signal, yielding a reduced dynamic range. Ultimately, the resulting ambiguities in phase estimation cause the GHZ protocol to be effectively insensitive to phases $\phi \gtrsim 1/N$. Consequently, the optimal interrogation time scales approximately as $1/(NZ)$, leading to a scaling of the effective measurement uncertainty of $\mathcal{O}(N^{-1/2})$, equivalent to the SQL (cf figure 2(b)). While the minimal ADEV of the GHZ protocol provides only a minor improvement over CSS and is outperformed by SSS, its shorter optimal interrogation time offers practical advantages. For instance, reduced probe times mitigate contrast losses and time dilation shifts caused by motional heating in ion crystals, thus improving the signal-to-noise ratio and the accuracy of such a clock [7].

OQI—As discussed in section 2.3, the OQI requires numerical optimization, as no analytical expressions are available for arbitrary ensemble sizes. Instead, we investigate the general scaling based on figure 2. For short interrogation times $T/Z \ll 1$, where $\delta\phi \lesssim 1/N$, the OQI is saturated by the GHZ protocol, achieving the HL $\sigma_{\text{HL}}(\tau) = 1/(N\sqrt{T\tau})$, as shown in figure 2(a). As the interrogation time increases and $\delta\phi \gtrsim N$, a characteristic *plateau* emerges in which the ADEV decreases only marginally with T . This plateau shifts to shorter interrogation times as the ensemble size N increases, reflecting the CTL of the GHZ protocol. Beyond this plateau, as the interrogation time increases further, the scaling of the ADEV with T converges back to $1/\sqrt{T}$, ultimately reaching its minimum σ_{min} at T_{min} . In the limit of large ensembles ($N \gg 1$), this minimum is determined by the π HL. While the OQI significantly outperforms SSS at T_{min} , SSS perform close to the OQI in the transition regime, especially for small ensembles. The relative gain of the OQI over SSS at T_{min} increases with the ensemble size, as illustrated in figure 2(b). In the asymptotic limit $N \gg 1$, the POI, introduced in section 2.3 and further detailed in appendix A.1.3, is optimal, saturating the π HL. In this regime, the OQI scales as $\mathcal{O}(N^{-0.97})$, closely approximating Heisenberg scaling. Ultimately, the CTL is approached at long interrogation times.

In the following, we essentially distinguish between two different regimes concerning the ensemble size N : the first regime covers systems ranging from $N = 1$ to some tens of atoms, as relevant for ion-traps [11–13] or tweezer-arrays [14–17]. The remainder of this section primarily focuses on bridging the gap between SSS and the OQI by identifying Ramsey protocols of increasing complexity that approximate the OQI within this regime. In contrast, for large ensembles ($N \gtrsim 100$), the regime of lattice clocks is reached and the asymptotic scaling is approximated [7, 18–21]. In this regime, dead time typically emerges as the dominant limitation [68], as discussed in detail in section 5.

3.2. Variational classes

Recent advances in quantum information have inspired the development of variational quantum circuits as versatile tools

for implementing interferometers with setup-specific quantum gates. Typically, each layer in these circuits comprises an entanglement-generating interaction and (single qubit) rotations that provide geometric flexibility. OAT [87] interactions have gathered significant attention, as they can be implemented in several experimental platforms [57, 79, 88–94] and corresponding circuits represent a natural extension of SSSs. Combined with collective rotations, OAT interactions form the building blocks of several variational quantum circuits [73–78]. While [75] offers a unified framework for generalized echo protocols in local phase estimation, encompassing numerous previously documented approaches [77, 79–86], this work investigates variational classes specifically tailored to Bayesian frequency metrology.

In general, any variational Ramsey protocol can be expressed as

$$P(x|\phi) = \text{Tr}(|x\rangle\langle x|\Lambda_{\phi,T}[\rho_{\text{in}}]) \quad (36)$$

$$\rho_{\text{in}} = \mathcal{U}_{\text{prep}}|\psi_0\rangle\langle\psi_0|\mathcal{U}_{\text{prep}}^\dagger \quad (37)$$

$$|x\rangle\langle x| = \mathcal{U}_{\text{meas}}^\dagger|M\rangle\langle M|\mathcal{U}_{\text{meas}}, \quad (38)$$

with arbitrary unitary preparation and measurement operations $\mathcal{U}_{\text{prep}}$ and $\mathcal{U}_{\text{meas}}$, respectively. While $\mathcal{U}_{\text{prep}}$ generates the initial state by acting on the ground state $|\psi_0\rangle = |\downarrow\rangle^{\otimes N}$, $\mathcal{U}_{\text{meas}}$ effectively determines the measurement X by transforming the projectors $|M\rangle\langle M|$ associated with Dicke states of spin $S = N/2$ and eigenvalue M of S_z . Since any alternative choice of $|\psi_0\rangle$ and Dicke basis $\{|M\rangle\langle M|\}$ can be incorporated into $\mathcal{U}_{\text{prep}}$ and $\mathcal{U}_{\text{meas}}$ by additional transformations, fixing $|\psi_0\rangle$ and $\{|M\rangle\langle M|\}$ does not limit the generality of the protocol. The unitaries $\mathcal{U}_{\text{prep}}$ and $\mathcal{U}_{\text{meas}}$ are constructed from n and m layers of the variational circuit, respectively. Consequently, n effectively determines the level of entanglement in the initial state, while m governs the measurement strategy and dynamic range, ultimately determining the minimal ADEV σ_{min} .

Previous results—Pioneering work on Bayesian variational Ramsey protocols was conducted in [73, 76]. These studies introduced a variational class constrained to be invariant under the x -parity transformation, resulting in an anti-symmetric signal. Each layer of the quantum circuit consisted of two OAT interactions applied along orthogonal directions, combined with a collective rotation about one of these axes. While this choice provided a diverse class of entanglement-generating unitaries in each layer, it imposed significant geometric constraints. Nevertheless, the quality of phase estimation was not compromised, as the main objective was to saturate the OQI in the asymptotic limit of deep circuits. The analysis primarily focused on ensembles with several tens of qubits and employed linear estimation strategies. Kaubruegger *et al* demonstrated that the minimal ADEV σ_{min} could be achieved with sufficiently deep circuits. However, this approach had two key drawbacks: the reliance on deep circuits due to restricted geometric flexibility, and the inclusion of two OAT interactions per layer, which are experimentally more challenging to implement than collective rotations.

Thurtell and Miyake [74] addressed these limitations by proposing a variational class where each layer comprises a single OAT interaction around the z -axis combined with global rotations. These rotations are designed to effectively transform the OAT interaction with respect to an arbitrary axis, thereby eliminating geometric constraints. This approach reduced both the circuit depth and the number of OAT interactions, while achieving results comparable to those in [73]. Nevertheless, a considerable number of OAT interactions remained necessary. Moreover, the analysis was conducted within the framework of general Bayesian phase estimation and thus, did not consider the trade-off with respect to the interrogation time in frequency metrology.

Variational Ramsey protocols—In the following, we aim to approximate the OQI within the framework of Bayesian estimation tailored to frequency metrology. Instead of exploring the convergence towards the OQI with many layers for state preparation and measurement, we focus on variational quantum circuits with minimal depth. We primarily consider small ensembles relevant to ion-traps and additionally investigate the transition toward tweezer-arrays, which have been predominantly studied in [73, 74]. Given the high degree of control achievable in these systems, this represents the regime where near-term experimental implementation is most likely. Moreover, variational protocols are less favorable in setups with many atoms, such as in lattice clocks, as we will discuss below. In contrast to earlier studies, relying on linear estimation strategies, we employ the optimal Bayesian estimator to fully exploit the potential of variational Ramsey protocols. This choice is motivated by the substantial improvements observed for the standard protocols, including enhanced squeezing for SSS at short interrogation times, an extended dynamic range for CSS and SSS, and an effective reduction in the circuit depth required to implement the GHZ protocol. Additionally, by using the optimal Bayesian estimator, we ensure saturation of the BCRB for any given initial state and measurement and thus minimizing the required circuit depth. A comprehensive comparison with the linear estimation strategy is provided in section 4.3.

Building on the advancements in [73, 74], we define the variational class of generalized Ramsey protocols considered in this work, as illustrated in figure 3(a), by

$$\begin{aligned} \mathcal{U}_{\text{prep}} &= \mathcal{R}_{\mathbf{n}} \left[\bigotimes_{j=1}^n \mathcal{T}_j \right] \mathcal{R}_{\frac{\pi}{2}} \\ \mathcal{U}_{\text{meas}} &= \mathcal{R}_{\mathbf{m}} \left[\bigotimes_{j=n+1}^{n+m} \mathcal{T}_j \right] \mathcal{R}_{\mathbf{n}}^\dagger, \end{aligned} \quad (39)$$

where we introduced the abbreviations $\mathcal{T}_j = \mathcal{T}_{\mathbf{k}_j}(\mu_j)$ and $\mathcal{R}_{\frac{\pi}{2}} = \mathcal{R}_y(-\frac{\pi}{2})$. The $\pi/2$ -pulse $\mathcal{R}_{\frac{\pi}{2}}$ in \mathcal{U}_{en} generates the CSS polarized in x -direction $|+\rangle^{\otimes N} = (|\uparrow\rangle + |\downarrow\rangle)^{\otimes N}/\sqrt{2^N} = \mathcal{R}_{\frac{\pi}{2}}|\downarrow\rangle^{\otimes N}$.

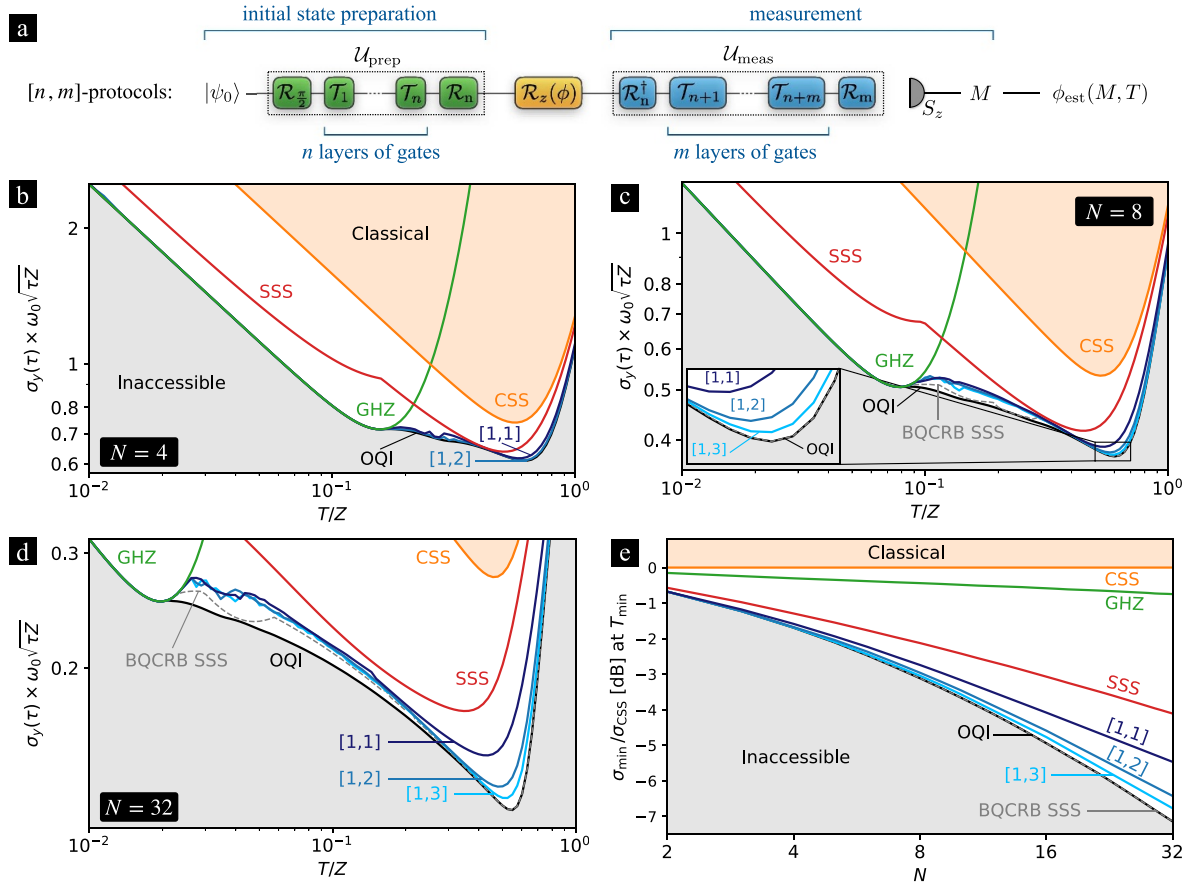


Figure 3. (a) Visualization of the variational Ramsey protocols defined in equation (39). The $\pi/2$ -pulse $\mathcal{R}_{\frac{\pi}{2}} = \mathcal{R}_y(-\pi/2)$ generates the coherent spin state (CSS) polarized in x -direction from the ground state $|\psi_0\rangle = |\downarrow\rangle^{\otimes N}$. Entanglement in the initial state and measurement is introduced via one-axis-twisting (OAT) interactions, denoted by $\mathcal{T}_j = \mathcal{T}_{\mathbf{k}_j}(\mu_j)$ with twisting strength μ_j around axis \mathbf{k}_j . During the free evolution time T , the phase ϕ is imprinted onto the initial state via a rotation around the z -axis. The rotations \mathcal{R}_n and \mathcal{R}_m result in an effective phase evolution around an arbitrary axis \mathbf{n} , $S_n = \mathcal{R}_n^\dagger S_z \mathcal{R}_n$, and an effective measurement of $S_m = \mathcal{R}_m^\dagger S_z \mathcal{R}_m$, respectively. Finally, the phase is estimated based on measurement outcome M of observable S_z . (b)–(d) Approximating the OQI using variational $[1, m]$ -classes (blue) for (a) $N=4$, (b) $N=8$ and (c) $N=32$. For comparison, the standard protocols are shown as they naturally emerge as specific quantum circuits within the variational classes. Additionally, the BQCRB of SSS is shown (dashed gray). With increasing N , the complexity of the variational circuits required to approach OQI performance increases. (e) Scaling of the gain in clock stability compared to CSS at the optimal interrogation time T_{\min} with N .

The rotations \mathcal{R}_n and \mathcal{R}_m result in an effective phase evolution around an arbitrary axis \mathbf{n} , $S_n = \mathcal{R}_n^\dagger S_z \mathcal{R}_n$, and an effective measurement of $S_m = \mathcal{R}_m^\dagger S_z \mathcal{R}_m$, respectively. Similarly, each OAT $\mathcal{T}_{\mathbf{k}}(\mu) = \mathcal{R}_{\mathbf{k}}^\dagger \mathcal{T}_z(\mu) \mathcal{R}_{\mathbf{k}}$ can be expressed as an OAT with respect to the z -axis and a rotation $\mathcal{R}_{\mathbf{k}}$. The resulting variational classes are not restricted by any geometric constraints. By employing the optimal Bayesian estimator, which can take arbitrary non-linear forms, there are likewise no restrictions imposed on the signals. Consequently, the signals often exhibit strongly non-sinusoidal shapes without any symmetry and no apparent relation to the phase (cf appendix A.5.2), which may initially seem counterintuitive. However, when combined with the corresponding estimator, this approach achieves a low phase estimation uncertainty, as we will see in this section. A similar behavior has already been observed for the GHZ protocol, where the estimator effectively mimics a parity measurement while the signal itself vanishes. In contrast, the use of

a linear estimator, as defined in equation (20), typically results in signals that are anti-symmetric, at least within the range of the prior distribution.

For a given protocol class $[n, m]$, the quantum circuit contains $n + m$ OAT interactions with associated twisting strengths μ_j . Together with the rotations $\mathcal{R}_n, \mathcal{R}_m, \mathcal{R}_{\mathbf{k}}$, which ensure geometric generality and are each characterized by two variational parameters, the total number of variational parameters is $4 + 3(n + m)$. Notably, the particular choice of the CSS $|+\rangle^{\otimes N}$ as the initial state allows us to fix the first OAT of U_{prep} along the z -axis without losing any generality. This simplification reduces the total number of variational parameters by two.

The variational class defined in equation (39) contains the standard Ramsey protocols as limiting cases. CSSs are recovered in the $[0, 0]$ -protocol, while SSSs are implemented within the $[1, 0]$ -class. The GHZ protocol emerges as a special

case, either within the $[1, 0]$ -class using the optimal Bayesian estimator, as discussed in the previous section, or as part of the $[1, 1]$ -class, as implemented in [79].

3.3. Optimal protocols

For fixed ensemble size N , circuit depth $[n, m]$ and prior phase width $\delta\phi$, the optimization of the quantum circuits introduced above is performed over all variational parameters. To enable a general discussion, we adopt the power-law scaling of the prior width with interrogation time T , as defined in equation (31). Results for exemplary ensemble sizes N as well as the scaling of the stability with N are presented in figure 3. The variational protocols are primarily compared to the OQI, as saturating it is the central goal of this section. Additionally, comparisons to standard Ramsey protocols are provided where relevant to highlight specific advantages and limitations.

General results—We begin by examining the general behavior and scaling of the variational classes, with a particular focus on the number of layers n and m . While $[n, 0]$ -protocols yield collective spin measurements with sinusoidal signals, increasing m allows for arbitrary signal shapes (cf appendix A.5.2), since no geometric constraints are imposed. As for the standard Ramsey protocols, variational protocols exhibit a clear trade-off between enhanced stability for increasing interrogation times and the CTL of the LO.

At long interrogation times close to the minimal ADEV σ_{\min} , the OQI is saturated by the BQCRB of SSS for any ensemble size N . Within the variational framework, this can be implemented using the $[1, m]$ -classes, as the optimal measurement of the BQCRB is approximated in the limit $m \gg 1$. Increasing the number of entangling layers n yields σ_{\min} comparable to that of the corresponding $[1, m]$ -protocols, consistent with findings in [74]. Consequently, to extend the dynamic range and approximate the OQI at long interrogation times requires to increase m .

In contrast, at short interrogation times, the dynamic range is negligible, and increasing the entanglement depth of the initial state, effectively determined by n , becomes beneficial. However, GHZ states are already optimal in this regime and saturate the HL. Thus, $n = 1$ remains sufficient.

At the plateau of the OQI, where GHZ states become ineffective, the $[1, m]$ -classes in general do not saturate the OQI. This regime becomes broader for larger ensembles, since the dynamic range of GHZ states reduces with N . Here, achieving the OQI requires asymptotically deep quantum circuits, which, however, is unfavorable. Even in the limit of $n + m \gg 1$, as considered in [73, 76], the variational class only gradually approximates the OQI with increasing complexity. Furthermore, in the plateau regime, the optimal variational parameters strongly depend on the prior width, causing substantial variations in the interferometer sequence. As a consequence, even minor modifications in the interrogation time can lead to profound changes in the form of both the signals and the associated estimation strategies. In particular, as the regime of GHZ states is exceeded, the twisting strengths decrease significantly, effectively reducing

the degree of entanglement to adapt to increased LO noise. Interestingly, this susceptibility diminishes with increasing circuit complexity m . This can be interpreted as follows: For low depth quantum circuits, the variational degree of freedom is limited and thus, the optimal states and measurements have to be extremely well tailored to a specific prior width to ensure a sufficiently high degree of entanglement and dynamic range at the same time. As the variational complexity increases, the variational space grows and reduces the susceptibility to small variations in the prior width. Additionally, this dependence gives rise to a large number of local minima, making global optimization tedious and facilitating numerical errors, indicated by the non-smooth curves.

As a consequence, we focus on approximating the OQI in all regimes except the plateau using variational $[1, m]$ -classes and strive for a minimal circuit depth m .

Protocol complexity and ensemble size—For the simplest case, $N = 2$, the GHZ protocol is optimal across most interrogation times because the critical prior width, $\delta\phi \sim 1/N$, is relatively large. Consequently, the region between the plateau and the minimal ADEV is narrow. In this transitional regime, SSS achieve the OQI, while the minimal ADEV as well as the plateau of the OQI are saturated by the simplest non-trivial variational class, the $[1, 1]$ -protocols. Hence, standard protocols are sufficient to saturate the OQI across a wide range of interrogation times.

For $N = 4$, illustrated in figure 3(b), the plateau of the OQI already broadens significantly. SSS perform close to the OQI and the variational protocols over a narrow range of interrogation times in the transition regime. The $[1, 1]$ -protocols remain optimal for a broad range of interrogation times, closely approaching the OQI at the plateau and at the minimal ADEV. Increasing the circuit complexity, the $[1, 2]$ -class approximates the OQI.

For $N = 8$, depicted in figure 3(c), the discrepancy between low-depth quantum circuits and the OQI at the plateau becomes more pronounced. Even the BQCRB of SSS, effectively represented by $[1, m]$ -classes for $m \gg 1$, approximates the OQI only closely. In the transition regime, where the OQI scaling reverts to $\sim 1/\sqrt{T}$, SSS approximate both the variational classes and the OQI, but the deviation grows with N , as discussed in section 3.1. At the minimal ADEV, $[1, 1]$ -protocols substantially extend the dynamic range, but nevertheless leave a noticeable gap to the OQI, which is largely closed by $[1, 2]$ -classes. Increasing the variational complexity further, the $[1, 3]$ -class approximates the OQI in the vicinity of σ_{\min} , but does not fully saturate it. Consequently, already for $N = 8$ relatively deep quantum circuits are required to saturate the OQI entirely. Since the gain diminishes with increasing m , and in order to keep the quantum circuit comparably simple, we do not increase the circuit depth further.

As N increases, this trend continues, as shown for $N = 32$ in figure 3(d). In this case, even the BQCRB of SSS exhibits significant deviations from the OQI at the plateau. In contrast, in the scaling regime of $\sim 1/\sqrt{T}$, and particularly at the minimal ADEV, the BQCRB saturates the OQI. The overall minimum is approximated by increasing m , but diminishing gains make

deeper circuits less advantageous. Hence, we restrict our analysis to variational classes $[1, m]$ with $m \leq 3$ as before. In general, the variational complexity required to saturate the OQI grows with N (cf figure 3(e)). These results align with the asymptotic analysis of OQI saturation in [73, 74, 76].

For large ensemble sizes ($N \gtrsim 100$), reaching the regime of optical lattice clocks, atom number fluctuations during interrogation become relevant [7]. Variational Ramsey protocols, optimized for fixed N , are highly sensitive to such fluctuations, making them less favorable in this regime. Instead, the POI emerges as a robust alternative, saturating the OQI in the limit of large N .

In summary, for systems with small ensembles N , such as ion-traps and tweezer-arrays, low-depth variational classes are sufficient to approximate the OQI. These protocols generally achieve optimal performance across all interrogation times, except at the OQI plateau. Already the simplest variational protocols from the $[1, 1]$ -class significantly enhance the stability at long interrogation times, particularly in the regime of the minimal ADEV. The circuit depth of $[1, m]$ protocols required to actually saturate the OQI at long interrogation times increases with N . However, the performance gain diminishes with m , presenting a trade-off between reduced instability and increasing complexity. To maintain a reasonable balance between dynamic range and circuit depth, we restrict our analysis to $m \leq 3$, acknowledging that the OQI can be fully saturated in the limit of deep circuits $m \gg 1$, as quantified by the BQCRB of SSS.

4. Application in the full feedback loop of an atomic clock

The Bayesian approach captures key aspects of atomic clock operation, including finite prior information, single-shot measurements, and the trade-off between enhanced stability achieved through longer interrogation times and the CTL of the LO. However, it models only a single clock cycle and neglects cumulative effects that arise in a full feedback loop. In particular, in the regime where the invertible domain of the main fringe is exceeded by the prior distribution and thus an unambiguous phase estimation is no longer possible, so called *fringe hops* might occur. In this scenario, the feedback loop passes to an adjacent fringe resulting in the clock running systematically wrong and consequently degrading the clock stability. Whether fringe hops or the CTL impose the dominant constraint depends on the specific Ramsey protocol and interrogation time. Since fringe hops are a feature only emerging in the context of a full feedback loop, they are not captured by the theoretical model presented above. While existing approaches, such as those in [68, 118], provide rough estimates for the effects of fringe hops based on single cycle properties, they are typically limited to sinusoidal signals and lack general applicability. A rigorous treatment of fringe hops requires modeling the complete feedback loop, as pursued in [69], but adapting this framework to variational Ramsey protocols lies beyond the scope of this work. Instead, we perform realistic

Monte Carlo simulations of the full feedback loop to validate our theoretical predictions on clock stability. These numerical simulations reflect the basic principles of atomic clock operation (cf section 2.1), building on the methods presented in [67, 68]. Further implementation details are provided in the appendix A.5.3. The prior width in the full feedback loop is determined iteratively, as discussed in section 2.6 and appendix A.5.4.

To start with, in section 4.1 we examine the limitations imposed by fringe hops and discuss the associated deviations between theoretical predictions and numerical simulations. In section 4.2, we investigate the clock stability within the full feedback loop of an atomic clock for various Ramsey protocols and ensemble sizes, identifying the protocols that perform best in the respective regimes. Furthermore, in section 4.3, we compare the linear estimation strategy with the optimal Bayesian estimator, focusing particularly on variational quantum circuits and the limitations imposed by fringe hops.

4.1. Limitation due to fringe hops

Results of numerical simulations, presented in figure 4, show good agreement with theoretical predictions across a wide range of interrogation times. However, significant deviations arise in two regimes.

First, for small ensembles at long interrogation times, fringe hops limit the clock stability rather than the CTL. As a result, the minimal ADEV σ_{\min} is not achieved for the standard protocols and variational classes. Instead, the best stability is observed at $T_{\text{sim}} < T_{\min}$, lying within the transition regime between the plateau of the OQI and σ_{\min} . However, as N increases, T_{sim} approaches the CTL at T_{\min} , resulting in improved stability. In particular, for $N \gtrsim 20$, fringe hops and the CTL spoil the stability at the same level and thus, the minimal ADEV is achieved for the standard protocols and variational classes. Notably, GHZ protocols remain limited by fringe hops regardless of N due to their inherently narrow dynamic range which decreases with the ensemble size.

Second, deviations arise in the regime of the plateau of the OQI, which primarily can be explained by three arguments: (i) similar to long interrogation times, fringe hops can occur in this regime. For instance, for $\delta\phi \lesssim 1/N$, the optimal variational protocols resemble the GHZ protocol. However, as argued before, fringe hops prevent GHZ protocols to achieve its minimal ADEV. Likewise, the optimal variational protocols may not attain the theoretical prediction as $\delta\phi \sim 1/N$. In this regime, $m \neq 0$ typically generates highly non-sinusoidal signals with reduced dynamic range compared to CSS and SSS (cf appendix A.5.2), resulting in severe limitations due to fringe hops. (ii) As described in the previous section, in the regime of the OQI plateau, the optimal variational parameters are highly sensitive to small changes in the interrogation time, where this susceptibility diminishes with the circuit complexity m . With increasing ensemble size, this limitation increases, since the plateau gets broader with N . Although the prior width is determined iteratively, modeling the actual

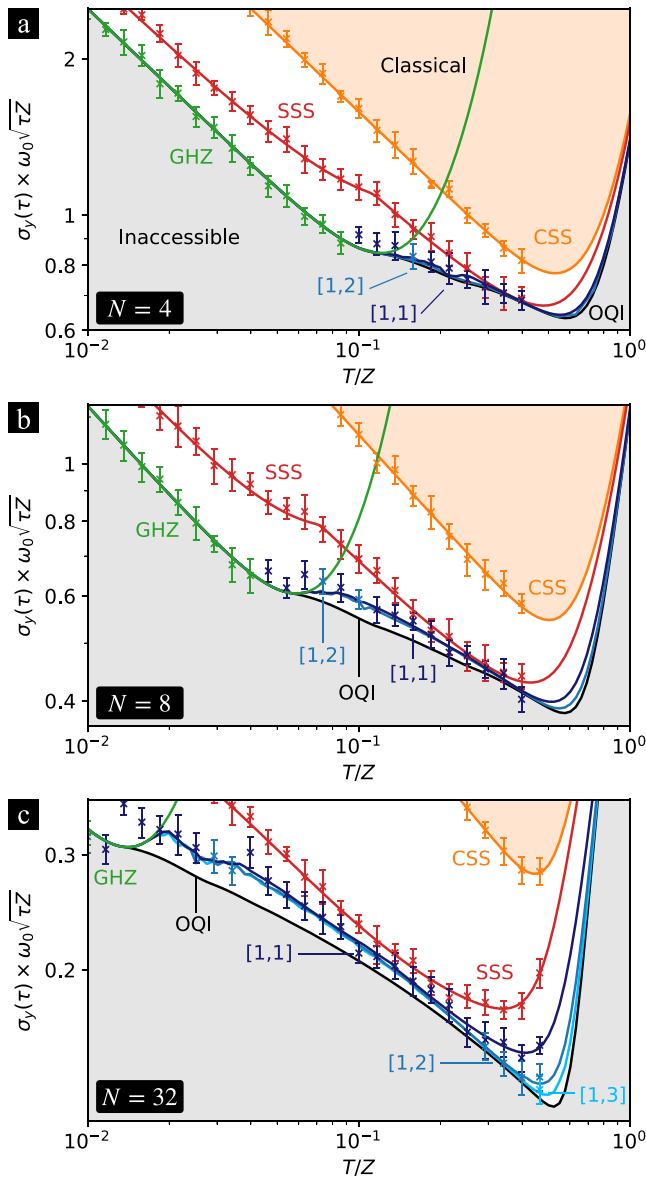


Figure 4. Numerical simulations of the full feedback loop in an atomic clock compared to the theoretical predictions of the Allan deviation for ensemble sizes (a) $N = 4$, (b) $N = 8$ and (c) $N = 32$. The variational Ramsey protocols $[n, m]$ consist of n and m layers of one-axis-twisting interactions for state preparation and measurement, respectively. Symbols represent the mean clock stability, while error bars indicate fluctuations over independent clock runs, arising from the stochastic nature of the Monte Carlo simulations. Fringe hops limit stability in the plateau regime and at long interrogation times, as discussed in the main text. The associated prior width is obtained iteratively. Further details on the numerical simulations are provided in appendix A.5.3.

prior distribution solely based on the width remains a simplified parameterization of the prior knowledge. Furthermore, this iterative evaluation of the prior width relies on a fixed interrogation sequence (cf appendix A.5.4), which may not capture the true prior width of variational protocols sufficiently accurately. Consequently, the optimization can lead to

variational protocols that are more susceptible to the true residual noise than predicted by the model, resulting in deviations between theoretical predictions and numerical simulations. (iii) Additionally, the assumption of a Gaussian prior distribution for the residual noise in each cycle may not reproduce the true dynamics appropriately. In particular, for small ensembles, the number of possible measurement outcomes is small and thus, the central limit theorem justifies this assumption to a reasonable level only in the asymptotic limit of many repetitions. Consequently, corresponding deviations reduce with increasing N .

As a consequence, stability can be compromised in both regimes. To address these limitations, we simulate the clock performance of several protocols for a fixed interrogation time T and variational class $[n, m]$, corresponding to distinct local minima in the parameter landscape, and select the protocol that achieves the best stability (see appendix A.5.3). Consequently, the best-performing protocol identified in simulations may differ from the theoretical optimum, leading to deviations between simulation and theory. Furthermore, the sensitivity landscape typically features numerous local minima and thus, it is not feasible to simulate all emerging protocols. In extreme cases, fringe hops may affect all simulated protocols, leading to complete stability loss. Hence, we show the least complex variational class $[1, m]$ that achieves theoretical predictions at the OQI plateau. At long interrogation times, we include simulation results of deeper quantum circuits where a substantial gain is observed.

4.2. Clock stability

Overall, numerical simulations align closely with theoretical predictions across a wide range of interrogation times. However, as discussed in the previous section, fringe hops impose the primary limitation at the OQI plateau. Additionally, for ensembles with $N \lesssim 20$ and long interrogation times, fringe hops limit the clock stability rather than the CTL of the LO. As a consequence, the minimal ADEV σ_{\min} is not achieved for small ensembles, and variational protocols provide marginal to no advantage over SSS in this regime. In particular, we distinguish between three regimes based on the ensemble size:

- (i) For very small ensembles with $N \lesssim 4$ (cf figure 4(a)), the GHZ protocol saturates the ultimate lower limit, represented by the OQI, for short interrogation times, while at long interrogation times, approaching the fringe hop limit T_{sim} , SSS become optimal. Hence, variational protocols provide an advantage over standard protocols only in the regime of the OQI plateau. In this regime, typically the simplest $[1, 1]$ -class already is sufficient to saturate the OQI. However, the optimal protocols vary significantly with interrogation time, increasing their susceptibility to fringe hops. Furthermore, the OQI plateau is relatively narrow for $N \lesssim 4$. Given the trade-off between potential stability gains and the experimental challenges

involved in implementing more complex Ramsey protocols, the GHZ protocol at short interrogation times and SSS at longer interrogation times remain the preferable choices for ensembles with $N \lesssim 4$.

- (ii) For intermediate ensembles with $N \lesssim 20$ (cf figure 4(b)), fringe hops continue to impose the fundamental limitation at long interrogation times. Furthermore, the regime of the OQI plateau expands, which in turn increases the region where variational protocols provide an advantage over SSS. Nevertheless, $[1, m]$ protocols remain fragile to fringe hops in this regime and additionally do not suffice to actually saturate the OQI. At long interrogation times approaching T_{sim} , which itself approximates T_{min} with increasing N , variational protocols offer improved stability compared to SSS. In this regime, again the simplest $[1, 1]$ -class is sufficient to achieve a relevant improvement, while the additional benefit of $[1, m]$ protocols with $m > 1$ is negligible when considering the fluctuations over independent clock runs. However, to achieve a gain compared to SSS at long interrogation times requires $T \sim T_{\text{sim}}$. For practical implementation in an experiment, [67] suggests to choose an interrogation time slightly shorter than T_{sim} , effectively providing a safety margin against fringe hops. As a result, similar to $N \leq 4$, variational protocols for $N \lesssim 20$ effectively enhance clock stability primarily within the OQI plateau regime, which remains less favorable in experimental settings, while GHZ states and SSS are beneficial at short and long interrogation times, respectively.
- (iii) As the ensemble size increases to $N \gtrsim 20$ (cf figure 4(c)), the limitations imposed by fringe hops and the CTL become comparable. In this regime, variational protocols succeed to achieve σ_{min} , resulting in a substantial gain in stability over SSS. Furthermore, as N grows, increasing the circuit complexity m of the $[1, m]$ protocols provides relevant gains in stability.

To conclude, variational protocols for clocks with only a few atoms, characteristic of ion-traps, primarily enhance stability within the OQI plateau. However, this regime is less favorable due to the strong dependence of variational parameters on interrogation time and increased susceptibility to fringe hops. In contrast, for clocks with several tens of atoms, typical of tweezer-arrays, variational Ramsey protocols offer a significant improvement in clock stability, particularly at long interrogation times. Here, low-depth quantum circuits are sufficient, as the benefits diminish with increasing m , resulting in a trade-off between increased complexity and extended dynamic range.

4.3. Comparison of linear and optimal Bayesian estimation

In section 3.1, we observed that the optimal Bayesian estimator achieves significant improvements over the linear estimator in several regimes for standard protocols. For instance, it provides a larger dynamic range at long interrogation times and enables stronger spin-squeezing at short interrogation

times due to its non-linearity. Nevertheless, the linear estimator delivers equivalent results at interrogation times where the signal can be linearized within the extent of the prior distribution, which typically corresponds to $T/Z \ll 1$. Moreover, the linear estimator simplifies numerical studies (see section 2.4) and has delivered remarkable results in previous works [67, 68], including applications in variational Ramsey interferometry [73, 74, 76]. Therefore, we compare the performance of the linear and optimal Bayesian estimators in the context of variational interrogation protocols to determine whether the potential advantages of the optimal Bayesian estimator, while significant in some regimes for standard protocols, translate into meaningful improvements in the case of variational quantum circuits.

In figure 5, we compare theoretical predictions of clock stability for optimized variational $[1, m]$ protocols employing both estimation strategies. Surprisingly, the linear estimator effectively achieves the same stability as the optimal Bayesian estimator. In particular, the optimal Bayesian estimator does not extend the dynamic range at long interrogation times and correspondingly does not enhance the minimal ADEV, while offering only a marginal enhancement in the plateau regime of the OQI, where GHZ protocols become ineffective. However, this gain is negligible, especially when considering the stability issues in this regime discussed in the previous sections. Consequently, in theory, the optimal Bayesian estimator does not provide a relevant improvement over the linear estimator, which is consistent with findings in [76] for exclusively anti-symmetric signals (cf supplementary discussion S9 in [76]).

While theoretical predictions offer valuable insights, their validation in realistic scenarios is essential for a comprehensive analysis, as discussed before. Figure 5 additionally presents numerical simulations of the full feedback loop. The standard protocols perform as predicted by theory, exhibiting the same limitation imposed by fringe hops at long interrogation times, as observed with the optimal Bayesian estimator. For larger ensemble sizes $N \gtrsim 20$, where the CTL and fringe hops constrain clock stability at the same level, the reduced dynamic range of the linear estimator for sinusoidal signals becomes relevant. As a result, the optimal Bayesian estimator achieves higher stabilities for CSS and SSS, as discussed in section 3.1. Again, variational protocols are constrained by fringe hops in two distinct regimes. (i) At the OQI plateau, the susceptibility to fringe hops is significantly enhanced for the linear estimator compared to the optimal Bayesian estimator, as indicated by larger deviations between theoretical predictions and numerical simulations, as well as broader regions where deeper quantum circuits are required to achieve theoretical expectations. However, given the strong variation in optimal Ramsey schemes and the emergence of fringe hops, operating a clock in these regimes may be experimentally unfavorable anyway, as discussed in previous sections. Thus, the stronger limitation imposed by fringe hops in this regime is of minor practical relevance. (ii) For $N \lesssim 20$, clock stability is limited by fringe hops at the same level for both estimation strategies, leading to comparable maximal interrogation times T_{sim} (cf figures 5(a) and (b)). In contrast to

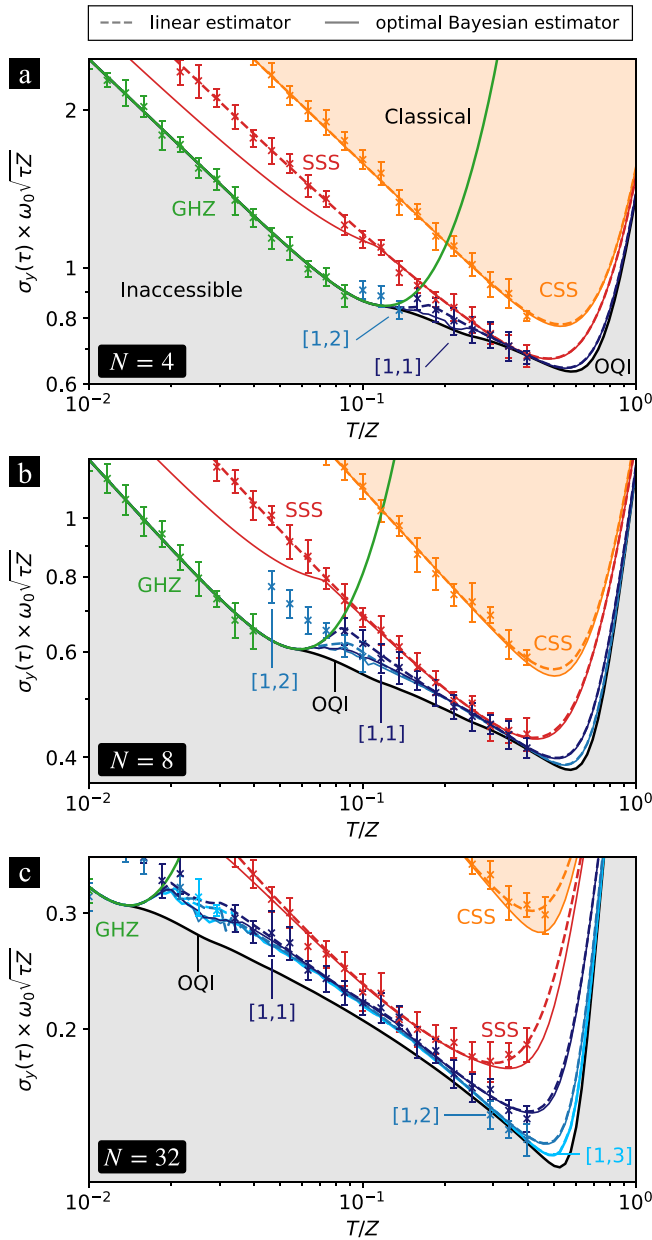


Figure 5. Numerical simulations of the full feedback loop in an atomic clock with a linear estimation strategy for ensemble sizes (a) $N = 4$, (b) $N = 8$ and (c) $N = 32$ characterized by the Allan deviation. The variational Ramsey protocols $[n, m]$ consist of n and m layers of one-axis-twisting interactions for state preparation and measurement, respectively. Lines depict theoretical predictions with the linear (dashed) and optimal Bayesian estimator (solid). Symbols represent the mean clock stability, while error bars indicate fluctuations over independent clock runs, arising from the stochastic nature of the process. Fringe hops limit stability in the plateau regime and at long interrogation times, as discussed in the main text. Further details on the numerical simulations are provided in appendix A.5.3.

the optimal Bayesian estimator, for larger ensembles with $N \gtrsim 20$, fringe hops remain the dominant limitation when using the linear estimation strategy. As a consequence, the minimal ADEV σ_{\min} is not achieved for $[1, m]$ protocols, as

illustrated in figure 5(c). Therefore, in clocks with a few tens of atoms, typically realized in tweezer-arrays, the linear estimation strategy causes fringe hops to impose a stricter constraint on clock stability than the coherence time of the LO.

In summary, the optimal Bayesian estimator guarantees to saturate the BCRB, thereby maximizing the use of the measurement data. Whether the linear estimation strategy can achieve comparable performance depends strongly on the particular interrogation scheme and must be evaluated for each specific scenario. For variational Ramsey protocols, as considered in this work, the optimal Bayesian estimator proves to be less susceptible to fringe hops. While this difference may be negligible in the regime of the OQI plateau, where these protocols are potentially unfavorable for experimental implementation, the critical ensemble size at which fringe hops and the CTL constrain the clock stability at the same level is larger when using the linear estimation strategy.

Moreover, it is important to note that the estimation strategy primarily affects the classical post-processing of the measurement outcome. Consequently, the complexity of the Ramsey sequence remains unchanged for both estimation strategies. The quantum circuit itself is only indirectly influenced, as the choice of estimator affects the optimal variational parameters. Typically, the optimal Bayesian estimator leads to smaller total twisting strengths $\mu = \sum_j |\mu_j|$, particularly for variational classes $[1, m]$ with $m > 1$. Hence, the linear estimation strategy effectively requires larger twisting strengths to compensate for the non-linearity of the optimal Bayesian estimator. As a result, quantum circuits employing the optimal Bayesian estimator achieve shorter gate durations, which may be of practical interest.

5. Dead time

In the previous sections, we extensively discussed the trade-off between QPN, which decreases with the interrogation time T (cf equation (26)), and the CTL of the LO, which constrains clock stability at long interrogation times. Here, we extend the discussion to account for the effect of dead time T_D in atomic clock operation. Dead time typically arises from processes such as probe preparation, measurement, and the application of feedback. During this period, frequency fluctuations of the LO remain unmonitored by the Ramsey interrogation and therefore cannot be measured or corrected. The cumulative effect of this lack of information degrades the long-term stability of the atomic clock, a phenomenon first described by Dick [119, 120], therefore commonly referred to as the Dick effect. The contribution of the Dick effect to the clock stability is directly inferred from the spectral noise density $S_y(f)$ of the LO and is given by [120, 121]

$$\sigma_{y,\text{Dick}}^2(\tau) = \frac{1}{\tau} \frac{T_C^2}{T^2} \sum_{k=1}^{\infty} S_y\left(\frac{k}{T_C}\right) \frac{\sin^2(\pi k T / T_C)}{\pi^2 k^2}, \quad (40)$$

where $T_C = T_D + T$ is the clock cycle duration. The impact of the Dick effect diminishes with longer interrogation times T ,

as it depends on the ratio T/T_C which decreases when the relative contribution of dead time is reduced. Taking dead time into account, the overall clock stability is determined by the interplay between QPN, CTL, and Dick noise. Specifically, it is characterized by the total ADEV

$$\sigma_{y,\text{tot}}(\tau) = \sqrt{\sigma_{y,\text{QPN}}^2(\tau) + \sigma_{y,\text{CTL}}^2(\tau) + \sigma_{y,\text{Dick}}^2(\tau)}, \quad (41)$$

where QPN and the CTL are combined in the Bayesian framework as $\sigma_y^2(\tau) = \sigma_{y,\text{QPN}}^2(\tau) + \sigma_{y,\text{CTL}}^2(\tau)$. While the Bayesian approach generally does not permit a strict separation of QPN and CTL contributions, except in specific cases such as the OQI or for CSS and SSS with a linear estimator (cf section 3.1), it is nevertheless advantageous to treat them formally as independent components to discuss their general scaling quantitatively. The trade-off characterized by $\sigma_{y,\text{tot}}(\tau)$ has been thoroughly studied for CSS and SSS with a linear estimator in [68], where QPN was characterized using local phase estimation theory, while the CTL was modeled via a stochastic differential equation describing the stabilized frequency of the LO. In contrast, in this work we adapt the discussion to the Bayesian framework, which provides an intuitive and comprehensive approach to treating these effects. Additionally, for comparison, we include the ultimate lower bound on clock stability represented by the OQI. After analyzing the general scaling of $\sigma_{y,\text{tot}}(\tau)$ for standard protocols, we consider various experimental platforms and discuss the effect of dead times characteristic of each setup. Furthermore, we explore the potential benefits of variational quantum circuits in these regimes.

5.1. Dead time in Bayesian frequency metrology

In addition to the contribution described by equation (40), dead time affects Bayesian frequency metrology in two distinct ways. First, and most notably, it modifies the scaling of the ADEV, associated with QPN and the CTL, as a function of the interrogation time T . Instead of the ideal $\sim 1/\sqrt{T}$ scaling, dead time reduces it to $\sim \sqrt{T_C/T^2}$, as apparent in equation (29). Second, dead time broadens the prior distribution of the phase due to unmonitored frequency fluctuations during T_D . Among these two effects, the modified scaling with T has a substantially larger impact, whereas the broadening of the prior distribution introduces only a relatively minor correction. Nevertheless, incorporating the implicit broadening is crucial for accurate modeling and for identifying optimal Ramsey protocols and estimation strategies, as the prior width strongly influences the optimal interrogation sequence, as explored in previous sections.

Although the prior width could, in principle, be adjusted iteratively to include dead time as for $T_D = 0$, this approach is computationally demanding. Moreover, our goal is to establish a direct connection between scenarios with ($T_D > 0$) and without ($T_D = 0$) dead time. Since the additional frequency fluctuations during dead time are unmonitored by the Ramsey interrogation, the broadening of the prior distribution during

dead time and during the Ramsey sequence T are independent processes. Treating the broadening of the phase distribution during dead time as a phase diffusion process, the modified prior distribution $\mathcal{P}(\phi) = (\mathcal{P}_D * \mathcal{P}_T)(\phi)$ is obtained by a convolution of the initial prior distribution $\mathcal{P}_T(\phi)$, resulting from the Ramsey interrogation time T with corresponding width $\delta\phi_T$ (cf section 2.6), and the distribution $\mathcal{P}_D(\phi)$ associated with dead time. In this context, $\mathcal{P}_D(\phi)$ effectively acts as a Green's function [122]. Although LO noise in general is correlated, the additional noise introduced during dead time within the full feedback loop is well approximated as white noise in the asymptotic limit of many clock cycles. Consequently, $\mathcal{P}_D(\phi)$ is modeled as a Gaussian distribution with zero mean and a width $\delta\phi_D$. As a result, the modified prior distribution $\mathcal{P}(\phi)$ remains Gaussian with zero mean and variance

$$(\delta\phi)^2 = (\delta\phi_D)^2 + (\delta\phi_T)^2. \quad (42)$$

To fully incorporate the impact of dead time into the Bayesian framework, we now relate the broadening of the phase distribution, characterized by $\delta\phi_D$, to the dead time T_D , akin to the dead time-free case in section 2.6. Rather than deriving a comprehensive model for arbitrary scenarios, we establish a relation $\delta\phi_D(T_D)$ that primarily aims to accurately predict behavior in the vicinity of the minimal ADEV σ_{\min} at interrogation time T_{\min} . To this end, the broadening of the prior width during dead time can be effectively modeled by translating the additional frequency fluctuations into hypothetical phase shifts, as if they had occurred during a Ramsey interrogation of duration T_D . In this context, the associated prior width is given by (cf appendix A.5.5)

$$(\delta\phi_D)^2 \simeq 2 \left(\frac{T_D}{Z} \right)^{2+\alpha}, \quad (43)$$

reflecting a power-law dependence, analogous to equation (31). Here, the parameter α again characterizes the nature of the frequency noise, with values $\alpha = -1, 0, 1$ corresponding to white, flicker, and random walk frequency noise, respectively.

As a consequence, adjusting the prior width according to equation (42) extends the Bayesian framework to incorporate dead time within the clock cycle, accounting for both the Ramsey interrogation time T and the dead time T_D . Therefore, aside from adapting the prior width to reflect dead time T_D , the findings from the previous sections remain directly applicable. Therefore, the primary remaining task is to analyze the impact of the Dick effect $\sigma_{y,\text{Dick}}(\tau)$ on overall clock stability.

5.2. General results

In general, the total clock stability reflects a trade-off between QPN, the CTL and the Dick effect, as described by equation (41). While the CTL emerges at long interrogation times, limiting the clock stability as T approaches the laser coherence time Z , both QPN and Dick noise decrease monotonically with the interrogation time T . Unlike QPN, which

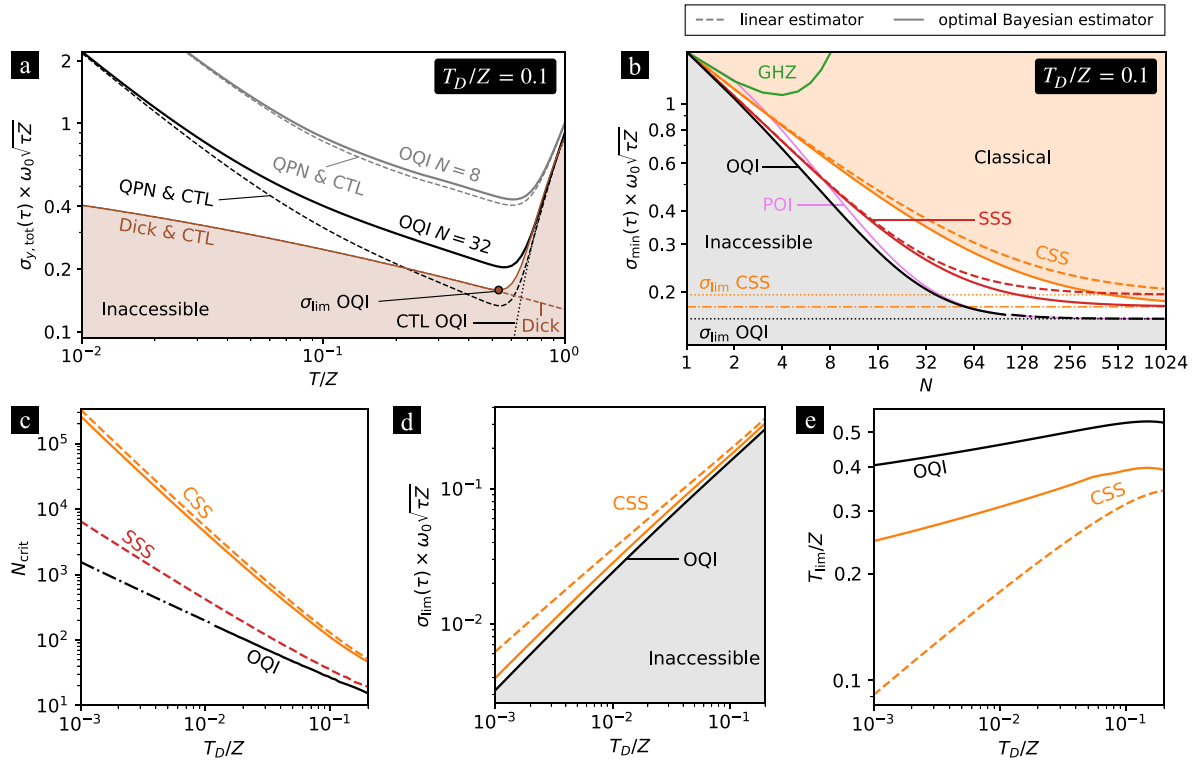


Figure 6. (a) Generic scaling of the dimensionless total Allan deviation $\sigma_{y,\text{tot}}(\tau) \times \omega_0 \sqrt{\tau Z}$ with the interrogation time T for dead time $T_D/Z = 0.1$. The total stability (solid) of the OQI for $N = 8$ (gray) and $N = 32$ (black) is shown in comparison to the trade-off between QPN and CTL (dashed). The N -independent lower limit σ_{lim} (symbol) is imposed by a trade-off (solid brown) between Dick noise (dashed brown) and CTL (dotted black). Consequently, the brown shaded area is inaccessible. (b) Scaling of the total dimensionless minimal Allan deviation $\sigma_{\text{min}}(\tau) \times \omega_0 \sqrt{\tau Z}$ with the ensemble size N for $T_D/Z = 0.1$. GHZ protocols (green) achieve no gain compared to CSS. For CSS (orange) and SSS (red), both the linear (dashed) and optimal Bayesian estimator (solid) are depicted. The gray shaded area represents the inaccessible stability region set by the OQI (black), while the orange shaded area indicates achievable stabilities using uncorrelated atoms. Dotted lines correspond to the lower limit σ_{lim} for the OQI and CSS with linear estimator, while the dashed dotted line denotes the lower limit for CSS using the optimal Bayesian estimator. SSS exhibit the same lower limit as discussed in the main text and seen from the convergence. The POI (violet) saturates the OQI for $N \gtrsim 50$. For the OQI and POI, numerical optimization is performed for $N \leq 100$, while the asymptotic behavior, represented by the π HL (black dashed-dotted), is shown for $N > 100$. (c) Critical ensemble size N_{crit} as a function of the dead time T_D/Z for the OQI (black), CSS (orange) and SSS (red). Again, dashed lines correspond to the linear estimator, while solid lines represent the optimal Bayesian estimator. The evaluation of the SSS with optimal Bayesian estimator requires the computation of the conditional probabilities (cf section 2.4) and thus is unfeasible for large N . For $N \gtrsim 100$, the asymptotic OQI, imposed by the π HL (black dashed-dotted), is shown. (d) Scaling of the total dimensionless lower limit $\sigma_{\text{lim}}(\tau) \times \omega_0 \sqrt{\tau Z}$ with dead time for the OQI (black) and CSS (orange). For the CSS, linear (dashed) and optimal Bayesian estimator (solid) are displayed. (e) Corresponding interrogation times T_{lim} , effectively characterizing the dynamic range. SSS result in the same σ_{lim} and T_{lim} as CSS, cf discussion in the main text.

reduces with larger ensembles, the CTL and Dick noise are independent of N . As a result, whether the minimal ADEV σ_{min} , achieved at optimal interrogation time T_{min} , arises from a trade-off between QPN and CTL or between the Dick effect and CTL depends on the particular dead time, ensemble size and Ramsey protocol [68].

For short dead times or small ensembles, QPN typically dominates Dick noise, leading to behavior that closely resembles the dead time-free case (cf $N = 8$ in figure 6(a)). Here, the clock stability is primarily determined by a trade-off between QPN and CTL and, therefore, depends on the ensemble size as well as the choice of Ramsey sequence. However, as dead time increases or QPN decreases, at some point, QPN is reduced to the level of Dick noise. Since Dick noise typically decreases more slowly with the interrogation

time than QPN, first, it becomes dominant at long interrogation times, limiting the minimal ADEV σ_{min} (cf $N = 32$ in figure 6(a)). Reducing QPN further, by either increasing the ensemble size or adapting the Ramsey interrogation, improves σ_{min} only marginally. In the regime where dead time effects strictly dominate over QPN, as is the case for large ensembles or long dead times, no further improvements in clock stability are possible, as Dick noise is independent of the particular Ramsey sequence and ensemble size. Therefore, we can define a lower limit σ_{lim} on the clock stability, at corresponding interrogation time T_{lim} , which is characterized by a trade-off between Dick noise and CTL. Since the CTL is protocol-dependent, σ_{lim} in general differs for distinct Ramsey protocols and is primarily determined by their respective dynamic range.

Figure 6(b) illustrates the scaling of σ_{\min} with the ensemble size N for the standard Ramsey protocols. For small ensembles, where QPN dominates, clock stability improves as N increases, as in the ideal scenario ($T_D = 0$). However, as the ensemble size grows, Dick noise becomes relevant, reducing the N -scaling and causing the clock stability to converge to σ_{\lim} . Unfortunately, explicit expressions for σ_{\lim} can only be derived for protocols where QPN and CTL are separable, such as the OQI or CSS and SSS with a linear estimator. Otherwise, the convergence towards σ_{\lim} with N has to be evaluated numerically. As argued before, CSS and SSS with a linear estimator exhibit the same CTL and, consequently, identical lower limits. A similar behavior is observed for both protocols using the optimal Bayesian estimator, which however, achieves an improved σ_{\lim} due to the larger dynamic range (cf section 3.1). GHZ protocols, already highly susceptible to LO noise in dead time-free scenarios, are further constrained by dead time, making them suitable only for small ensembles and short dead times. In the asymptotic limit, the performance of the POI again saturates the OQI.

To characterize the transition between the regimes dominated by either QPN or Dick noise, we define a critical ensemble size N_{crit} , at which the ADEV $\sigma_y(\tau)$ (cf equation (29)), arising from QPN and the CTL, saturates σ_{\lim} at T_{lim} . Beyond N_{crit} , Dick noise dominates over QPN and thus spoils the N -scaling of σ_{\min} , which ultimately converges towards σ_{\lim} , without substantial improvements as N increases. Since N_{crit} depends explicitly on QPN, it differs for distinct Ramsey protocols and estimation strategies, as generically illustrated in figure 6(c). For instance, N_{crit} for SSS is substantially smaller than for CSS, since they exhibit the same CTL, but SSS have a substantially smaller QPN. Consequently, the required ensemble size to approach σ_{\lim} is substantially smaller for SSS compared to CSS, with a reduction of up to two orders of magnitude for short T_D , while still maintaining a significant difference even at long dead times. In contrast, the difference between OQI and SSS is relatively small, amounting to less than one order of magnitude for short dead times and becoming effectively negligible as T_D increases.

For a particular dead time T_D , the lower limit σ_{\lim} is determined solely by the CTL of the Ramsey protocol and estimation strategy, essentially reflecting the dynamic range. Consequently, CSS and SSS achieve the same lower limit for a specific estimator. Moreover, as shown in figure 6(d), the enhancement of σ_{\lim} for the OQI compared to CSS or SSS is relatively small. Interestingly, increasing the dynamic range of CSS and SSS by substituting the linear by the optimal Bayesian estimation strategy yields a greater gain than the advantage provided by the OQI over CSS or SSS with the optimal Bayesian estimator.

As T_D increases, the potential enhancement of σ_{\lim} diminishes further. This can be understood as follows: In general, dead time shifts T_{\min} , where the minimal ADEV σ_{\min} is achieved, to longer interrogation times. This is shown in figure 6(e) and primarily results from the impact of $\sigma_{y,\text{Dick}}$ (cf figure 6(a)). However, in this regime, the difference in the CTL for distinct Ramsey schemes decreases with increasing T (cf

figure 2(a)), thereby reducing the advantage associated with a larger dynamic range. While the OQI allows unbiased phase estimation over $[-\pi, +\pi]$, the optimal Bayesian strategy for the CSS and SSS resembles the *arcsin* estimator and thus covers the range $[-\pi/2, +\pi/2]$ (cf section 3.1). As a result, the OQI and CSS or SSS with optimal Bayesian estimator exhibit a similar behavior, where the corresponding gain only marginally reduces with T_D . In contrast, the deviation between the linear and optimal Bayesian estimators for CSS and SSS reduces substantially with T_D , since the corresponding N_{crit} becomes smaller, leading to a reduced gain in dynamic range for the optimal Bayesian estimator, as discussed in section 3.1.

To summarize, for small ensembles N or short dead times T_D , clock stability is primarily limited by QPN, closely resembling the dead time-free case. However, as N or T_D increases, the Dick effect becomes the dominant noise and ultimately limits the clock stability. Beyond the critical ensemble size N_{crit} , which decreases with T_D , the minimal ADEV σ_{\min} converges to the lower limit σ_{\lim} . In this regime, further improvements in clock stability by increasing the ensemble size or adapting the Ramsey sequence are marginal. As a consequence, clocks with large ensembles $N \gg 1$ limited by Dick noise sufficiently well approximate the lower limit σ_{\lim} by employing CSS or SSS.

5.3. Setup specific dead times

Building on the general discussion of dead time effects on clock stability in standard protocols, this section focuses on examining specific examples relevant to particular experimental setups, such as ion-traps, tweezer-arrays and lattice clocks, ranging from a few to thousands of atoms.

In general, atomic clock operation involves three key time scales: the laser coherence time Z , the dead time T_D and the interrogation time T . In a given experimental setup, Z and T_D are primarily independent but fixed, defining a specific ratio T_D/Z . In contrast, T remains an adjustable parameter, which is implicitly constrained by Z . As a consequence, findings on clock stability cannot be trivially rescaled with respect to various laser coherence times Z , as in the dead time-free scenario, or dead times T_D , since a modification of Z or T_D results in a change of the ratio T_D/Z , which in turn has a substantial impact on the clock stability, as discussed in the previous section.

In experimental settings, the dead time T_D and interrogation time T commonly are expressed in terms of the dimensionless duty cycle

$$\eta = \frac{T}{T_C} = \frac{T}{T + T_D}, \quad (44)$$

which quantifies the relative contribution of the interrogation time T to the total duration of the clock cycle $T_C = T_D + T$. Hence, a larger duty cycle η corresponds to a reduced relative impact of dead time. However, it is crucial to note that nevertheless a specific ratio of T_D and Z is always assumed implicitly.

While the laser coherence time Z is independent of the atomic reference, dead time strongly depends on the particular experimental platform. To address this, we investigate the clock stability for typical dead times across the three predominant regimes: ion-traps, tweezer-arrays, and lattice clocks. Each of these platforms exhibits distinct time scale dynamics and operational characteristics that significantly influence clock performance. Ion-traps provide the highest degree of control, including rapid cooling and no need for reloading due to the deep trap depths, resulting in relatively short dead times [7]. Although recent advancements in Coulomb crystals have facilitated multi-ion clocks [12, 13], ion-traps remain inherently limited in scalability, typically operating with only a few ions. In contrast, optical lattice clocks employ large ensembles of hundreds to thousands of atoms, enabling high precision at the cost of experimental challenges such as atom number fluctuations and interatomic collisions [7]. Furthermore, these systems exhibit substantially longer dead times due to processes such as loading the lattice and cooling the atoms [18–21]. Additionally, dead time has a particularly pronounced impact on clock stability in lattice clocks, as QPN is typically suppressed well below the Dick effect, as discussed in the previous section. Tweezer-arrays bridge between these two contrary approaches, offering a balance between the high control in ion-traps and the scalability inherent in lattice clocks [14–17]. By incorporating ensembles of several tens of atoms, they offer a promising compromise between precision, scalability, and operational efficiency.

In the following, we investigate the clock stability for representative dead times T_D and laser coherence times Z across these three distinct regimes. As discussed before, this is equivalently expressed by fixing the ratio T_D/Z . Starting with ion-traps, which feature relatively short dead times, we explicitly examine the interplay between laser coherence time Z , dead time T_D , and interrogation time T , or equivalently the duty cycle η , using state-of-the-art parameter values to develop an intuitive understanding of the relationship between these time scales. Subsequently, we progressively increase the dead time for setups representing tweezer-arrays and lattice clocks, illustrating how dead time increasingly constrains clock performance and how the optimal Ramsey protocols change accordingly.

Besides the experimental platform, dead time is also affected by the particular interrogation sequence. In practice, the preparation time required for different quantum states—particularly entangled states—can vary significantly depending on the specific initial state. Likewise, the measurement time can vary substantially between different measurement strategies, especially for correlated measurement transformations as pursued for the variational Ramsey protocols. For instance, the conventional Ramsey protocol—utilizing CSSs and a projective spin measurement—can be performed relatively quickly, as it relies on standard collective rotations of the spin system. In contrast, SSSs already require non-linear interactions or measurement-based feedback for state preparation, which introduce additional time overheads. While

the variational $[1, m]$ protocols—investigated in the previous sections—have comparable state preparation times to SSS, as they are also generated by a single OAT interaction, each additional layer of the quantum circuit implementing the effective measurement involves interaction times that scale with the corresponding twisting strength. The OQI, which represents the ultimate lower bound and currently lacks a specific experimental implementation, potentially requires even more demanding resources. As a consequence, for a particular experimental setup, dead times arising from state preparation and measurement have to be considered for each interrogation scheme. However, the resulting dead times are highly individual for each experimental setup and are thus challenging to quantify in general. To nevertheless provide theoretical insight, we compare different Ramsey protocols assuming a fixed dead time for each experimental platform, independent of the protocol complexity. While the results derived in the following provide general insight, this assumption should be kept in mind, as it may lead to an overestimation of the performance of highly entangled states in practical scenarios.

Ion-traps—In ion-traps, dead times of about $T_D = 100$ ms are routinely implemented in experiments. Moreover, state-of-the-art clock lasers achieve laser coherence times Z of several seconds. However, in practice, this impressive level of coherence is often not entirely maintained as the laser propagates between the cavity, which not necessarily is located close to the trap or even in the same laboratory, and the ions. While optical path-length stabilization should, in principle, preserve coherence all the way to the ions, experimental imperfections, such as phase noise within the vacuum chamber, typically lead to a degradation of this quality. Therefore, we assume a laser coherence time of $Z = 2$ s at the location of the ions. Hence, in this exemplary scenario we obtain a ratio $T_D/Z = 0.05$. Furthermore, figure 4 demonstrates that fringe hops limit the clock stability at interrogation times around $T_{\text{sim}} \sim 0.4 - 0.5 \times Z$. Consequently, the maximal achievable duty cycle, given by $\eta_{\text{max}} = T_{\text{sim}}/(T_{\text{sim}} + T_D)$, is on the order of 90%.

As discussed in the previous section, for small ensembles and short dead times, clock stability typically resembles the dead time-free scenario ($T_D = 0$), as illustrated in figure 7(a) for $N = 8$ and $T_D/Z = 0.05$. In this case, QPN remains the primary limitation, while the Dick effect has only a marginal impact, leading to a behavior similar to that shown in figure 4(b). However, in comparison, the plateau of the OQI in the presence of dead time is substantially less pronounced and thus, fringe hops impose a less stringent limitation in this regime. As in the $T_D = 0$ case, SSS approximate the OQI in the transition region between the plateau and σ_{min} . At long interrogation times, fringe hops remain the primary constraint, limiting the interrogation time to $T_{\text{sim}} < T_{\text{min}}$. Furthermore, variational protocols effectively provide no significant enhancement in clock stability around T_{sim} . As a result, for optical atomic clocks based on ion-traps, GHZ states and SSS approach the OQI over a broad range of interrogation

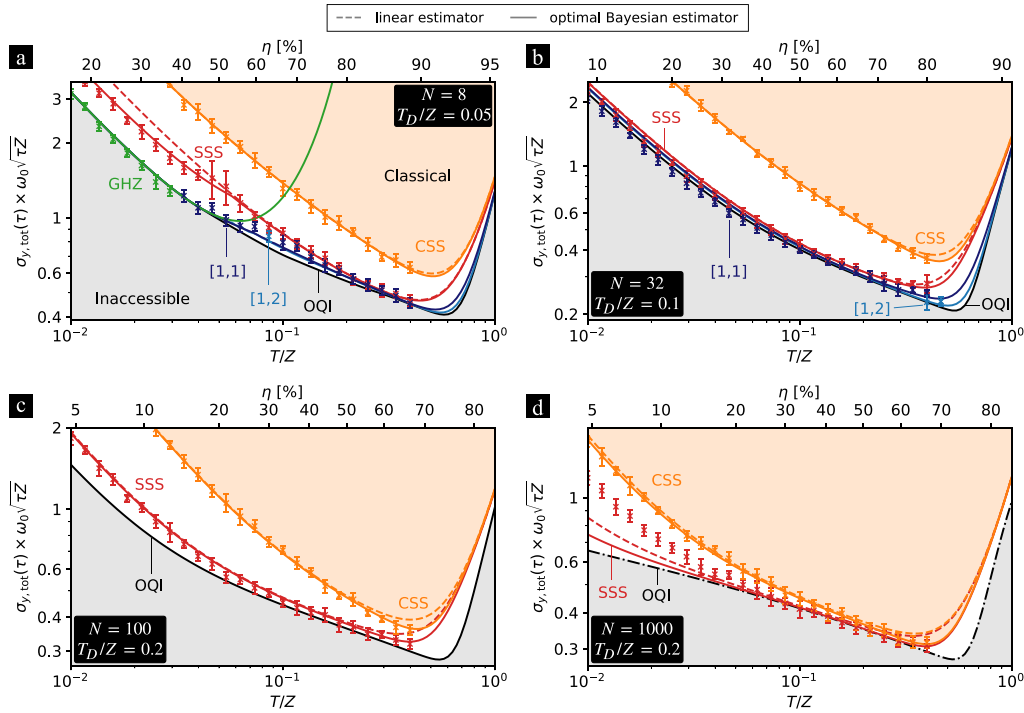


Figure 7. Theoretical predictions and numerical simulations of various Ramsey protocols for (a) $N = 8$ and $T_D/Z = 0.05$, (b) $N = 32$ and $T_D/Z = 0.1$, (c) $N = 100$ and $T_D/Z = 0.2$, (d) $N = 1000$ and $T_D/Z = 0.2$. The variational Ramsey protocols $[n, m]$ consist of n and m layers of one-axis-twisting interactions for state preparation and measurement, respectively. Theory curves (lines) are displayed for the linear (dashed) and optimal Bayesian estimator (solid). Symbols represent numerical simulations in the full feedback loop of an atomic clock employing the optimal Bayesian estimation strategy. In both cases, the total Allan deviation is rescaled with respect to the atomic transition frequency ω_0 , total averaging time τ and laser coherence time Z . The lower x -axis represents the interrogation time T relative to Z , while the upper x -axis denotes the dimensionless duty cycle η . The gray shaded area represents the inaccessible stability region set by the OQI limit (black), while the orange shaded area indicates achievable stabilities using uncorrelated atoms. For $N = 8$ (a) and $N = 32$ (b) the performance of variational quantum circuits (blue) is shown in addition to the standard protocols, namely GHZ states (green), CSS (orange) and SSS (red). For $N = 1000$, the asymptotic regime is reached and thus, the OQI is approximated by the π HL.

times, while the deviation from the OQI or variational classes within the plateau are reduced compared to the dead time-free scenario.

Tweezer-arrays—For tweezer arrays, we consider a representative case with $N = 32$ in figure 7(b), assuming an increased dead time of $T_D/Z = 0.1$. Within the framework of the previous example, this corresponds to an absolute dead time of $T_D = 200$ ms and a maximal achievable duty cycle of approximately $\eta_{\max} = 80\%$. As already evident in figure 6(a), dead time in this regime imposes a significant limitation on clock stability. While GHZ states essentially are ineffective, SSS already perform close to the OQI for short and intermediate interrogation times. Variational protocols offer only marginal improvements in stability, with a noticeable enhancement observed only for $[1, m]$ protocols in the vicinity of T_{\min} . However, this gain is significantly smaller than in the dead time-free case, and unlike the $T_D = 0$ scenario, fringe hops constrain clock stability at interrogation times $T_{\text{sim}} \lesssim T_{\min}$. Additionally considering a safety margin for fringe hops, as discussed in section 4, the improvement becomes effectively negligible. Consequently, SSS emerge as a robust Ramsey sequence, achieving clock stabilities close to the OQI in this regime.

Interestingly, for short interrogation times, deviations between theoretical predictions and numerical simulations appear. These discrepancies stem from the assumed prior width in the presence of dead time, which is intended to provide a reliable model primarily for interrogation times in the vicinity of σ_{\min} .

Crossover regime—Typically, the boundaries with respect to the ensemble size N between different platforms are not sharply defined. To explore the transition between tweezer-arrays and lattice clocks, we examine the case of $N = 100$ in figure 6(c), with an increased dead time $T_D/Z = 0.2$. In the example above, this corresponds to $T_D = 400$ ms and an associated maximal duty cycle of approximately $\eta_{\max} = 65\%$. Such an increase in dead times is characteristic of lattice clocks, as discussed before, but can also result from various processes such as the overhead of operating multiple tweezer arrays simultaneously, the potential need for reloading due to shallower trap depths or extended cooling times. Moreover, inhomogeneous interactions may be relevant, as addressed in [73]. In this regime, variational classes are no longer favorable, as discussed in previous sections. A key characteristic of this regime is that dead time becomes the dominant limitation. However, while CSS have not yet fully converged to the lower

bound σ_{lim} , SSS already approximate it. As a result, SSS perform close to the OQI across all interrogation times, except at T_{min} , where their limited dynamic range becomes apparent. Additionally, the choice of estimation strategy for standard protocols gains importance, as the optimal Bayesian estimator yields significantly higher clock stability at long interrogation times compared to the linear estimator.

Lattice clocks—Finally, we investigate the regime of lattice clocks with large ensembles $N \gg 1$, where QPN is reduced well below the Dick noise. Figure 7(d) illustrates the case of $N = 1000$ with $T_{\text{D}}/Z = 0.2$. In this regime, both CSS and SSS closely approximate the lower limit σ_{lim} . As a result, at long interrogation times, both protocols achieve comparable clock stability, whereas SSS provide a significant advantage at short interrogation times. Furthermore, the optimal Bayesian estimator results in a substantially higher stability in the vicinity of T_{min} compared to the linear estimation strategy. Notably, deviations from theoretical predictions and numerical simulations appear for the SSS at short interrogation times due to the choice of prior width (cf section 5.1). Additionally, since SSS introduce correlations between atoms, unlike CSS, numerical approximations are required to simulate the full feedback loop for $N \gg 1$, which can further contribute to discrepancies.

In summary, for small ensembles representing ion-traps, the behavior closely resembles the dead time-free case. Here, standard protocols as GHZ states or SSS already achieve clock stabilities comparable to the OQI for a wide range of interrogation times. As the ensemble size N or dead time T_{D} increases, Dick noise becomes the dominant limitation, effectively reducing the potential enhancement offered by variational quantum circuits compared to SSS. In particular, dead time results in SSS performing close to the OQI for a variety of scenarios. In the regime of large ensembles $N \gg 1$, characteristic for lattice clocks, CSS likewise converge to the lower limit σ_{lim} at long interrogation times and thus, are sufficient to approximate the OQI. As a consequence, dead time significantly constraints clock stability, where the degree of limitation increases with the ensemble size.

6. Outlook

At this point, rather than reiterating the detailed insights and results from the previous sections, we refer the reader to the introduction for a comprehensive overview. We also highlight the primary assumptions underlying the statements and conclusions of this work. This review is framed within the context of frequency metrology tailored to optical atomic clock time standards—or equivalently, frequency standards—that employ generalized Ramsey protocols to stabilize the frequency output of the laser. Specifically, we focus on single-ensemble clocks in which the atomic reference is periodically interrogated using the same protocol in each clock cycle. Furthermore, we assume that laser noise imposes the primary limitation. For this purpose, we adopted the framework of Bayesian estimation theory to directly incorporate frequency fluctuations into the theoretical model. Our primary metric is

the ADEV, which represents the achievable stability. Other metrics—such as the sensor bandwidth—are not considered relevant in this context, however, they might be important in different scenarios such as for gravitational wave detection.

Approaches to mitigate laser noise limitations—In the following, we briefly discuss promising avenues for further improvements in optical atomic clocks. Overall, laser noise constrains clock stability in three distinct ways: via the laser CTL, the emergence of fringe hops and dead time effects. In addition to ongoing technological improvements in laser stability [66, 123, 124], several interrogation schemes have been proposed that go beyond the conventional single-ensemble approach, where the atomic reference is interrogated with the same protocol in every clock cycle, by employing adaptive schemes [71, 96, 125, 126] and multi-ensemble strategies. For instance, dynamical decoupling sequences [127] and synchronous differential clock comparisons [128–131] have been demonstrated to extend interrogation times well beyond the laser coherence time. Other approaches involve active feedback and feedforward on the laser [132, 133], or cascaded clock operation that allows for increasingly long interrogation times [59, 133–136]. Furthermore, dead time free clock operation can be achieved by asynchronously interrogating two atomic ensembles [24, 137, 138]. As proposed by Rosenband and Leibbrandt in reference [139], partitioning atoms into multiple ensembles with distinct interrogation times can exponentially improve clock stability relative to the atom number. Moreover, synchronous out-of-phase interrogations expand the invertible phase range and enhance the dynamic range [118, 140]. Although these approaches extend beyond the scope of the present work, many of their underlying principles can be integrated with the Ramsey protocols discussed here, potentially mitigating the limitations imposed by laser noise and enabling longer interrogation times.

Decoherence and noise—Beyond the effects of laser noise, other decoherence processes can have a significant impact on the clock performance, depending on the specific parameters of the experiment. While numerical optimization in the context of local (frequentist) phase estimation with decoherence already is a fundamental challenge on its own—as discussed in detail, for example, in the appendix of [65] for spontaneous decay—additionally introducing the averages over the prior phase distribution, inherent in Bayesian estimation theory, further increases the complexity. Consequently, incorporating decoherence effects directly into the theoretical framework of Bayesian phase estimation—as pursued in this work—would combine the numerical challenges of both approaches. Furthermore, the resulting optimal protocols may change substantially and thus would require a thorough investigation, similar to the discussion of dead time. However, this is beyond the scope of this progress report on frequency metrology limited by laser noise. Instead, we briefly review the literature considering decoherence effects.

In the context of Bayesian phase estimation, the robustness of optimal Ramsey protocols to dephasing was investigated by

adding dephasing to the optimal protocol in the decoherence-free scenario. For instance, [72] considers additional collective dephasing that is not associated with laser noise. Since collective dephasing is phenomenologically similar to the treatment of laser noise within the Bayesian framework, it affects stability in much the same way. The impact of uncorrelated single-atom dephasing in the Bayesian framework has been explored in [73], where it was observed that for moderate dephasing strengths the overall behavior remains qualitatively unchanged, although stability is naturally degraded. However, the benefit provided by variational quantum circuits, or more complex Ramsey schemes in general, over SSS diminishes substantially as the dephasing strength increases, leading to a behavior akin to that observed for dead time. Similarly, a comparable pattern was reported in [74] for correlated single-atom dephasing.

In contrast, decoherence effects are directly incorporated into the theoretical model in the context of local (frequentist) phase estimation, which makes the resulting optimal Ramsey protocols naturally robust against the respective noise process. In scenarios where single-atom dephasing imposes the dominant noise source, the regime explored in the seminal work by Huelga *et al* in [64] is entered. In the asymptotic limit of large ensembles, SSS are known to be optimal [141], whereas for small ensembles, the optimal interrogation scheme can be efficiently determined numerically by exploiting permutational invariance [142]. Likewise, SSS remain asymptotically optimal for setups limited by spontaneous decay. In contrast, for small ensembles, highly entangled GHZ-like states approach the ultimate bounds, as demonstrated in [65]. Beyond decoherence during the Ramsey dark time, noise affecting the twisting operations has also been considered in [74, 75]. As expected, deeper quantum circuits, which generally require stronger total twisting strengths, exhibit higher susceptibility to noise. Braverman *et al* [143] demonstrates that the potential improvements in clock stability through spin squeezing are significantly reduced if the squeezing is non-unitary. It further highlights that contrast loss during the squeezing process is less detrimental to clock stability than during the Ramsey dark time.

Overall, these observations reinforce the conclusions of this work: Standard protocols with low complexity—employing CSSs, GHZ states and SSSs in particular—achieve stabilities compatible with the ultimate limit across a wide range of scenarios, whereas deeper quantum circuits—generating arbitrary states and measurements—offer a significant advantage only in very specific parameter regimes. This is consistent with a well-established observation in quantum sensing: In general, no single interrogation scheme is universally optimal. Instead, there exists an optimal protocol for a given application and its specific system parameters.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

We thank P Schmidt, T Mehlstäubler, N Huntemann, U Sterr, C Lisdat, R Kaubrügger, D Vasilyev and P Zoller for valuable discussions on the topics of this progress article. We acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project-ID 274200144—SFB 1227 (Projects A06 and A07) and Project-ID 390837967—EXC 2123, and by the Quantum Valley Lower Saxony Q1 project (QVLS- Q1) through the Volkswagen foundation and the ministry for science and culture of Lower Saxony.

Appendix

A.1. Derivations of Bayesian bounds

In the following, we derive the Bayesian bounds summarized in the main text.

A.1.1. BCRB. To derive the BCRB, following [100], we assume standard regularity conditions

$$\sum_x \partial_\phi P(x|\phi) = \partial_\phi \sum_x P(x|\phi) = 0, \quad (\text{A1})$$

where the last equality results from normalization of the conditional probabilities $P(x|\phi)$, and vanishing of the prior at the boundaries

$$\lim_{\phi \rightarrow \pm\infty} \mathcal{P}(\phi) = 0. \quad (\text{A2})$$

By defining the function

$$f(\phi, x) = \sqrt{\mathcal{P}(\phi)P(x|\phi)}[\phi - \phi_{\text{est}}(x)], \quad (\text{A3})$$

the BMSE can be expressed as a squared norm

$$(\Delta\phi)^2 = \int d\phi \sum_x f^2(\phi, x). \quad (\text{A4})$$

Furthermore, we define

$$g(\phi, x) = \frac{1}{\sqrt{\mathcal{P}(\phi)P(x|\phi)}} \frac{d\mathcal{P}(\phi)P(x|\phi)}{d\phi} \quad (\text{A5})$$

such that

$$\int d\phi \sum_x g^2(\phi, x) = \int d\phi \sum_x \frac{1}{\mathcal{P}(\phi)P(x|\phi)} \times \left(\mathcal{P}(\phi) \frac{dP(x|\phi)}{d\phi} + P(x|\phi) \frac{d\mathcal{P}(\phi)}{d\phi} \right)^2 \quad (\text{A6})$$

$$= \int d\phi \mathcal{P}(\phi) \sum_x \frac{1}{P(x|\phi)} \left(\frac{dP(x|\phi)}{d\phi} \right)^2 + \int d\phi \frac{1}{\mathcal{P}(\phi)} \left(\frac{d\mathcal{P}(\phi)}{d\phi} \right)^2 \sum_x P(x|\phi) \quad (\text{A7})$$

$$+ 2 \int d\phi \frac{d\mathcal{P}(\phi)}{d\phi} \sum_x \frac{dP(x|\phi)}{d\phi} \quad (\text{A8})$$

$$= \overline{\mathcal{F}} + \mathcal{I}. \quad (\text{A9})$$

In the last step, we introduced the average Fisher information $\overline{\mathcal{F}}$ and prior knowledge \mathcal{I} defined in equations (8) and (9), respectively. Furthermore, the last term vanishes as a consequence of the regularity condition equation (A1). Using partial integration, equation (A2) and normalization of the probability distributions, we evaluate

$$\int d\phi \sum_x f(\phi, x) g(\phi, x) = \int d\phi \sum_x [\phi - \phi_{\text{est}}(x)] \frac{d\mathcal{P}(\phi) P(x|\phi)}{d\phi} \quad (\text{A10})$$

$$= \left[\sum_x [\phi - \phi_{\text{est}}(x)] \mathcal{P}(\phi) P(x|\phi) \right]_{-\infty}^{+\infty} - \int d\phi \mathcal{P}(\phi) \sum_x P(x|\phi) \quad (\text{A11})$$

$$= -1. \quad (\text{A12})$$

Finally, application of the Cauchy–Schwarz inequality yields

$$\left(\int d\phi \sum_x f(\phi, x) g(\phi, x) \right)^2 \leq \left(\int d\phi \sum_x f^2(\phi, x) \right) \times \left(\int d\phi \sum_x g^2(\phi, x) \right) \quad (\text{A13})$$

which, with the definitions from above, is equivalent to $1 \leq (\Delta\phi)^2 [\overline{\mathcal{F}} + \mathcal{I}]$ and ultimately results in the van Trees inequality equation (7).

A.1.2. BQCRB. To derive the BQCRB, we follow [72]. We start by rewriting the BMSE as

$$(\Delta\phi)^2 = \int d\phi \mathcal{P}(\phi) \sum_x \text{Tr}(\Lambda_{\phi, T}[\rho_{\text{in}}] \Pi_x) [\phi - \phi_{\text{est}}(x)]^2 \quad (\text{A14})$$

$$= (\delta\phi)^2 + \text{Tr} \left(\int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}[\rho_{\text{in}}] \sum_x \Pi_x \phi_{\text{est}}^2(x) \right) - 2\text{Tr} \left(\int d\phi \mathcal{P}(\phi) \phi \Lambda_{\phi, T}[\rho_{\text{in}}] \sum_x \Pi_x \phi_{\text{est}}(x) \right) \quad (\text{A15})$$

$$= (\delta\phi)^2 + \text{Tr}(\overline{\rho} L_2) - 2\text{Tr}(\overline{\rho}' L_1) \quad (\text{A16})$$

where $(\delta\phi)^2$ represents the variance of the prior distribution, $\overline{\rho} = \int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}[\rho_{\text{in}}]$ denotes the average state and $\overline{\rho}' = \int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}[\rho_{\text{in}}] \phi$. Furthermore, we have combined the measurement $\{\Pi_x\}$ and estimator ϕ_{est} by defining the operators $L_1 = \sum_x \Pi_x \phi_{\text{est}}(x)$ and $L_2 = \sum_x \Pi_x \phi_{\text{est}}^2(x)$.

In a first step, following [72, 144, 145], we will demonstrate that, without loss of optimality, the measurement can be restricted to the class of PVMs, i.e. projective von Neumann measurements $\Pi_x = |x\rangle\langle x|$, with orthonormal eigenstates $|x\rangle$, $\langle x|x'\rangle = \delta_{x,x'}$, of the observable X with eigenvalue x . We denote the projective strategy by L_1^{PVM} , where $L_1^{\text{PVM}} = L_1 = \sum_x \phi_{\text{est}}(x) |x\rangle\langle x|$ effectively corresponds to the eigendecomposition. Based on equation (A16), we have to show that $\text{Tr}(\overline{\rho} L_2^{\text{PVM}}) \leq \text{Tr}(\overline{\rho} L_2)$ to prove that we do not lose any optimality by restricting to the projective strategy. Using that L_1 is hermitian and $\Pi_x \geq 0$, we can consider the inequality

$$\sum_x (\phi_{\text{est}}(x) - L_1) \Pi_x (\phi_{\text{est}}(x) - L_1) \geq 0 \quad (\text{A17})$$

$$\sum_x \phi_{\text{est}}^2(x) \Pi_x - \sum_x \phi_{\text{est}}(x) \Pi_x L_1 - L_1 \sum_x \phi_{\text{est}}(x) \Pi_x + L_1 \sum_x \Pi_x L_1 \geq 0 \quad (\text{A18})$$

$$L_2 - L_1^2 \geq 0 \quad (\text{A19})$$

$$L_2 \geq L_1^2, \quad (\text{A20})$$

where we have identified L_1 and L_2 in equation (A18) and used the completeness relation $\sum_x \Pi_x = \mathbb{1}$. However, equality in equation (A20) is achieved specifically for the projective strategy, since $L_2^{\text{PVM}} = \sum_x \phi_{\text{est}}^2(x) |x\rangle\langle x| = (L_1^{\text{PVM}})^2$. Therefore, $\text{Tr}(\overline{\rho} L_2^{\text{PVM}}) \leq \text{Tr}(\overline{\rho} L_2)$ and it is always optimal to choose the measurement to be projective.

In a second step, we derive the BQCRB equation (12). Choosing the projective strategy discussed above and accordingly labeling $L = L_1$ and thus $L_2 = L^2$, the BMSE reads

$$(\Delta\phi)^2 = (\delta\phi)^2 + \text{Tr}(\overline{\rho} L^2) - 2\text{Tr}(\overline{\rho}' L). \quad (\text{A21})$$

Hence, the task of finding the optimal measurement and estimation maps to the optimization of the operator L , containing both. Variation of L according to $L \mapsto L + \epsilon \delta L$ with infinitesimal parameter ϵ and hermitian δL yields

$$(\Delta\phi)^2 = (\delta\phi)^2 + \text{Tr}(\overline{\rho} [L^2 + \epsilon L \delta L + \epsilon \delta L L + \epsilon^2 \delta L^2]) - 2\text{Tr}(\overline{\rho}' [L + \epsilon \delta L]). \quad (\text{A22})$$

Differentiation with respect to ϵ and evaluation at $\epsilon = 0$ results in

$$0 = \text{Tr}([\overline{\rho} L + L \overline{\rho} - 2\overline{\rho}'] \delta L). \quad (\text{A23})$$

Since equation (A23) must hold for any δL , it implies

$$\overline{\rho}' = \frac{1}{2} (L \overline{\rho} + \overline{\rho} L), \quad (\text{A24})$$

reproducing equation (13). Substituting this expression for $\bar{\rho}'$ in equation (A21), we find

$$(\Delta\phi)^2 = (\delta\phi)^2 + \text{Tr}(\bar{\rho}L^2) - 2\text{Tr}(\bar{\rho}'L) \quad (\text{A25})$$

$$= (\delta\phi)^2 + \text{Tr}(\bar{\rho}L^2) - \text{Tr}(\bar{\rho}L^2) - \text{Tr}(L\bar{\rho}L) \quad (\text{A26})$$

$$= (\delta\phi)^2 - \text{Tr}(\bar{\rho}L^2), \quad (\text{A27})$$

which corresponds to the BQCRB equation (12).

Finally, we will express the BQCRB in terms of the QFI of the average state $\bar{\rho}$. Assuming a unitary phase evolution according to equation (5), corresponding to the von Neumann equation

$$\partial_\phi \Lambda_{\phi,T}[\rho_{\text{in}}] = -i[S_z, \Lambda_{\phi,T}[\rho_{\text{in}}]], \quad (\text{A28})$$

and a Gaussian prior distribution as defined in equation (6), we can rewrite $\bar{\rho}'$ as

$$\bar{\rho}' = \int d\phi \mathcal{P}(\phi) \phi \Lambda_{\phi,T}[\rho_{\text{in}}] \quad (\text{A29})$$

$$= -(\delta\phi)^2 \int d\phi (\partial_\phi \mathcal{P}(\phi)) \Lambda_{\phi,T}[\rho_{\text{in}}] \quad (\text{A30})$$

$$= -(\delta\phi)^2 [\mathcal{P}(\phi) \Lambda_{\phi,T}[\rho_{\text{in}}]]_{-\infty}^{+\infty} + (\delta\phi)^2 \int d\phi \mathcal{P}(\phi) \partial_\phi \Lambda_{\phi,T}[\rho_{\text{in}}] \quad (\text{A31})$$

$$= -i(\delta\phi)^2 \left[S_z, \int d\phi \mathcal{P}(\phi) \Lambda_{\phi,T}[\rho_{\text{in}}] \right] \quad (\text{A32})$$

$$= -i(\delta\phi)^2 [S_z, \bar{\rho}] \quad (\text{A33})$$

where we exploited the property $\partial_\phi \mathcal{P}(\phi) = -(\delta\phi)^{-2} \phi \mathcal{P}(\phi)$ of a Gaussian prior distribution. Furthermore, we used partial integration in the second step and equation (A2) as well as equation (A28) in equation (A31). With equations (A24) and (A33), we obtain

$$\frac{1}{2} (L\bar{\rho} + \bar{\rho}L) = -i(\delta\phi)^2 [S_z, \bar{\rho}]. \quad (\text{A34})$$

Substituting $L := (\delta\phi)^2 L_{\text{local}}$, the BQCRB equation (A27) and the implicit equation (A34) become

$$(\Delta\phi)^2 = (\delta\phi)^2 \left[1 - (\delta\phi)^2 \text{Tr}(\bar{\rho}L_{\text{local}}^2) \right] \quad (\text{A35})$$

$$\frac{1}{2} (L_{\text{local}}\bar{\rho} + \bar{\rho}L_{\text{local}}) = -i[S_z, \bar{\rho}]. \quad (\text{A36})$$

Comparison to the QFI approach in local phase estimation shows that L_{local} defines the SLD [144, 146, 147] and thus, $\text{Tr}(\bar{\rho}L_{\text{local}}^2) = \mathcal{F}_Q[\bar{\rho}]$ corresponds to the QFI of the average state $\bar{\rho}$, resulting in equation (14)

$$(\Delta\phi_{\text{BQCRB}})^2 = (\delta\phi)^2 \left[1 - (\delta\phi)^2 \mathcal{F}_Q[\bar{\rho}] \right]. \quad (\text{A37})$$

A.1.3. OQI. The OQI represents the ultimate lower bound of the BMSE. However, no general expressions for arbitrary ensemble sizes exist, but rather complex optimization procedures are required.

Iterative algorithm—In the following, we outline an algorithm introduced in [72, 98], which iteratively optimizes the initial probe state ρ_{in} and measurement $\{\Pi_x\}$. Although numerical optimization becomes challenging with increasing ensemble size, this algorithm enables efficient computation at least for small N . For a given input probe state ρ_{in} , the optimal projective measurement and estimation strategy L can be determined according to the previous discussion on the BQCRB. Conversely, for a given L , the optimal ρ_{in} can be evaluated as follows. Rewriting equation (A16), the BMSE can be expressed as

$$(\Delta\phi)^2 = (\delta\phi)^2 + \text{Tr} \left(\int d\phi \mathcal{P}(\phi) \Lambda_{\phi,T}[\rho_{\text{in}}] (L^2 - 2\phi L) \right). \quad (\text{A38})$$

Defining the adjoint quantum channel $\Lambda_{\phi,T}^\dagger$ via $\text{Tr}(\Lambda_{\phi,T}[\rho]A) = \text{Tr}(\rho\Lambda_{\phi,T}^\dagger[A])$ for an arbitrary operator A , the BMSE becomes

$$(\Delta\phi)^2 = (\delta\phi)^2 + \text{Tr} \left(\rho_{\text{in}} \int d\phi \mathcal{P}(\phi) \Lambda_{\phi,T}^\dagger [L^2 - 2\phi L] \right). \quad (\text{A39})$$

Consequently, the optimal input probe state $\rho_{\text{in}} = |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|$ corresponds to the eigenvector $|\psi_{\text{in}}\rangle$ of the operator $\int d\phi \mathcal{P}(\phi) \Lambda_{\phi,T}^\dagger [L^2 - 2\phi L]$ associated with its most negative eigenvalue. In the iterative algorithm, starting from an arbitrary state, repeatedly the optimal measurement and the corresponding optimal probe state are determined, until the BMSE converges to the OQI.

CTL—In the following, we derive equation (17). Considering a 2π -periodic quantum channel with respect to the phase ϕ as imposed by equation (5), the OQI allows for unambiguous phase estimation within the range $[-\pi, +\pi]$. Exceeding this invertible regime, an estimation error is accumulated which increases with the distance from the primary Ramsey fringe. In particular, an estimation error of $\epsilon_k = (2\pi k)^2$ is accumulated if the phase slips in the region $[-(2k+1)\pi, -(2k-1)\pi]$ or $[(2k-1)\pi, +(2k+1)\pi]$ for $k \in \mathbb{N}$. The estimation error associated with these events can be modeled by

$$\left(\Delta\phi_{\text{CTL}}^{\text{OQI}} \right)^2 = \sum_{k=1}^{\infty} \epsilon_k P_k \quad (\text{A40})$$

which effectively represents the average of the estimation error ϵ_k weighted with its corresponding probability

$$P_k = \int_{-(2k+1)\pi}^{-(2k-1)\pi} d\phi \mathcal{P}(\phi) + \int_{(2k-1)\pi}^{(2k+1)\pi} d\phi \mathcal{P}(\phi). \quad (\text{A41})$$

Consequently, equation (A40) constitutes an asymptotic limit for broad prior distributions. In the context of atomic clocks,

this regime corresponds to long interrogation times, where the coherence time of the LO will become relevant and ultimately limits the clock stability. Therefore, we will denote equation (17) as the CTL of the OQI. Assuming a Gaussian prior distribution as defined in equation (6), the probabilities P_k can be evaluated explicitly to read

$$\begin{aligned} P_k &= 2 \int_{(2k-1)\pi}^{(2k+1)\pi} d\phi \mathcal{P}(\phi) = 2 \int_0^{(2k+1)\pi} d\phi \mathcal{P}(\phi) \\ &\quad - 2 \int_0^{(2k-1)\pi} d\phi \mathcal{P}(\phi) = \operatorname{erf}\left(\frac{(2k+1)\pi}{\sqrt{2}\delta\phi}\right) \\ &\quad - \operatorname{erf}\left(\frac{(2k-1)\pi}{\sqrt{2}\delta\phi}\right) \end{aligned} \quad (\text{A42})$$

where we substituted $t = \frac{\phi}{\sqrt{2}\delta\phi}$ and used the error function $\operatorname{erf}(z) = \int_0^z dre^{-t^2}$. In the relevant regime of prior widths considered in this work, where typically only the adjacent fringes around $\phi = 0$ contribute, the prior distribution $\mathcal{P}(\phi)$ is effectively limited to the region $[-3\pi, +3\pi]$. As a result, the CTL simplifies significantly compared to the general form in equation (A40), which accounts for contributions from all Ramsey fringes. In this restricted regime, the CTL reduces to

$$\begin{aligned} (\Delta\phi_{\text{CTL}}^{\text{OQI}})^2 &= 4\pi^2 \left[\int_{-\infty}^{-\pi} d\phi \mathcal{P}(\phi) + \int_{\pi}^{\infty} d\phi \mathcal{P}(\phi) \right] \\ &= 4\pi^2 \left[1 - \int_{-\pi}^{\pi} d\phi \mathcal{P}(\phi) \right] \\ &= 4\pi^2 \left[1 - \operatorname{erf}\left(\frac{\pi}{\sqrt{2}\delta\phi}\right) \right]. \end{aligned} \quad (\text{A43})$$

Asymptotic scaling (πHL)—With increasing ensemble size, the numerical algorithm presented above becomes computationally challenging. However, in the asymptotic limit ($N \gg 1$), an explicit analytical expression for the OQI can be derived. Assuming unitary phase evolution as described by equation (5) and restricting to the invertible range $[-\pi, +\pi]$, it has been shown for arbitrary prior distributions [95, 99, 101, 102] that the ultimate lower bound is given by πHL , as defined in equation (18). An intuitive derivation for Gaussian prior distributions is given in [99] and is reproduced here. Based on equation (14), the optimization of the BQCRB over all input probe states ρ_{in} is equivalent to evaluating the QFI for the averaged state $\bar{\rho}$. This averaging can be formally associated with a collective dephasing process, where the dephasing rate is identified with the variance of the prior distribution [72, 99]. Combining this perspective with the asymptotic result for collective dephasing in [63], the asymptotic OQI can be expressed as

$$\begin{aligned} (\Delta\phi_{\text{OQI}})^2 &\stackrel{N \gg 1}{\simeq} (\delta\phi)^2 \left[1 - \frac{1}{1 + \frac{\pi^2}{N^2(\delta\phi)^2}} \right] \stackrel{N \gg 1}{\simeq} (\delta\phi)^2 \\ &\quad \times \left[1 - \left(1 - \frac{\pi^2}{N^2(\delta\phi)^2} \right) \right] = \frac{\pi^2}{N^2} \end{aligned} \quad (\text{A44})$$

where we used the expansion $\frac{1}{1+x} \stackrel{x \ll 1}{\simeq} 1 - x$. This result is valid for Gaussian prior distributions with widths $\delta\phi \ll N$, which encompasses all relevant widths in the asymptotic regime $N \gg 1$. It is therefore reasonable to expect that this result generalizes to arbitrary prior distributions, as the fundamental characteristics of the estimation problem in this regime remain largely unaffected by the specific shape of the prior [95, 99].

POI—Finally, we aim to identify the protocol that saturates the asymptotic limit of the OQI. As discussed above, simultaneously determining the optimal measurement, input state, and estimation strategy is a highly non-trivial problem. However, assuming a flat prior distribution and a periodic cost function in the interval $[-\pi, +\pi]$, the concept of covariant measurements [101, 146] provides an explicit solution for the optimal measurement operator, the so-called phase operator [70, 73, 101–107]. The phase operator Φ is defined as

$$\begin{aligned} \Phi &= \sum_{s=-N/2}^{N/2} \phi_s |s\rangle\langle s| \\ \phi_s &= \frac{2\pi s}{N+1} \\ |s\rangle &= \frac{1}{\sqrt{N+1}} \sum_{M=-N/2}^{N/2} e^{-i\phi_s M} |M\rangle \end{aligned} \quad (\text{A45})$$

where ϕ_s are the eigenvalues with corresponding eigenstates $|s\rangle$, constructed from the eigenstates $|M\rangle$ of S_z with eigenvalue M and total spin $N/2$. An interferometer based on Φ is referred to as POI. Furthermore, under these assumptions, the optimal input states in the asymptotic regime ($N \gg 1$), known as *sine* states [70, 73, 101–107] and saturating the πHL , can also be explicitly determined

$$|\psi_{\Phi}\rangle = \sqrt{\frac{2}{N+1}} \sum_{M=-N/2}^{N/2} \sin\left(\frac{\pi(M+1/2)}{N+1}\right) |M\rangle. \quad (\text{A46})$$

However, since the assumptions of a periodic cost function and flat prior distribution are contrary to the framework introduced in section 2.2, namely a global BMSE with phases $-\infty < \phi < +\infty$ and Gaussian prior distributions, these states are not necessarily optimal in the approach pursued in this work. Therefore, the optimal initial states and measurements must be explicitly evaluated. Nevertheless, it is instructive to investigate the performance of the POI and compare to the standard protocols as well as variational classes discussed in the main text. Notably, this scenario is contrary to the BQCRB, since the measurement is fixed by Φ , while we aim to optimize over all initial states. For a fixed prior width, the optimal state $|\psi_{\Phi}\rangle$ can be identified by adapting the iterative algorithm presented above and building on methods from [73]: Starting with an arbitrary initial state $|\psi_{\text{in}}^{(0)}\rangle$, such as $|\psi_{\text{in}}^{(0)}\rangle = |s=0\rangle$, the optimal Bayesian estimator (cf section 2.4) $\phi_{\text{est}}^{(0)}(s)$ is computed. Based on $\phi_{\text{est}}^{(0)}(s)$, the subsequent input probe

state $|\psi_{\text{in}}^{(1)}\rangle$ in the iterative algorithm is evaluated by selecting the eigenstate corresponding to the most negative eigenvalue of the operator $\int d\phi \mathcal{P}(\phi) \Lambda_{\phi, T}^\dagger [L^2 - 2\phi L]$ defined in equation (A39). This ensures that the state $|\psi_{\text{in}}^{(1)}\rangle$ is optimal for a given measurement and estimator. This process is repeated until convergence to the optimal state $|\psi_{\Phi}\rangle$, tailored to the framework considered in this work, is achieved. Numerical evaluation of this iterative algorithm shows that the POI saturates the OQI in the limit of large ensembles within the framework of this work, as discussed in the main text and depicted in figure 2(b).

A.2. Estimators

In the following, we derive explicit expressions for the linear and optimal Bayesian estimators, as presented in section 2.4.

A.2.1. Linear estimator. With the linear estimator defined in equation (20) as $\phi_{\text{est}}^{\text{linear}}(x) = a \cdot x$, the BMSE is expressed as

$$\begin{aligned} (\Delta\phi)^2 &= (\delta\phi)^2 - 2a \int d\phi \mathcal{P}(\phi) \phi \sum_x x P(x|\phi) \\ &\quad + a^2 \int d\phi \mathcal{P}(\phi) \sum_x x^2 P(x|\phi) \end{aligned} \quad (\text{A47})$$

$$\begin{aligned} &= (\delta\phi)^2 - 2a \int d\phi \mathcal{P}(\phi) \phi \langle X(\phi) \rangle \\ &\quad + a^2 \int d\phi \mathcal{P}(\phi) \langle X^2(\phi) \rangle. \end{aligned} \quad (\text{A48})$$

Here, the moments of the observable X are defined by $\langle X^n(\phi) \rangle = \sum_x x^n P(x|\phi)$. The optimal scaling factor a is determined by minimizing the BMSE. Differentiating equation (A48) and solving for a yields

$$a = \frac{\int d\phi \mathcal{P}(\phi) \phi \langle X(\phi) \rangle}{\int d\phi \mathcal{P}(\phi) \langle X^2(\phi) \rangle}. \quad (\text{A49})$$

Hence, from equation (A48), the corresponding BMSE is given by equation (23), i.e.

$$(\Delta\phi)^2 = (\delta\phi)^2 - \frac{[\int d\phi \mathcal{P}(\phi) \phi \langle X(\phi) \rangle]^2}{\int d\phi \mathcal{P}(\phi) \langle X^2(\phi) \rangle}. \quad (\text{A50})$$

Due to the linearity of the estimator, the scaling factor and BMSE only depend on the first and second moments of the observable X , rather than the full statistical model $P(x|\phi)$. This dependence significantly simplifies practical computations, while retaining reliable performance in several situations. Nevertheless, the linear estimation strategy is not optimal in general.

A.2.2. Optimal Bayesian estimator. To start with, we expand equation (4)

$$\begin{aligned} (\Delta\phi)^2 &= \sum_x P(x) \left[\int d\phi P(\phi|x) \phi^2 - 2\phi_{\text{est}}(x) \int d\phi P(\phi|x) \phi \right. \\ &\quad \left. + \phi_{\text{est}}^2(x) \int d\phi P(\phi|x) \right]. \end{aligned} \quad (\text{A51})$$

As before, the first term results in the prior variance $(\delta\phi)^2$, while the last integral simplifies to unity due to the normalization of the posterior distribution. To minimize the BMSE, the optimal Bayesian estimator has to minimize the term in the brackets for each measurement outcome x , since $P(x) \geq 0$ and $\phi_{\text{est}}(x)$ is independent for different x . Differentiation and solving for the estimator yields the optimal Bayesian estimator given in equation (24), $\phi_{\text{est}}^{\text{opt}}(x) = \int d\phi P(\phi|x) \phi$ [70]. Thus, the optimal Bayesian estimator corresponds to the mean posterior phase. With this result, the BMSE becomes

$$(\Delta\phi)^2 = (\delta\phi)^2 - \sum_x P(x) (\phi_{\text{est}}^{\text{opt}}(x))^2. \quad (\text{A52})$$

Equivalently, the BMSE can be expressed in terms of the statistical model $P(x|\phi)$ and prior distribution $\mathcal{P}(\phi)$ according to Bayes theorem equation (3), resulting in equation (25). Unlike the linear estimator, the optimal Bayesian estimator as well as the corresponding BMSE depend explicitly on the statistical model, rather than just the first and second moments of the observable. While this dependence ensures optimality, it also increases computational complexity.

A.3. ADEV

In the following, we characterize the frequency fluctuations $\omega(t) = \omega_0 - \omega_{\text{LO}}(t)$, arising from deviations of the atomic transition frequency ω_0 and LO frequency $\omega_{\text{LO}}(t)$. To facilitate comparisons between LO at different ω_0 , it is convenient to introduce the dimensionless fractional frequency deviation

$$y(t) = \frac{\omega(t)}{\omega_0} = \frac{\omega_0 - \omega_{\text{LO}}(t)}{\omega_0}. \quad (\text{A53})$$

In general, the LO produces a continuous noisy frequency trace $y(t)$. However, in many applications, including the operation of an atomic clock, only a sequence of discrete frequency measurements averaged over individual clock cycles of duration T_C is recorded. We assume that each clock cycle consists of dead time T_D , during which the atomic reference is not interrogated, followed by an interrogation time T . Consequently, the frequency trace is divided into equal intervals of duration $T_C = T_D + T$. The fractional frequency value recorded at the end of cycle k is obtained by

$$y_k = \frac{1}{T_C} \int_{(k-1)T_C}^{kT_C} dt y(t) = \frac{1}{T_C} \left[\int_{(k-1)T_C}^{(k-1)T_C+T_D} dt y(t) + \int_{(k-1)T_C+T_D}^{kT_C} dt y(t) \right]. \quad (\text{A54})$$

A common method to characterizing statistical processes involves calculating the mean value \bar{y} and the variance s^2 , defined by

$$\bar{y} = \frac{1}{n} \sum_{k=1}^n y_k \quad (\text{A55})$$

$$s^2 = \frac{1}{n-1} \sum_{k=1}^n (y_k - \bar{y})^2, \quad (\text{A56})$$

where n denotes the total number of fractional frequency values y_k . However, the standard variance is only a meaningful measure for uncorrelated noise. If the noise is correlated, the deviation from its mean value is no longer stationary [109] and thus, the standard variance might be non-convergent. Consequently, it is not recommended to characterize frequency standards or atomic clocks using the standard variance.

To determine the clock stability, we consider the fractional frequency fluctuations averaged over a total measurement time $\tau = mT_C$, corresponding to m individual clock cycles,

$$\begin{aligned} \bar{y}_j &= \frac{1}{\tau} \int_{(j-1)\tau}^{j\tau} dt y(t) = \frac{1}{\tau} \sum_{k=(j-1)m+1}^{jm} \int_{(k-1)T_C}^{kT_C} dt y(t) \\ &= \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} y_k. \end{aligned} \quad (\text{A57})$$

The most widely used time-domain metric for evaluating the stability of frequency standards and atomic clocks [6, 109–111] is represented by the ADEV [108, 148]. It is defined as the square root of the Allan variance [6, 108]

$$\sigma_y^2(\tau) = \frac{1}{2} \langle (\bar{y}_{j+1} - \bar{y}_j)^2 \rangle, \quad (\text{A58})$$

where $\langle \cdot \rangle$ denotes statistical averaging. While the standard deviation measures the spread of uncorrelated fluctuations around a mean value, the ADEV focuses on differences between consecutive measurements, making it sensitive to time-correlated noise. This distinction allows the ADEV to reveal trends and patterns in the temporal behavior of frequency fluctuations, which the standard deviation would average out or fail to capture. Unlike the standard variance, the Allan variance converges for most types of noise encountered in frequency standards. As it quantifies the fractional frequency fluctuations, a lower value indicates reduced instability, or equivalently, improved stability. The ADEV $\sigma_y(\tau)$ depends on the averaging time τ , providing insights into noise characteristics on different timescales. Short averaging times $\tau \sim T_C$ reveal short-term stability, while large $\tau \gg T_C$

describe long-term stability. Therefore, analyzing the full τ -dependence of $\sigma_y(\tau)$ is essential to assess the performance of different oscillators.

In practice, for finite data sets, the statistical averaging is realized as [109]

$$\sigma_y^2(\tau) = \frac{1}{2(M-1)} \sum_{j=1}^{M-1} (\bar{y}_{j+1} - \bar{y}_j)^2 \quad (\text{A59})$$

where $M = \frac{m}{n}$ represents the number of consecutive frequency intervals with length $\tau = mT_C$. The quantity usually addressed is the square root of the Allan variance, namely the ADEV.

To summarize, the ADEV is an essential metric for assessing the stability of atomic clocks and frequency standards. It provides a measure of fractional frequency fluctuations over various timescales, distinguishing between short-term and long-term stability. Unlike the standard variance, which is only meaningful for uncorrelated noise, the ADEV is well-suited to analyze correlated noise, capturing trends and temporal patterns in the fluctuations. This makes it an indispensable tool for comparing and optimizing the performance of different oscillators.

A.4. Sensitivity of standard protocols

In the following, we derive the sensitivities of the CSS, SSS and GHZ protocols introduced in the main text.

A.4.1. CSSs. For a measurement of the collective spin operator S_y and unitary phase evolution through a rotation around the z -axis, according to equation (5), the first and second moments of the observable are given by

$$\langle X(\phi) \rangle = \langle S_y(\phi) \rangle = \langle S_y \rangle \cos(\phi) + \langle S_x \rangle \sin(\phi) \quad (\text{A60})$$

$$\begin{aligned} \langle X^2(\phi) \rangle &= \langle S_y^2(\phi) \rangle = \langle S_y^2 \rangle \cos^2(\phi) \\ &\quad + \langle S_y S_x + S_x S_y \rangle \sin(\phi) \cos(\phi) + \langle S_x^2 \rangle \sin^2(\phi), \end{aligned} \quad (\text{A61})$$

where the expectation values $\langle \cdot \rangle$ are evaluated with respect to the initial state $|\psi_{\text{in}}\rangle$, independent of the phase ϕ . Assuming a Gaussian prior distribution, as defined in equation (6), the integrals in equation (23) become

$$\begin{aligned} \int d\phi \mathcal{P}(\phi) \phi \langle X(\phi) \rangle &= \int d\phi \mathcal{P}(\phi) \phi [\langle S_y \rangle \cos(\phi) \\ &\quad + \langle S_x \rangle \sin(\phi)] = \langle S_x \rangle (\delta\phi)^2 e^{-(\delta\phi)^2/2} \end{aligned} \quad (\text{A62})$$

$$\begin{aligned} \int d\phi \mathcal{P}(\phi) \langle X^2(\phi) \rangle &= \int d\phi \mathcal{P}(\phi) [\langle S_y^2 \rangle \cos^2(\phi) \\ &\quad + \langle S_y S_x + S_x S_y \rangle \sin(\phi) \cos(\phi) + \langle S_x^2 \rangle \sin^2(\phi)] \end{aligned} \quad (\text{A63})$$

$$= e^{-(\delta\phi)^2} \left[\langle S_y^2 \rangle \cosh((\delta\phi)^2) + \langle S_x^2 \rangle \sinh((\delta\phi)^2) \right], \quad (\text{A64})$$

where terms with odd integrands vanish directly. Thus, the optimal linear scaling factor, corresponding BMSE and effective measurement uncertainty are given by

$$a = \frac{\langle S_x \rangle (\delta\phi)^2 e^{(\delta\phi)^2/2}}{\langle S_y^2 \rangle \cosh((\delta\phi)^2) + \langle S_x^2 \rangle \sinh((\delta\phi)^2)}$$

$$(\Delta\phi)^2 = (\delta\phi)^2 \times \left[1 - (\delta\phi)^2 \frac{\langle S_x \rangle^2}{\langle S_y^2 \rangle \cosh((\delta\phi)^2) + \langle S_x^2 \rangle \sinh((\delta\phi)^2)} \right]$$

$$(\Delta\phi_M)^2 = \frac{\langle S_y^2 \rangle}{\langle S_x \rangle^2} \cosh((\delta\phi)^2) + \frac{\langle S_x^2 \rangle}{\langle S_x \rangle^2} \sinh((\delta\phi)^2) - (\delta\phi)^2. \quad (\text{A65})$$

For the conventional Ramsey protocol, a CSS polarized in x -direction is prepared by a $\pi/2$ -pulse applied to the collective ground state

$$|\text{CSS}\rangle = \mathcal{R}_y(-\frac{\pi}{2}) |\downarrow\rangle^{\otimes N} = |+\rangle^{\otimes N} = \left[\frac{1}{\sqrt{2}} (|\downarrow\rangle + |\uparrow\rangle) \right]^{\otimes N}, \quad (\text{A66})$$

which represents N uncorrelated atoms, each in an equal superposition of the ground and excited state. CSS and their properties are discussed in detail in [112–114]. With expectation values

$$\langle S_x \rangle = \frac{N}{2}, \quad \langle S_y \rangle = \langle S_z \rangle = 0,$$

$$\langle S_x^2 \rangle = \frac{N^2}{4}, \quad \langle S_y^2 \rangle = \frac{N}{4} = \langle S_z^2 \rangle, \quad \langle S_x S_y \rangle = 0 = \langle S_x S_z \rangle, \quad (\text{A67})$$

we derive

$$a_{\text{CSS}} = \frac{2e^{(\delta\phi)^2/2}}{\cosh((\delta\phi)^2) + N \sinh((\delta\phi)^2)} \quad (\text{A68})$$

$$(\Delta\phi_{\text{CSS}})^2 = (\delta\phi)^2 \times \left[1 - (\delta\phi)^2 \frac{N}{\cosh((\delta\phi)^2) + N \sinh((\delta\phi)^2)} \right] \quad (\text{A69})$$

$$(\Delta\phi_M^{\text{CSS}})^2 = \frac{\cosh((\delta\phi)^2)}{N} + \sinh((\delta\phi)^2) - (\delta\phi)^2. \quad (\text{A70})$$

Rewriting the first term, we recover the result from [67]

$$(\Delta\phi_M^{\text{CSS}})^2 = \frac{e^{(\delta\phi)^2}}{N} + \left(1 - \frac{1}{N} \right) \sinh((\delta\phi)^2) - (\delta\phi)^2. \quad (\text{A71})$$

A.4.2. SSSs. The application of an OAT interaction $\mathcal{T}_z(\mu) = \exp(-i\frac{\mu}{2}S_z^2)$ with small twisting strength μ to the CSS, defined in equation (A66), generates a SSS. To align the

minimal spin variance along the y -axis, an additional rotation $\mathcal{R}_x(\theta)$ around the x -axis by an angle θ is applied. Thus, the initial state reads

$$|\text{SSS}\rangle = \mathcal{R}_x(\theta) \mathcal{T}_z(\mu) |\text{CSS}\rangle. \quad (\text{A72})$$

These states are introduced and discussed in detail in [87]. In comparison to CSS, the SSS differs primarily in its polarization and spin variances, while other properties remain unchanged. Hence, the optimal linear scaling factor, BMSE and effective measurement uncertainty are given by equation (A65) with expectation values

$$\langle S_x \rangle = \frac{N}{2} \cos^{N-1}\left(\frac{\mu}{2}\right) \quad (\text{A73})$$

$$\langle S_x^2 \rangle = \frac{N}{4} \left\{ N \left[1 - \cos^{2N-2}\left(\frac{\mu}{2}\right) \right] - \frac{1}{2} (N-1) A \right\} + \langle S_x \rangle^2 \quad (\text{A74})$$

$$\langle S_y^2 \rangle = \frac{N}{4} \left\{ 1 + \frac{1}{4} (N-1) \left[A - \sqrt{A^2 + B^2} \right] \right\}, \quad (\text{A75})$$

where $A = 1 - \cos^{N-2}(\mu)$ and $B = 4 \sin\left(\frac{\mu}{2}\right) \cos^{N-2}\left(\frac{\mu}{2}\right)$.

A.4.3. GHZ states. The GHZ state [115] is defined by

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} [|\downarrow\rangle^{\otimes N} + |\uparrow\rangle^{\otimes N}], \quad (\text{A76})$$

which represents an equal superposition of the collective ground and excited states and thus, is maximally entangled. After the free evolution, the state reads

$$|\psi_\phi\rangle = \mathcal{R}_z(-\frac{\pi}{2N}) \mathcal{R}_z(\phi) |\text{GHZ}\rangle = \frac{1}{\sqrt{2}} \left[e^{i\frac{N}{2}\phi - i\frac{\pi}{4}} |\downarrow\rangle^{\otimes N} + e^{-i\frac{N}{2}\phi + i\frac{\pi}{4}} |\uparrow\rangle^{\otimes N} \right], \quad (\text{A77})$$

where the additional rotation $\mathcal{R}_z(-\frac{\pi}{2N})$ is applied to shift the optimal working point to $\phi_0 = 0$, since the prior is centered around $\phi = 0$. Equivalently, the prior distribution could be shifted by $\pi/2N$. The expectation value of the parity $\Pi = (-1)^N \sigma_x^{\otimes N}$ is given by

$$\langle \Pi(\phi) \rangle = (-1)^N \sin(N\phi). \quad (\text{A78})$$

Since $\sigma_x^2 = \mathbb{1}$, the second moment directly yields $\langle \Pi^2(\phi) \rangle = 1$. Hence, the integrals in equation (23) become

$$\int d\phi \mathcal{P}(\phi) \phi \langle X(\phi) \rangle = (-1)^N \int d\phi \mathcal{P}(\phi) \phi \sin(N\phi) = (-1)^N N (\delta\phi)^2 e^{-N^2(\delta\phi)^2/2} \quad (\text{A79})$$

$$\int d\phi \mathcal{P}(\phi) \langle X^2(\phi) \rangle = 1. \quad (\text{A80})$$

Consequently, the corresponding optimal linear scaling factor, BMSE and effective measurement uncertainty are given by

$$\begin{aligned}
a_{\text{GHZ}} &= (-1)^N N (\delta\phi)^2 e^{-N^2(\delta\phi)^2/2} \\
(\Delta\phi_{\text{GHZ}})^2 &= (\delta\phi)^2 \left[1 - N^2 (\delta\phi)^2 e^{-N^2(\delta\phi)^2/2} \right] \\
(\Delta\phi_M^{\text{GHZ}})^2 &= \frac{e^{N^2(\delta\phi)^2}}{N^2} - (\delta\phi)^2. \tag{A81}
\end{aligned}$$

Due to the binary nature of the parity measurement, the linear estimator is already optimal and thus, saturates the BCRB and coincides with the optimal Bayesian estimator.

However, the parity measurement can also be mimicked by a projective spin measurement and application of the corresponding optimal Bayesian estimator: for N even, a Ramsey pulse is applied after the free evolution time, implemented by a rotation of $\pi/2$ around the x -axis, resulting in the final state

$$\begin{aligned}
|\psi_f\rangle &= \mathcal{R}_x\left(\frac{\pi}{2}\right)|\psi_\phi\rangle = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2^N}} \left[e^{i\frac{N}{2}\phi - i\frac{\pi}{4}} (|\downarrow\rangle - i|\uparrow\rangle)^{\otimes N} \right. \\
&\quad \left. + e^{-i\frac{N}{2}\phi + i\frac{\pi}{4}} (|\uparrow\rangle - i|\downarrow\rangle)^{\otimes N} \right]. \tag{A82}
\end{aligned}$$

For N odd, calculations are analogous with final rotation around the y -axis. Finally, a projective measurement of S_z is performed. Note that the final Ramsey pulse can equivalently be absorbed in the observable, leading to an effective measurement of S_y , as for the CSS and SSS protocol. The conditional probabilities are evaluated to read

$$P(x = +\frac{N}{2} - N_- | \phi) = \frac{1}{2^N} \binom{N}{N_-} \left[1 + (-1)^{\frac{N}{2} + N_-} \sin(N\phi) \right] \tag{A83}$$

where N_- denotes the number of atoms in the ground state. Interestingly, the conditional probabilities for N_- and $N - N_-$ are equal (since N is even), resulting in a vanishing signal $\langle X(\phi) \rangle \equiv 0$. Nevertheless, with

$$\begin{aligned}
P(x = +\frac{N}{2} - N_-) &= \int d\phi P(x|\phi) \mathcal{P}(\phi) \\
&= \frac{1}{2^N} \binom{N}{N_-}, \tag{A84}
\end{aligned}$$

the optimal Bayesian estimator is given by

$$\begin{aligned}
\phi_{\text{est}}(x = +\frac{N}{2} - N_-) &= \frac{1}{P(x)} \int d\phi P(x|\phi) \mathcal{P}(\phi) \phi \\
&= (-1)^{\frac{N}{2} + N_-} N (\delta\phi)^2 e^{-N(\delta\phi)^2/2}. \tag{A85}
\end{aligned}$$

and an efficient estimation is possible. Consequently, the optimal estimation strategy distinguishes between even and odd numbers of atoms in the ground state and thus, effectively mimics a parity measurement.

As a final step, we determine the BQCRB for the GHZ state. Using equation (A77), we find

$$\begin{aligned}
\Lambda_{\phi,T}[\rho_{\text{in}}] &= \frac{1}{2} \left[|\downarrow\rangle\langle\downarrow|^{\otimes N} + e^{iN\phi} |\downarrow\rangle\langle\uparrow|^{\otimes N} + e^{-iN\phi} |\uparrow\rangle\langle\downarrow|^{\otimes N} \right. \\
&\quad \left. + |\uparrow\rangle\langle\uparrow|^{\otimes N} \right], \tag{A86}
\end{aligned}$$

which leads to the average state

$$\begin{aligned}
\bar{\rho} &= \frac{1}{2} \left[|\downarrow\rangle\langle\downarrow|^{\otimes N} + e^{-N^2(\delta\phi)^2/2} \right. \\
&\quad \left. (|\downarrow\rangle\langle\uparrow|^{\otimes N} + |\uparrow\rangle\langle\downarrow|^{\otimes N}) + |\uparrow\rangle\langle\uparrow|^{\otimes N} \right]. \tag{A87}
\end{aligned}$$

Interestingly, equation (A87) is no longer pure due to the averaging and effectively corresponds to a real 2×2 -matrix. Hence, using equation (14), the BQCRB directly follows from

$$\mathcal{F}_Q[\bar{\rho}] = N^2 e^{-N^2(\delta\phi)^2} \tag{A88}$$

and results in the same value as in equation (A81).

A.5. Numerical methods

This section presents the numerical methods employed in this work, including the optimization of the variational quantum circuits, the simulation of the full feedback loop in an atomic clock, the iterative determination of the prior width and the incorporation of dead time noise into the prior phase distribution.

A.5.1. Optimization. The $2 + 3(n + m)$ variational parameters of the quantum circuits introduced in section 3.2 are optimized using the *scipy* library in Python [149]. Due to the vast number of local minima, a global optimization method is required. Specifically, we employ a *differential evolution* approach. As discussed in section 4, for fixed interrogation time T and class $[n, m]$, multiple local minima are identified in different regions to mitigate limitations caused by fringe hops. Following [68, 75], figure A1(a) illustrates a generic optimization landscape in the $\mu_1 - \mu_2$ -plane for the $[1, 1]$ -class with $N = 8$ at an interrogation time of $T/Z = 0.1$. The $\mu_1 - \mu_2$ -plane is divided into four quadrants (I–IV), with regimes of strong twistings (V–VII) considered separately. Notably, region (VII) is effectively equivalent to (VI) due to the periodic nature of OAT interactions. For variational protocols $[1, m]$ with deeper circuits ($m = 2, 3$), the twisting strength landscape expands and is divided into eight quadrants considering the twisting strengths μ_j . Local minima identified in these landscapes, represented by distinct symbols, correspond to specific protocols simulated within the full feedback loop of an atomic clock (see section 4.1).

A.5.2. Ramsey signals. The standard protocols, namely CSS, SSS and GHZ protocols, exhibit sinusoidal signals. While CSS and SSS have a dynamic range of $[-\pi/2, +\pi/2]$, allowing for unbiased phase estimation within this interval, the phase is imprinted N times faster for the GHZ state, leading to a correspondingly N times smaller dynamic range. In contrast, the application of variational quantum circuits with multiple layers for state preparation and measurement can generate arbitrary signals. Within the variational classes introduced in section 3.2, there are no restrictions on the geometry, and therefore the signal shape is not constrained. While $[n, 0]$ -protocols yield collective spin measurements that result

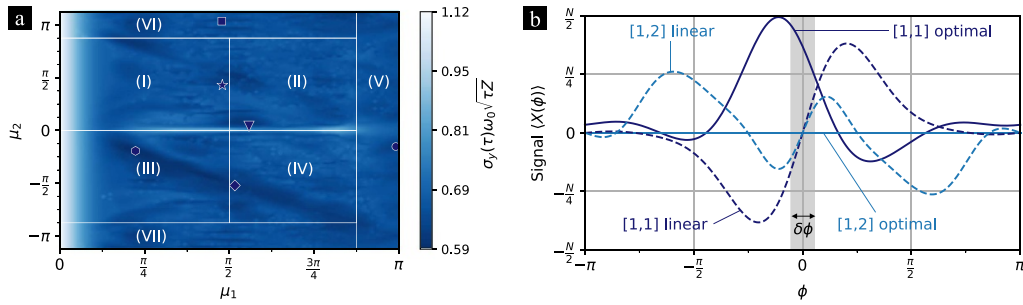


Figure A1. (a) Optimization landscape in the μ_1 - μ_2 -plane for the $[1, 1]$ class with $N = 8$ at an interrogation time of $T/Z = 0.1$. The Allan deviation $\sigma_y(\tau)$ is rescaled by the averaging time τ , coherence time Z and transition frequency ω_0 . Darker areas correspond to better stability. The optimization areas (I–VII) are separated by white lines, while local minima within these regions are illustrates by symbols. In theory, the lowest instability is achieved by the protocol indicated by the hexagon in area (III), while the other local minima result in a comparable clock stability. (b) Comparison of signals $\langle X(\phi) \rangle$ of the optimal $[1, 1]$ and $[1, 2]$ protocols for the linear (dashed) and optimal Bayesian estimator (solid), with $N = 8$ at an interrogation time of $T/Z = 0.1$. The gray shaded region represents the spread of the prior distribution, with its width $\delta\phi$ corresponding to the specific interrogation time.

in sinusoidal signals, increasing m allows for arbitrary signal shapes. Typically, when $m \neq 0$, highly non-sinusoidal signals are generated, where the dynamic range adapts to the interrogation time and the associated frequency fluctuations of the LO.

Generic Ramsey signals for the $[1, 1]$ and $[1, 2]$ protocols are illustrated in figure A1(b) for $N = 8$ at interrogation time $T/Z = 0.1$. The signals, associated with the optimal variational parameters, are compared for both the linear and optimal Bayesian estimation strategies. Specifically, the signal for the $[1, 1]$ -class with the optimal Bayesian estimator corresponds to the hexagon in figure A1(a).

In principle, the estimation strategy does not affect the signal directly, as it is determined solely by the initial state, phase imprint, and measurement. However, the choice of the estimator influences the optimization of the variational parameters, which in turn affects the signal. Consequently, the linear estimation strategy typically results in anti-symmetric signals, at least within the range of the prior distribution. In contrast, the optimal Bayesian estimator can become highly non-linear. As a result, these signals often exhibit strongly non-sinusoidal shapes, lacking symmetry and any apparent relation to the phase. While this may initially seem counterintuitive, this approach achieves lower phase estimation uncertainty when combined with the corresponding estimator, as discussed in section 3.2. An example is presented in section 3.1 for the GHZ protocol with projective spin measurement, where the estimator effectively mimics a parity measurement, while the signal itself vanishes. A similar behavior is observed for the $[1, 2]$ protocol with optimal Bayesian estimator in figure A1(b).

A.5.3. Clock simulation. This section provides a brief overview of the methods used to simulate an atomic clock. The Monte Carlo simulation of the full feedback loop builds on [67, 68]. Throughout this work, we assume clock operations with identical interrogation sequences in each clock cycle. Specifically, the interrogation time T , dead time T_D and

Ramsey protocol are fixed in a single clock run. Consequently, the trace of average frequency deviations $\{\omega_k\}$, where k denotes the index of the clock cycle, can be generated in advance for a given spectral noise density $S_y(f)$ or LO ADEV $\sigma_{y,LO}(\tau)$. In principle, the frequency traces can be obtained by Fourier transformation of the corresponding noise in the Fourier frequency domain. However, this approach becomes computationally expensive for long traces with length $n \gg 1$. While white and random walk frequency noise can be generated using standard methods [109], flicker frequency noise is efficiently generated by a sum of multiple damped random walks [67]. Given the frequency trace $\{\omega_k\}$, the clock operation is simulated by implementing the full feedback loop, closely following the basic clock operation outlined in section 2.1. In each clock cycle k , the accumulated phase $\phi_k = \omega_k T$ is used to evaluate the conditional probabilities $P(x|\phi_k)$. A random measurement outcome x_k is then sampled based on this statistics. Finally, the frequency is corrected by the servo according to $\omega_{\text{corr},k}$ inferred from the phase estimate $\phi_{\text{est}}(x_k)$. The servo is implemented by a general linear predictor, presented in [67], taking into account 50 previous frequency estimates. This procedure is repeated in each clock cycle for $n = 10^7$ cycles in a single clock run. Afterwards, the clock stability is determined for the stabilized frequency trace.

Although the ADEV $\sigma_y(\tau)$ depends on the total averaging time τ , clock stability is typically characterized by a single value, assuming that the ADEV scales as $\sim 1/\sqrt{\tau}$ for $\tau \gg 1$ s (cf equation (29)). Conventionally, this is the ADEV at unit averaging time $\sigma_y(\tau = 1$ s). Equivalently, the ADEV can be rescaled with the total averaging time, i.e. $\sigma_y(\tau)\sqrt{\tau}$. However, the delayed feedback in a clock operation leads to a large deviation of theoretical prediction and simulation or experiment at $\tau \sim 1$ s, while the long term stability $\sigma_y(\tau) \propto 1/\sqrt{\tau}$ is achieved for $\tau \gg 1$ s. Therefore, the stability is evaluated at $\tau \gg 1$ and extrapolated to its hypothetical value at $\tau = 1$ s based on the scaling $\sigma_y(\tau) \propto 1/\sqrt{\tau}$.

Since clock operation involves stochastic processes, such as frequency trace generation and measurement sampling, results

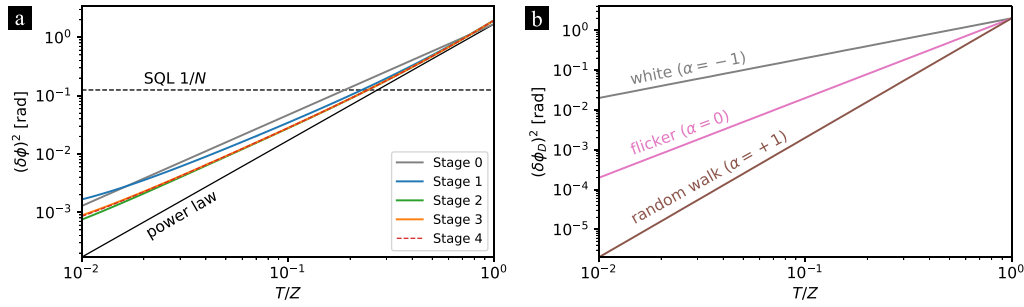


Figure A2. (a) Convergence of the prior variance $(\delta\phi)^2$ in the iterative approach for $N = 8$. The distinct iteration stages are illustrated by colored lines. Additionally, the power law scaling equation (31) and the SQL $1/N$ are shown for comparison. (b) Additional noise due to dead time characterized by the associated prior variance $(\delta\phi_D)^2$ given in equation (43) for white (gray), flicker (pink) and random walk frequency noise (brown). Mean values are averaged over 10 independent runs.

vary across different clock runs. To ensure robust stability estimates, each setup, defined by fixed ensemble size N , interrogation time T , dead time T_D , and Ramsey sequence, is simulated over 10 independent clock runs. Symbols in the corresponding figures therefore represent mean values, while error bars indicate standard deviations. To include a protocol in the results, we impose the stringent criterion that no fringe hops occur across 10^8 total clock cycles, since even a single fringe hop results in a complete loss of clock stability.

Fringe hops represent a substantial limitation at long interrogation times, where they can dominate even over the CTL for small ensembles. Additionally, fringe hops arise for the variational protocols at the plateau of the OQI (cf section 4.1). To address this, we simulate several protocols from the optimization (see above) for a given interrogation time T and variational class $[1, m]$. As a result, the theoretically optimal protocol may be limited by fringe hops in the full feedback loop and thus, a different protocol performs best in simulations, leading to substantial deviations between theoretical predictions and numerical simulations. To give an example, the variational protocol associated with the hexagon in area (III) in figure A1 achieves the lowest ADEV in theory, while it is limited by fringe hops in the numerically simulated full feedback loop. Instead, the protocol corresponding to the circle in area (V) performs best in numerical simulations, resulting in a significant deviation to theoretical prediction. In extreme cases, all simulated protocols may suffer from fringe hops, and no data points are shown at these interrogation times. However, the susceptibility to fringe hops at the OQI plateau decreases with increasing circuit complexity m (cf section 4.1). Thus, simulation results for protocols with the lowest depth $m \leq 3$ that are not limited by fringe hops are shown.

A.5.4. Iterative prior width. Equation (31) provides a good approximation of the prior phase width $\delta\phi$ in the regime of large ensembles N and long interrogation times T , as demonstrated in [67, 68, 73]. However, as discussed in the main text, the prior width $\delta\phi$ and estimation error $\Delta\phi$ mutually influence each other in the full feedback loop of an atomic clock. Moreover, any model of the prior width can only capture the

true residual noise to a certain degree. Consequently, an on-device optimization, as utilized in [76], would most accurately reflect the experimental conditions and thus, yield the best results. However, this approach has several disadvantages. First, it precludes theoretical predictions and *ab initio* studies of clock stability, making it impossible to exclude protocols prone to fringe hops, for instance. Second, it is exceptionally demanding in terms of experimental time. While the variational parameters need only be optimized for individual clock runs, evaluating the ADEV as a cost function requires a sufficiently long averaging time τ for each optimization step to achieve the long-term scaling according to $1/\sqrt{\tau}$. Unlike Bayesian phase estimation, which can focus on single interrogation cycles, on-device optimization for clock stability must account for time-varying frequency deviations ω across cycles. As a result, on-device optimization using the ADEV as a cost function is impractical.

To overcome these challenges, we focus on modeling the prior knowledge according to equation (6) and iteratively adjust the prior width $\delta\phi$ to account for the closed feedback loop dynamics. The general strategy involves simulating the full feedback loop multiple times and using the results from previous simulations to estimate the prior width for the subsequent iteration stage. This procedure is repeated until convergence is achieved. In each iteration stage, the frequency deviation ω at the end of the Ramsey dark time is recorded and interpolated as a function of the interrogation time at a fixed ensemble size. Directly applying this iterative method to the variational protocols would lead to the same issues discussed above. Therefore, it is advantageous to use fixed and reliable protocols, such as CSS and SSS, to ensure consistency. Additionally, comparing results across protocols would be cumbersome, as each protocol yields a distinct prior width and corresponding OQI. Instead, we approximate the prior width $\delta\phi$ for a fixed ensemble size through the following iterative stages:

- **Stage 0 (Initialization):** Start with a heuristic prior width, where $\delta\phi$ is interpolated linearly on a log-log scale between $(\delta\phi)^2 = (T/Z)^{4/3}N^{-1/4}$ for $T/Z = 0.01$ and the value given by equation (31) for $T/Z = 1$. Using this prior width,

- simulate the CSS protocol with the optimal Bayesian estimator and record the resulting frequency deviations $\{\omega_k\}$.
- Stage 1 (Refinement): Use the recorded $\{\omega_k\}$ from the previous simulation to determine the corresponding prior phase distribution. Fit this distribution to a Gaussian as described in equation (6) to obtain an updated prior width $\delta\phi$. Plot $\delta\phi$ as a function of interrogation time and fit it with a fifth-order polynomial. Exclude prior widths for interrogation times where fringe hops limit stability and additionally add the value from equation (31) at $T/Z = 1$. Simulate the SSS protocol with the updated prior width.
 - Stages 2, 3, ... (Iteration): Repeat the refinement process.

Convergence is typically achieved after stage 3, even for small ensembles, as the prior width from stage 4 introduces only negligible adjustments. This convergence is generically illustrated in figure A2(a). Hence, the prior width from stage 3 is adopted to model a realistic atomic clock scenario used in sections 4 and 5.3. While this iterative approach provides a reasonable approximation of the closed feedback loop dynamics, it remains a simplification. Consequently, deviations between theoretical predictions and numerical simulations may still arise, as discussed above and in the main text.

A.5.5. Prior width with dead time. As discussed in the main text, the additional noise introduced during dead time can be approximated as white noise in the asymptotic limit of many clock cycles. The corresponding prior width $\delta\phi_D$ is determined by simulating the uncorrected frequency trace of the LO and quantifying the noise accumulated during the cycle duration T_D . Specifically, the new frequency deviations $\omega_{\text{new},k} = \omega_k - \omega_{k-1}$ are recorded for each cycle k , representing the differences between consecutive cycles. Using the recorded $\omega_{\text{new},k}$, the phase distribution associated with a hypothetical phase shift during T_D is evaluated, and the corresponding prior width $\delta\phi_D$ is extracted. Simulations confirm the power-law scaling predicted by equation (43), as illustrated in figure A2(b).

ORCID iDs

T Kielinski  0009-0002-4694-7593

K Hammerer  0000-0002-7179-0666

References

- [1] Pezzè L, Smerzi A, Oberthaler M K, Schmied R and Treutlein P 2018 Quantum metrology with nonclassical states of atomic ensembles *Rev. Mod. Phys.* **90** 035005
- [2] Ye J and Zoller P 2024 Essay: quantum sensing with atomic, molecular and optical platforms for fundamental physics *Phys. Rev. Lett.* **132** 190001
- [3] Lombardi M A 2003 Time and frequency *Encyclopedia of Physical Science and Technology* (Elsevier) pp 783–801
- [4] Levine J 1999 Introduction to time and frequency metrology *Rev. Sci. Instrum.* **70** 2567–96
- [5] Audoin C and Guinot B 2001 *The Measurement of Time* (Cambridge University Press)
- [6] Riehle F 2003 *Frequency Standards: Basics and Applications* (Wiley) (<https://doi.org/10.1002/3527605991>)
- [7] Ludlow A D, Boyd M M, Ye J, Peik E and Schmidt P O 2015 Optical atomic clocks *Rev. Mod. Phys.* **87** 637
- [8] Colombo S, Pedrozo-Peñañiel E and Vuletić V 2022 Entanglement-enhanced optical atomic clocks *Appl. Phys. Lett.* **121** 210502
- [9] Udem T, Holzwarth R and Hänsch T W 2002 Optical frequency metrology *Nature* **416** 233–7
- [10] Ye J and Cundiff S T 2005 *Femtosecond Optical Frequency Comb: Principle, Operation and Applications* (Kluwer Academic Publishers) (<https://doi.org/10.1007/b102450>)
- [11] Steinel M, Shao H, Filzinger M, Lipphardt B, Brinkmann M, Didier A, Mehlstäubler T, Lindvall T, Peik E and Huntemann N 2023 Evaluation of a $^{88}\text{Sr}^+$ optical clock with a direct measurement of the blackbody radiation shift and determination of the clock frequency *Phys. Rev. Lett.* **131** 1079
- [12] Pelzer L, Dietze K, Martínez-Lahuerta V J, Krinner L, Kramer J, Dawel F, Spethmann N C H, Hammerer K and Schmidt P O 2024 Multi-ion frequency reference using dynamical decoupling *Phys. Rev. Lett.* **133** 033203
- [13] Hauser H N *et al* 2025 $^{115}\text{In}^+ - ^{172}\text{Yb}^+$ coulomb crystal clock with 2.5×10^{-18} systematic uncertainty *Phys. Rev. Lett.* **134** 023201
- [14] Madjarov I S *et al* 2019 An atomic-array optical clock with single-atom readout *Phys. Rev. X* **9** 041052
- [15] Norcia M A, Young D J, Eckner W J, Anderegg L, Schkolnik V and Kaufman A M 2019 Second-scale coherence on an optical clock transition in a tweezer array *Science* **366** 93
- [16] Young A W, Eckner W J, Milner W R, Kedar D, Norcia M A, Oelker E, Schine N, Ye J and Kaufman A M 2020 Half-minute-scale atomic coherence and high relative stability in a tweezer clock *Nature* **588** 408–13
- [17] Shaw A L, Finkelstein R, Tsai R B-S, Scholl P, Yoon T H, Choi J and Endres M 2024 Multi-ensemble metrology by programming local rotations with atom movements *Nat. Phys.* **20** 195
- [18] Takamoto M *et al* 2015 Frequency ratios of Sr, Yb and Hg based optical lattice clocks and their applications *C. R. Phys.* **16** 489
- [19] Katori H, Takamoto M, Pal'chikov V G and Ovsiannikov V D 2003 Ultrastable optical clock with neutral atoms in an engineered light shift trap *Phys. Rev. Lett.* **91** 173005
- [20] Al-Masoudi A, Dörscher S, Häfner S, Sterr U and Lisdat C 2015 Noise and instability of an optical lattice clock *Phys. Rev. A* **92** 063814
- [21] Katori H 2011 Optical lattice clocks and quantum metrology *Nat. Photon.* **5** 203–10
- [22] Abdel-Hafiz M *et al* 2019 Guidelines for developing optical clocks with 10^{-18} fractional frequency uncertainty (arXiv:1906.11495)
- [23] Oelker E *et al* 2019 Demonstration of 4.8×10^{-17} stability at 1s for two independent optical clocks *Nat. Photon.* **13** 714–9
- [24] Schioppo M *et al* 2016 Ultrastable optical clock with two cold-atom ensembles *Nat. Photon.* **11** 48–52
- [25] Nicholson T *et al* 2015 Systematic evaluation of an atomic clock at 2×10^{-18} total uncertainty *Nat. Commun.* **6** 6896
- [26] Bloom B J, Nicholson T L, Williams J R, Campbell S L, Bishof M, Zhang X, Zhang W, Bromley S L and Ye J 2014 An optical lattice clock with accuracy and stability at the 10^{-18} level *Nature* **506** 71–75
- [27] Pedrozo-Peñañiel E *et al* 2020 Entanglement on an optical atomic-clock transition *Nature* **588** 414–8
- [28] Aeppli A, Kim K, Warfield W, Safronova M S and Ye J 2024 Clock with 8×10^{-19} systematic uncertainty *Phys. Rev. Lett.* **133** 023401

- [29] Keller J, Kalincev D, Burgermeister T, Kulosa A P, Didier A, Nordmann T, Kiethe J and Mehlstäubler T 2019 Probing time dilation in coulomb crystals in a high-precision ion trap *Phys. Rev. Appl.* **11** 011002
- [30] Aharon N, Spethmann N, Leroux I D, Schmidt P O and Retzker A 2019 Robust optical clock transitions in trapped ions using dynamical decoupling *New J. Phys.* **21** 083040
- [31] Sanner C, Huntemann N, Lange R, Tamm C, Peik E, Safronova M S and Porsev S G 2019 Optical clock comparison for Lorentz symmetry testing *Nature* **567** 204–8
- [32] Huang Y, Guan H, Bian W, Ma L, Liang K, Li T and Gao K 2017 A comparison of two $^{40}\text{Ca}^+$ single-ion optical frequency standards at the 5×10^{-17} level and an evaluation of systematic shifts *Appl. Phys. B* **123** 166
- [33] Dimarcq N *et al* 2024 Roadmap towards the redefinition of the second *Metrologia* **61** 012001
- [34] Takamoto M, Ushijima I, Ohmae N, Yahagi T, Kokado K, Shinkai H and Katori H 2020 Test of general relativity by a pair of transportable optical lattice clocks *Nat. Photon.* **14** 411–5
- [35] Chou C W, Hume D B, Rosenband T and Wineland D J 2010 Optical clocks and relativity *Science* **329** 1630–3
- [36] Bothwell T, Kennedy C J, Aeppli A, Kedar D, Robinson J M, Oelker E, Staron A and Ye J 2022 Resolving the gravitational redshift across a millimetre-scale atomic sample *Nature* **602** 420–4
- [37] Dreissen L S, Yeh C-H, Fürst H A, Grensemann K C and Mehlstäubler T E 2022 Improved bounds on Lorentz violation from composite pulse Ramsey spectroscopy in a trapped ion *Nat. Commun.* **13** 7314
- [38] Safronova M S 2019 The search for variation of fundamental constants with clocks *Ann. Phys., Lpz.* **531** 1800364
- [39] Roberts B M *et al* 2020 Search for transient variations of the fine structure constant and dark matter using fiber-linked optical atomic clocks *New J. Phys.* **22** 093010
- [40] Derevianko A and Pospelov M 2014 Hunting for topological dark matter with atomic clocks *Nat. Phys.* **10** 933–6
- [41] Safronova M S, Budker D, DeMille D, Kimball D F J, Derevianko A and Clark C W 2018 Search for new physics with atoms and molecules *Rev. Mod. Phys.* **90** 025008
- [42] Kennedy C J, Oelker E, Robinson J M, Bothwell T, Kedar D, Milner W R, Marti G E, Derevianko A and Ye J 2020 Precision metrology meets cosmology: improved constraints on ultralight dark matter from atom-cavity frequency comparisons *Phys. Rev. Lett.* **125** 201302
- [43] Schuldt T *et al* 2021 Optical clock technologies for global navigation satellite systems *GPS Solut.* **25** 83
- [44] Schmidt T D, Schuldt T, Michalak G, Surof J, Poliak J, Giorgi G, Braxmaier C, Meurer M and Günther C 2023 Optical technologies for future global navigation satellite systems 2023 *IEEE/ION Position, Location and Navigation Symp. (PLANS)* (IEEE) (<https://doi.org/10.1109/plans53410.2023.10140082>)
- [45] Droste S, Ozimek F, Udem T, Predehl K, Hänsch T W, Schnatz H, Grosche G and Holzwarth R 2013 Optical-frequency transfer over a single-span 1840 km fiber link *Phys. Rev. Lett.* **111** 110801
- [46] Lisdat C *et al* 2016 A clock network for geodesy and fundamental science *Nat. Commun.* **7** 12443
- [47] Grotti J *et al* 2018 Geodesy and metrology with a transportable optical clock *Nat. Phys.* **14** 437–41
- [48] Mehlstäubler T E, Grosche G, Lisdat C, Schmidt P O and Denker H 2018 Atomic clocks for geodesy *Rep. Prog. Phys.* **81** 064401
- [49] Grotti J *et al* 2024 Long-distance chronometric leveling with a portable optical clock *Phys. Rev. Appl.* **21** L061001
- [50] Itano W M, Bergquist J C, Bollinger J J, Gilligan J M, Heinzen D J, Moore F L, Raizen M G and Wineland D J 1993 Quantum projection noise: population fluctuations in two-level systems *Phys. Rev. A* **47** 3554
- [51] Gardiner C W and Zoller P 2004 *Quantum Noise (Springer Series in Synergetics)* 3rd edn (Springer)
- [52] Pezzé L and Smerzi A 2009 Entanglement, nonlinear dynamics and the Heisenberg limit *Phys. Rev. Lett.* **102** 100401
- [53] Wineland D J, Bollinger J J, Itano W M, Moore F L and Heinzen D J 1992 Spin squeezing and reduced quantum noise in spectroscopy *Phys. Rev. A* **46** R6797
- [54] Wineland D J, Bollinger J J, Itano W M and Heinzen D J 1994 Squeezed atomic states and projection noise in spectroscopy *Phys. Rev. A* **50** 67
- [55] Franke J, Muleady S R, Kautbruegger R, Kranzl F, Blatt R, Rey A M, Joshi M K and Roos C F 2023 Quantum-enhanced sensing on optical transitions through finite-range interactions *Nature* **621** 740–5
- [56] Robinson J M, Miklos M, Tso Y M, Kennedy C J, Bothwell T, Kedar D, Thompson J K and Ye J 2024 Direct comparison of two spin-squeezed optical clock ensembles at the 10^{-17} level *Nat. Phys.* **20** 208–13
- [57] Eckner W J, Darkwah Oppong N, Cao A, Young A W, Milner W R, Robinson J M, Ye J and Kaufman A M 2023 Realizing spin squeezing with Rydberg interactions in an optical clock *Nature* **621** 734–9
- [58] Finkelstein R, Tsai R B-S, Sun X, Scholl P, Direkci S, Gefen T, Choi J, Shaw A L and Endres M 2024 Universal quantum operations and ancilla-based read-out for tweezer clocks *Nature* **634** 321–7
- [59] Cao A *et al* 2024 Multi-qubit gates and Schrödinger cat states in an optical clock *Nature* **634** 315–20
- [60] Shaji A and Caves C M 2007 Qubit metrology and decoherence *Phys. Rev. A* **76** 032111
- [61] Demkowicz-Dobrzański R, Kołodyński J and Guţă M 2012 The elusive Heisenberg limit in quantum-enhanced metrology *Nat. Commun.* **3** 1063
- [62] Escher B M, de Matos Filho R L and Davidovich L 2011 General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology *Nat. Phys.* **7** 406–11
- [63] Knysh S I, Chen E H and Durkin G A 2014 True limits to precision via unique quantum probe (arXiv:1402.0495)
- [64] Huelga S F, Macchiavello C, Pellizzari T, Ekert A K, Plenio M B and Cirac J I 1997 Improvement of frequency standards with quantum entanglement *Phys. Rev. Lett.* **79** 3865–8
- [65] Kielinski T, Schmidt P O and Hammerer K 2024 GHZ protocols enhance frequency metrology despite spontaneous decay *Sci. Adv.* **10** eadr1439
- [66] Matei D G *et al* 2017 1.5 μm lasers with sub-10 mHz linewidth *Phys. Rev. Lett.* **118** 263202
- [67] Leroux I D, Scharnhorst N, Hannig S, Kramer J, Pelzer L, Stepanova M and Schmidt P O 2017 On-line estimation of local oscillator noise and optimisation of servo parameters in atomic clocks *Metrologia* **54** 307
- [68] Schulte M, Lisdat C, Schmidt P O, Sterr U and Hammerer K 2020 Prospects and challenges for squeezing-enhanced optical atomic clocks *Nat. Commun.* **11** 5955
- [69] Fraas M 2016 An analysis of the stationary operation of atomic clocks *Commun. Math. Phys.* **348** 363–93
- [70] Demkowicz-Dobrzański R, Jarzyna M and Kołodyński J 2015 Quantum limits in optical interferometry *Progress in Optics* (Elsevier) pp 345–435
- [71] Han C, Ma Z, Qiu Y, Fang R, Wu J, Zhan C, Li M, Huang J, Lu B and Lee C 2024 Atomic clock locking with Bayesian quantum parameter estimation: scheme and experiment *Phys. Rev. Appl.* **22** 044058

- [72] Macieszczak K, Fraas M and Demkowicz-Dobrzański R 2014 Bayesian quantum frequency estimation in presence of collective dephasing *New J. Phys.* **16** 113002
- [73] Kaubruegger R, Vasilyev D V, Schulte M, Hammerer K and Zoller P 2021 Quantum variational optimization of Ramsey interferometry and atomic clocks *Phys. Rev. X* **11** 041045
- [74] Thurtell T G and Miyake A 2024 Optimizing one-axis twists for variational Bayesian quantum metrology *Phys. Rev. Res.* **6** 023179
- [75] Scharnagl M S, Kielinski T and Hammerer K 2023 Optimal Ramsey interferometry with echo protocols based on one-axis twisting *Phys. Rev. A* **108** 062611
- [76] Marciniak C D, Feldker T, Pogorelov I, Kaubruegger R, Vasilyev D V, van Bijnen R, Schindler P, Zoller P, Blatt R and Monz T 2022 Optimal metrology with programmable quantum sensors *Nature* **603** 604–9
- [77] Li J, da Silva G R M, Kain S and Shahrar S M 2023 Generalized echo squeezing protocol with near-Heisenberg-limit sensitivity and strong robustness against detection noise and variation in squeezing parameter *Phys. Rev. A* **107** 032610
- [78] Schulte M, Martínez-Lahuerta V J, Scharnagl M S and Hammerer K 2020 Ramsey interferometry with generalized one-axis twisting echoes *Quantum* **4** 268
- [79] Leibfried D, Barrett M D, Schaetz T, Britton J, Chiaverini J, Itano W M, Jost J D, Langer C and Wineland D J 2004 Toward Heisenberg-limited spectroscopy with multiparticle entangled states *Science* **304** 1476
- [80] Davis E, Bentsen G and Schleier-Smith M 2016 Approaching the Heisenberg limit without single-particle detection *Phys. Rev. Lett.* **116** 053601
- [81] Fröwis F, Sekatski P and Dür W 2016 Detecting large quantum Fisher information with finite measurement precision *Phys. Rev. Lett.* **116** 090801
- [82] Macrì T, Smerzi A and Pezzè L 2016 Loschmidt echo for quantum metrology *Phys. Rev. A* **94** 010102
- [83] Nolan S P, Szigeti S S and Haine S A 2017 Optimal and robust quantum metrology using interaction-based readouts *Phys. Rev. Lett.* **119** 193601
- [84] Colombo S, Pedrozo-Peñañiel E, Adiyatullin A F, Li Z, Mendez E, Shu C and Vuletić V 2022 Time-reversal-based quantum metrology with many-body entangled states *Nat. Phys.* **18** 925–30
- [85] Burd S C, Srinivas R, Bollinger J J, Wilson A C, Wineland D J, Leibfried D, Slichter D H and Allcock D T C 2019 Quantum amplification of mechanical oscillator motion *Science* **364** 1163
- [86] Anders F, Pezzè L, Smerzi A and Klempt C 2018 Phase magnification by two-axis counter-twisting for detection-noise robust interferometry *Phys. Rev. A* **97** 043813
- [87] Kitagawa M and Ueda M 1993 Squeezed spin states *Phys. Rev. A* **47** 5138
- [88] Benhelm J, Kirchmair G, Roos C F and Blatt R 2008 Towards fault-tolerant quantum computing with trapped ions *Nat. Phys.* **4** 463–6
- [89] Blatt R and Wineland D 2008 Entangled states of trapped atomic ions *Nature* **453** 1008–15
- [90] Hines J A, Rajagopal S V, Moreau G L, Wahrman M D, Lewis N A, Marković O and Schleier-Smith M 2023 Spin squeezing by Rydberg dressing in an array of atomic ensembles *Phys. Rev. Lett.* **131** 063401
- [91] Sørensen A, Duan L-M, Cirac J I and Zoller P 2001 Many-particle entanglement with Bose–Einstein condensates *Nature* **409** 63–66
- [92] Estève J, Gross C, Weller A, Giovanazzi S and Oberthaler M K 2008 Squeezing and entanglement in a Bose–Einstein condensate *Nature* **455** 1216–9
- [93] Gross C, Zibold T, Nicklas E, Estève J and Oberthaler M K 2010 Nonlinear atom interferometer surpasses classical precision limit *Nature* **464** 1165–9
- [94] Riedel M F, Böhi P, Li Y, Hänsch T W, Sinatra A and Treutlein P 2010 Atom-chip-based generation of entanglement for quantum metrology *Nature* **464** 1170–3
- [95] Górecki W, Demkowicz-Dobrzański R, Wiseman H M and Berry D W 2020 π -corrected Heisenberg limit *Phys. Rev. Lett.* **124** 030501
- [96] Direkci S, Finkelstein R, Endres M and Gefen T 2024 Heisenberg-limited Bayesian phase estimation with low-depth digital quantum circuits (arXiv:2407.06006 [quant-ph])
- [97] Van Trees H L 1968 *Detection, Estimation and Modulation Theory: Detection, Estimation and Linear Modulation Theory: Part I* (Wiley)
- [98] Macieszczak K 2013 Quantum Fisher information: variational principle and simple iterative algorithm for its efficient computation (arXiv:1312.1356)
- [99] Jarzyna M and Demkowicz-Dobrzański R 2015 True precision limits in quantum metrology *New J. Phys.* **17** 013010
- [100] Gill R D and Levit B Y 1995 Applications of the van Trees inequality: a Bayesian Cramér–Rao bound *Bernoulli* **1** 59
- [101] Bužek V, Derka R and Massar S 1999 Optimal quantum clocks *Phys. Rev. Lett.* **82** 2207
- [102] Berry D W and Wiseman H M 2000 Optimal states and almost optimal adaptive measurements for quantum interferometry *Phys. Rev. Lett.* **85** 5098
- [103] Summy G and Pegg D 1990 Phase optimized quantum states of light *Opt. Commun.* **77** 75
- [104] Pegg D T and Barnett S M 1988 Unitary phase operator in quantum mechanics *Europhys. Lett.* **6** 483
- [105] Sanders B C and Milburn G J 1995 Optimal quantum measurements for phase estimation *Phys. Rev. Lett.* **75** 2944
- [106] Derka R, Bužek V and Ekert A K 1998 Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement *Phys. Rev. Lett.* **80** 1571
- [107] Luis A and Peřina J 1996 Optimum phase-shift estimation and the quantum description of the phase difference *Phys. Rev. A* **54** 4564
- [108] Allan D 1966 Statistics of atomic frequency standards *Proc. IEEE* **54** 221
- [109] Riley W and Howe D 2008 *Handbook of Frequency Stability Analysis* (National Institute of Standards and Technology) (available at: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=50505)
- [110] Rutman J and Walls F 1991 Characterization of frequency stability in precision frequency sources *Proc. IEEE* **79** 952
- [111] Barnes J A *et al* 1971 Characterization of frequency stability *IEEE Trans. Instrum. Meas.* **IM-20** 105
- [112] Arecchi F T, Courtens E, Gilmore R and Thomas H 1972 Atomic coherent states in quantum optics *Phys. Rev. A* **6** 2211
- [113] Radcliffe J M 1971 Some properties of coherent spin states *J. Phys. A: Gen. Phys.* **4** 313
- [114] Hioe F T 1974 Coherent states and Lie algebras *J. Math. Phys.* **15** 1174–7
- [115] Greenberger D M, Horne M A and Zeilinger A 1989 Going beyond Bell’s theorem *Bell’s Theorem, Quantum Theory and Conceptions of the Universe* (Springer) pp 69–72
- [116] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 Optimal frequency measurements with maximally correlated states *Phys. Rev. A* **54** R4649
- [117] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Hänsel W, Hennrich M and

- Blatt R 2011 14-qubit entanglement: creation and coherence *Phys. Rev. Lett.* **106** 130506
- [118] Li W, Wu S, Smerzi A and Pezzè L 2022 Improved absolute clock stability by the joint interrogation of two atomic ensembles *Phys. Rev. A* **105** 053116
- [119] Dick G J 1987 Local oscillator induced instabilities in trapped ion frequency standards *Proc. 19th Annual Precise Time and Time Interval Systems and Applications Meeting (Redondo Beach, California, 1–3 December)* p 133
- [120] Dick G, Prestage J, Greenhall C and Maleki L 1990 Local oscillator induced degradation of medium-term stability in passive atomic frequency standards *Proc. 22th Annual Precise Time and Time Interval Systems and Applications Meeting (Virginia, 4–6 December)*
- [121] Santarelli G, Audoin C, Makdissi A, Laurent P, Dick G and Clairon A 1998 Frequency stability degradation of an oscillator slaved to a periodically interrogated atomic resonator *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45** 887–94
- [122] Gardiner C 2010 *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer)
- [123] Robinson J M, Oelker E, Milner W R, Zhang W, Legero T, Matei D G, Riehle F, Sterr U and Ye J 2019 Crystalline optical cavity at 4K with thermal-noise-limited instability and ultralow drift *Optica* **6** 240
- [124] Kedar D, Yu J, Oelker E, Staron A, Milner W R, Robinson J M, Legero T, Riehle F, Sterr U and Ye J 2023 Frequency stability of cryogenic silicon cavities with semiconductor crystalline coatings *Optica* **10** 464
- [125] Mullan M and Knill E 2014 Optimizing passive quantum clocks *Phys. Rev. A* **90** 042310
- [126] Zhou J, Huang J, Wei J, Han C and Lee C 2024 High-dynamic-range atomic clocks with dual Heisenberg-limited precision scaling (arXiv:2411.14944)
- [127] Dörscher S, Al-Masoudi A, Bober M, Schwarz R, Hobson R, Sterr U and Lisdat C 2020 Dynamical decoupling of laser phase noise in compound atomic clocks *Commun. Phys.* **3** 185
- [128] Nardelli N V *et al* 2021 Differential spectroscopy of atomic clocks for improved measurement instability *Conf. on Lasers and Electro-Optics* (Optica Publishing Group) p STh4A.4
- [129] Hume D B and Leibbrandt D R 2016 Probing beyond the laser coherence time in optical clock comparisons *Phys. Rev. A* **93** 032138
- [130] Zheng X, Dolde J, Lochab V, Merriman B N, Li H and Kolkowitz S 2022 Differential clock comparisons with a multiplexed optical lattice clock *Nature* **602** 425–30
- [131] Clements E R, Kim M E, Cui K, Hankin A M, Brewer S M, Valencia J, Chen J-S, Chou C-W, Leibbrandt D R and Hume D B 2020 Lifetime-limited interrogation of two independent $^{27}\text{Al}^+$ clocks using correlation spectroscopy *Phys. Rev. Lett.* **125** 243602
- [132] Pezzè L and Smerzi A 2020 Heisenberg-limited noisy atomic clock using a hybrid coherent and squeezed state protocol *Phys. Rev. Lett.* **125** 210503
- [133] Kim M E *et al* 2022 Improved interspecies optical clock comparisons through differential spectroscopy *Nat. Phys.* **19** 25–29
- [134] Kessler E M, Kómár P, Bishof M, Jiang L, Sørensen A S, Ye J and Lukin M D 2014 Heisenberg-limited atom clocks based on entangled qubits *Phys. Rev. Lett.* **112** 190403
- [135] Borregaard J and Sørensen A S 2013 Efficient atomic clocks operated with several atomic ensembles *Phys. Rev. Lett.* **111** 090802
- [136] Bowden W, Vianello A, Hill I R, Schioppo M and Hobson R 2020 Improving the Q factor of an optical atomic clock using quantum nondemolition measurement *Phys. Rev. X* **10** 041052
- [137] Biedermann G W, Takase K, Wu X, Deslauriers L, Roy S and Kasevich M A 2013 Zero-dead-time operation of interleaved atomic clocks *Phys. Rev. Lett.* **111** 170802
- [138] Takamoto M, Takano T and Katori H 2011 Frequency comparison of optical lattice clocks beyond the Dick limit *Nat. Photon.* **5** 288–92
- [139] Rosenband T and Leibbrandt D R 2013 Exponential scaling of clock stability with atom number (arXiv:1303.6357)
- [140] Zheng X, Dolde J and Kolkowitz S 2024 Reducing the instability of an optical lattice clock using multiple atomic ensembles *Phys. Rev. X* **14** 011006
- [141] Ulam-Orgikh D and Kitagawa M 2001 Spin squeezing and decoherence limit in Ramsey spectroscopy *Phys. Rev. A* **64** 052106
- [142] Fröwis F, Skotiniotis M, Kraus B and Dür W 2014 Optimal quantum states for frequency estimation *New J. Phys.* **16** 083010
- [143] Braverman B, Kawasaki A and Vuletić V 2018 Impact of non-unitary spin squeezing on atomic clock performance *New J. Phys.* **20** 103019
- [144] Helstrom C W 1969 Quantum detection and estimation theory *J. Stat. Phys.* **1** 231–52
- [145] Personick S 1971 Application of quantum estimation theory to analog communication over quantum channels *IEEE Trans. Inf. Theory* **17** 240
- [146] Holevo A 2011 *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale) (<https://doi.org/10.1007/978-88-7642-378-9>)
- [147] Paris M G A 2009 Quantum estimation for quantum technology *Int. J. Quantum Inf.* **07** 125–37
- [148] Allan D W 1987 Should the classical variance be used as a basic measure in standards metrology? *IEEE Trans. Instrum. Meas.* **IM-36** 646
- [149] Virtanen P *et al* 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python *Nat. Methods* **17** 261–72