Quantum Variational Optimization of Ramsey Interferometry and Atomic Clocks

Raphael Kaubruegger[*,1,2] Denis V. Vasilyev[*,1,2] Marius Schulte,[3] Klemens Hammerer,[3] and Peter Zoller[1,2]

[1]*Center for Quantum Physics, University of Innsbruck, 6020 Innsbruck, Austria*
[2]*Institute for Quantum Optics and Quantum Information of the Austrian Academy of Sciences, 6020 Innsbruck, Austria[*]*
[3]*Institute for Theoretical Physics, Institute for Gravitational Physics (Albert
Einstein Institute), Leibniz University Hannover, 30167 Hannover, Germany*

We discuss quantum variational optimization of Ramsey interferometry with ensembles of N entangled atoms, and its application to atomic clocks based on a Bayesian approach to phase estimation. We identify best input states and generalized measurements within a variational approximation for the corresponding entangling and decoding quantum circuits. These circuits are built from basic quantum operations available for the particular sensor platform, such as one-axis twisting, or finite range interactions. Optimization is defined relative to a cost function, which in the present study is the Bayesian mean squared error of the estimated phase for a given prior distribution, i.e. we optimize for a finite dynamic range of the interferometer. In analogous variational optimizations of optical atomic clocks, we use the Allan deviation for a given Ramsey interrogation time as the relevant cost function for the long-term instability. Remarkably, even low-depth quantum circuits yield excellent results that closely approach the fundamental quantum limits for optimal Ramsey interferometry and atomic clocks. The quantum metrological schemes identified here are readily applicable to atomic clocks based on optical lattices, tweezer arrays, or trapped ions. While in the present work variationally optimized circuits are found with classical simulations, optimization can also be performed 'on' the (physical) quantum sensor, also in regimes not accessible to classical computations and in presence of imperfections.

I. INTRODUCTION

Recent progress in quantum technology of sensors has provided us with the most precise measurement devices available in physical sciences. Examples include the development of optical clocks [1], atom [2] and light [3] interferometers, and magnetic field sensing [4]. These achievements have opened the door to novel applications from the practical to the scientific. Atomic clocks and atomic interferometers allow height measurements in relativistic geodesy [5–8] or fundamental tests of our understanding of the laws of nature [9–11], such as time variation of the fine structure constant. In the continuing effort to push the boundaries of quantum sensing, entanglement as a key element of quantum physics gives the opportunity to reduce quantum fluctuations inherent in quantum measurements below the standard quantum limit (SQL), i.e. what is possible with uncorrelated constituents [12]. Squeezed light improves gravitational wave detection [13], allows lifescience microscopy below the photodamage limit [14], further, squeezing has been demonstrated in atom interferometers [15–30]. However, beyond the SQL, quantum physics imposes ultimate limits on quantum sensing, and one of the key challenges is to identify, and in particular devise experimentally realistic strategies defining *optimal quantum sensors* [31].

In our discussion below we will focus on optimal Ramsey interferometry, where 'optimality' is defined as achieving the best average signal to noise ratio (SNR) for phase estimation in a single-shot measurement. The distinguishing feature of the present work is that we consider optimal Ramsey interferometry with *finite dynamic range*, i.e. we wish to achieve optimal sensitivity for phases ϕ in a given finite interval of width $\delta\phi$ [31–37] as is relevant for numerous applications including atomic clocks [38–48]. To implement this optimal Ramsey interferometry we devise an approach based on variational quantum circuits [49–56]. Here, entangled input states and the entangled measurement protocols [57–63] defining the generalized Ramsey interferometer are represented as variational quantum circuits built from 'natural' quantum resources available on the specific sensor platform (see Fig. 1), which are optimized in light of a given cost function defining the optimal interferometer. As we will show, already low-depth variational quantum circuits can provide excellent approximations to the optimal interferometer. Intermediate scale atomic quantum devices [64, 65], acting as programmable quantum sensors [66], present the opportunity to implement these low-depth quantum circuits, defining an experimental route towards *optimal* Ramsey interferometry.

As noted above, optimality of a quantum sensing protocol is defined via a cost function $\mathcal{C}$ which is identified in context of a specific metrological task. In our study of variational N-atom Ramsey interferometry, we wish to optimize for phase estimation accuracy defined as the mean squared error $\epsilon(\phi)$ relative to the actual phase ϕ, averaged with respect to a prior distribution $\mathcal{P}(\phi)$ with width $\delta\phi$, which represents the finite *finite dynamic range* of the interferometer. Thus the cost function is $\mathcal{C} \equiv (\Delta\phi)^2 = \int d\phi\, \epsilon(\phi)\mathcal{P}(\phi)$. This corresponds to a Bayesian approach to optimal interferometry where the prior width of the phase distribution $\delta\phi$ is updated through measurement to $\Delta\phi$ characterizing the posterior distribution. As outlined in Fig. 1, our
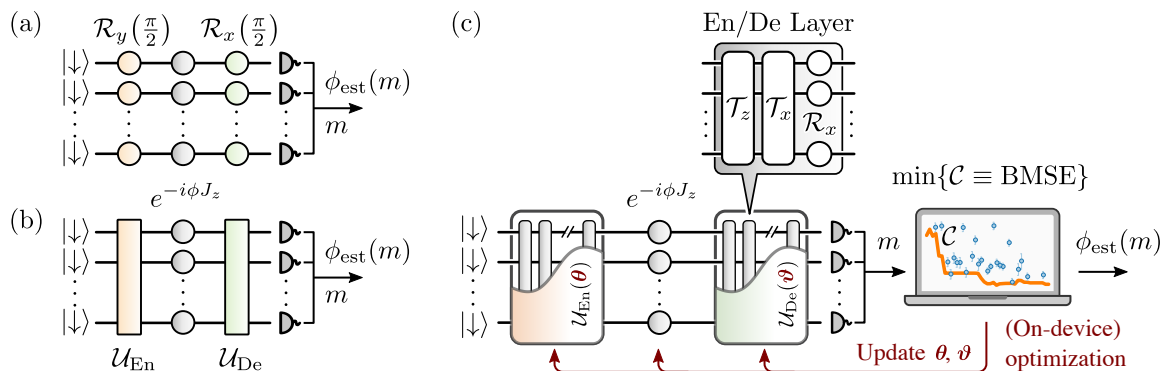
FIG. 1. (a) Quantum circuit representation of Ramsey interferometer with uncorrelated atoms. The phase ϕ is imprinted on the atomic spin-superposition prepared by global $\pi/2$-rotation around y-axis, $\mathcal{R}_y(\pi/2)$. Consequent rotation, $\mathcal{R}_x(\pi/2)$, and measurement of difference m of atoms in eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ in z-basis allows estimating the phase ϕ using an estimator function $\phi_{\text{est}}(m)$. (b) Quantum circuit of a generalized Ramsey interferometer with generic *entangling* and *decoding* operations $\mathcal{U}_{\text{En}}$ and $\mathcal{U}_{\text{De}}$, respectively. Our variational approach (c) consists of an ansatz, where optimal $\mathcal{U}_{\text{En}}$ and $\mathcal{U}_{\text{De}}$ are approximated by low-depth circuits. These are built from 'layers' of elementary operations, which are provided by the given platform. We specify the variationally optimized quantum sensor by circuits $\mathcal{U}_{\text{En}}(\boldsymbol{\theta})$ and $\mathcal{U}_{\text{De}}(\boldsymbol{\vartheta})$ [see Eqs. (6) and (7)], of depth n_{En} and n_{De}, respectively. Here $\boldsymbol{\theta} \equiv \{\theta_i\}$ and $\boldsymbol{\vartheta} \equiv \{\vartheta_i\}$ are vectors of variational parameters to be optimized for a given strategy represented by a cost function $\mathcal{C}$ defined here as Bayesian mean squared error (BMSE) [see Eqs. (2) and (10)]. We illustrate the approach with a variational circuit built from global spin rotations $\mathcal{R}_x$ and one-axis-twisting gates $\mathcal{T}_{x,z}$ available in neutral atom and ion quantum simulation platforms, as discussed in Sec. II B. The circuits optimization, shown as a feedback loop (in red), can be performed on a classical computer, or, if the complexity of underlying quantum many-body problem exceeds capabilities of classical computers, on the sensor itself, thus, leading to a (relevant) quantum advantage, see Sec. II G.

variational approach to optimal Ramsey interferometry seeks to minimize $\mathcal{C}$ over variational quantum circuits, and thus identifying optimal input states and measurements for a given $\delta\phi$. Note that in the present work we optimize a metrological cost function for the complete quantum sensing protocol with variational quantum circuits. We distinguish this from variational state preparation schemes, e.g. variational squeezed state preparation of Ref. [66], where a squeezing parameter was optimized as cost function.

We contrast our Bayesian approach of identifying a metrological cost function to a Fisher information approach, which optimizes accuracy locally at a *specific* value of the phase, corresponding to the limit $\delta\phi \to 0$ [3]. Discussions of fundamental limits in quantum sensing are often phrased in terms of quantum Fisher information and the quantum Cramèr-Rao bound leading to definition of the Heisenberg limit (HL) [67–69]. This identifies GHZ states [70], saturating the HL, as the optimal states for Ramsey interferometry. Furthermore, this leads to the conclusion that adding a decoding step (see Fig. 1) is not beneficial for quantum metrology since a separable measurement is optimal in this context [68]. This conclusion, however, is not applicable to phase estimation with *finite* prior width since the GHZ state interferometry in single-shot scenarios is optimal only for estimation of phase values in an interval $\delta\phi_{\text{GHZ}} \sim 1/N$, which shrinks as number of atoms N increases [37, 71], see Sec. II F below. In fact, for large priors $\delta\phi$ tailored quantum input states will differ greatly from squeezed spin states (SSS) [72, 73] or GHZ states [3, 31], and a nontrivial mea-

surement is required for an optimal metrological protocol. Our variational approach to optimal Ramsey interferometry (see Fig. 1) finds these optimal entangling and decoding circuits [74].

Our discussion of optimal single-shot Ramsey interferometry [75] has immediate relevance for atomic clocks [12, 76–79]. An optical atomic clock operates by locking the frequency of an oscillator, represented by a classical laser field with fluctuating frequency $\omega_L(t)$, to the transition frequency ω_A of an ensemble of N isolated atoms [1]. The locking of the laser to the atomic transition is achieved by repeatedly measuring the accumulated phase $\phi = \int_0^T dt[\omega_L(t) - \omega_A]$ in Ramsey interferometry with interrogation time T. Importantly, the width $\delta\phi$ of the distribution of this phase increases with the Ramsey time T. It is therefore critical to achieve a good phase estimate in conjunction with a wide dynamic range for making an accurate inference about the frequency deviation, and ultimately for stabilizing the clock laser to the atomic transition. Our variational approach for Bayesian phase estimation is made to satisfy these requirements, and provides optimal quantum states and measurements minimizing the instability in atomic clocks as measured by its Allan deviation. We predict significant improvements over previously known one-shot non-adaptive strategies. Our predictions are backed up by comprehensive numerical simulations of the clock laser and its stabilization to the atomic reference in a closed feedback loop [78, 79].

In the following, we first develop the general theory of variationally optimized Ramsey interferometry based on Bayesian phase estimation in Sec. II, and then apply this

theory to the specific problem of an optical atomic clock in Sec. III.

II. QUANTUM VARIATIONAL OPTIMIZATION OF RAMSEY INTERFEROMETRY

For concreteness, we consider estimation of the phase ϕ in an atomic interferometer consisting of an ensemble of N identical two-level atoms described as spin-1/2 particles [12]. The general idea developed in the following applies to any $SU(2)$ interferometer. The interferometer encodes the phase in the atomic state by evolving according to $|\psi_\phi\rangle = \exp(-i\phi J_z)|\psi_{\rm in}\rangle$. Here $|\psi_{\rm in}\rangle$ is an initial probe state [80], and $J_{x,y,z} = 1/2\sum_{k=1}^{N}\sigma_k^{x,y,z}$ is the collective spin with $\sigma^{x,y,z}$ the Pauli operators. The task is to determine the unknown phase ϕ by performing a measurement on the atoms.

A. Bayesian approach to phase interferometry

The most general measurement is described by a positive operator valued measure (POVM), that is a set $\{\Pi_x\}$ of positive Hermitian operators such that $\int dx\,\Pi_x = \mathbb{1}$. The parameter ϕ is estimated on the basis of a measurement result x using an estimator function $\phi_{\rm est}(x)$. The phase estimation accuracy is characterized by a mean squared error (MSE) with respect to the actual phase ϕ

$$\epsilon(\phi) = \int dx\,[\phi - \phi_{\rm est}(x)]^2 p(x|\phi), \qquad (1)$$

where $p(x|\phi) = {\rm Tr}\{\Pi_x|\psi_\phi\rangle\langle\psi_\phi|\}$ is the conditional probability of the measurement outcome x [3]. In our discussion we consider the phase ϕ to be defined on the interval $-\infty < \phi < \infty$ [81].

In order to find an interferometer performing the most accurate measurement of the phase ϕ we cannot minimize the MSE (1) for all values of ϕ simultaneously. First, the atomic interferometer is only sensitive to phase values modulo 2π as $\exp(-i\phi J_z)$, and hence also $p(x|\phi)$ is periodic. Thus, it can not distinguish arbitrary phases. Second, an initial state and measurement working well for one phase value might be insensitive to another value. Thus we consider an estimation error minimized for a weighted range of phase values relevant for a given sensor and measurement task. In the following we will adopt a Bayesian approach where the estimation error is averaged over a prior phase distribution $\mathcal{P}(\phi)$. The cost function of interest is thus defined as the MSE averaged over the prior distribution, defining the Bayesian mean squared error (BMSE)

$$\mathcal{C} \equiv (\Delta\phi)^2 = \int_{-\infty}^{\infty} d\phi\,\epsilon(\phi)\mathcal{P}(\phi). \qquad (2)$$

The prior distribution $\mathcal{P}(\phi)$ reflects the statistical properties of the unknown phase ϕ hence it is, in general, sensor and task dependent.

Optimal interferometry is based on minimizing the cost function (2) over $|\psi_{\rm in}\rangle$, $\{\Pi_x\}$, and $\phi_{\rm est}(x)$ for the given prior distribution. For simplicity, we will focus on prior distribution as a normal distribution centered around zero

$$\mathcal{P}_{\delta\phi}(\phi) = \frac{1}{\sqrt{2\pi(\delta\phi)^2}}\exp\left[-\frac{\phi^2}{2(\delta\phi)^2}\right]. \qquad (3)$$

This problem was addressed in [31], where the optimal quantum interferometer has been identified. Below we optimize the cost function (2) within a variational quantum algorithmic approach.

B. Variational Ramsey interferometry

Our goal is to find an implementation of the optimal interferometer given a restricted set of quantum gates available on an experimental platform such as neutral atoms or trapped ions. We will show that low-depth variational quantum circuits of given depth [see Fig. 1(c)] are excellent approximations to optimal interferometry, and can yield significant improvements over SQL defined for uncorrelated atoms.

In the most general form the variational interferometer, illustrated in Fig. 1(b), can be defined by a generic entangling unitary operation $\mathcal{U}_{\rm En}$ preparing an entangled input state from the initial product state $|\psi_0\rangle = |\downarrow\rangle^{\otimes N}$

$$|\psi_{\rm in}\rangle = \mathcal{U}_{\rm En}|\psi_0\rangle, \qquad (4)$$

and a decoding operation $\mathcal{U}_{\rm De}$ transforming the projective measurement of a typical observable J_z, with eigenbasis $|m\rangle$, into a generic projection

$$\Pi_{x_m} \equiv |x_m\rangle\langle x_m| = \mathcal{U}_{\rm De}^\dagger|m\rangle\langle m|\mathcal{U}_{\rm De}. \qquad (5)$$

Here we consider the subspace spanned by the states $|m\rangle$, $m \in \{-N/2,\ldots,N/2\}$, which are completely symmetric under permutations of N atoms, and $|\psi_0\rangle = |-N/2\rangle$. The measurement amounts to counting the difference m of atoms in state $|\uparrow\rangle$ and $|\downarrow\rangle$. As shown in [34], this assumption can be made without loss of generality. The basis states $|m\rangle$ are given by the eigenstates of total spin of maximum length, $j = N/2$, thus satisfying $\boldsymbol{J}^2|m\rangle = j(j+1)|m\rangle$ and $J_z|m\rangle = m|m\rangle$. As shown in [31], the optimal POVM may be restricted to the class of standard projection von Neumann measurements $\Pi_x = |x\rangle\langle x|$, $\langle x|x'\rangle = \delta_{xx'}$. Thus the measurement of the collective spin component J_z transformed by a decoder $\mathcal{U}_{\rm De}$ represents the measurement problem in full generality.

We assume that the programmable quantum sensor provides us with a set of native resource Hamiltonians $\{H_R^{(i)}\}$. The unitaries generated by these Hamiltonians determine a corresponding native set of quantum gates as variational ansatz for $\mathcal{U}_{\rm En}$ and $\mathcal{U}_{\rm De}$. A generic example is provided by global rotations $\mathcal{R}_\mu(\theta) = \exp(-i\theta J_\mu)$ and the infinite range one-axis-twisting (OAT) interaction [73]

$\mathcal{T}_\mu(\theta) = \exp(-i\theta J_\mu^2)$ with $\mu = x, y, z$. Such interactions have been realized on quantum simulation platforms [15–29, 82, 83], and very recently also on an optical clock transition [30]. Within this set of gates we constrain the quantum circuits to be invariant under the spin x-parity transformation ensuring an anti-symmetric estimator at and around $\phi = 0$ (see App. B). The most general circuits satisfying the x-parity constraint for a fixed number $n_{\rm En}$ and $n_{\rm De}$ of layers of entangling and decoding gates are

$$\mathcal{U}_{\rm En}(\boldsymbol{\theta}) = \Big[\mathcal{R}_x(\theta_{n_{\rm En}}^{(3)})\mathcal{T}_x(\theta_{n_{\rm En}}^{(2)})\mathcal{T}_z(\theta_{n_{\rm En}}^{(1)})\cdots$$
$$\mathcal{R}_x(\theta_1^{(3)})\mathcal{T}_x(\theta_1^{(2)})\mathcal{T}_z(\theta_1^{(1)})\Big]\mathcal{R}_y\big(\tfrac{\pi}{2}\big), \quad (6)$$

and

$$\mathcal{U}_{\rm De}(\boldsymbol{\vartheta}) = \mathcal{R}_x\big(\tfrac{\pi}{2}\big)\Big[\mathcal{T}_z(\vartheta_1^{(1)})\mathcal{T}_x(\vartheta_1^{(2)})\mathcal{R}_x(\vartheta_1^{(3)})\cdots$$
$$\mathcal{T}_z(\vartheta_{n_{\rm De}}^{(1)})\mathcal{T}_x(\vartheta_{n_{\rm De}}^{(2)})\mathcal{R}_x(\vartheta_{n_{\rm De}}^{(3)})\Big]. \quad (7)$$

Here the subscripts on the parameters indicate the layer containing the same three gates and the superscript identifies the gate within the layer. The complexity of the circuit is thus classified by $(n_{\rm En}, n_{\rm De})$, and we have $3(n_{\rm En}+n_{\rm De})$ (global) variational parameters in a $(n_{\rm En}, n_{\rm De})$–circuit, independent of N. Note that here $\mathcal{U}_{\rm En}$ and $\mathcal{U}_{\rm De}$ commute with particle exchange. The Hilbert space dimension for dynamics in the symmetric subspace is linear in N, which allows us to study theoretically the scaling for large particle numbers N below — in contrast to the case of finite range interactions in Sec. II G.

We note that conventional Ramsey interferometry with uncorrelated atoms corresponds to the $(0,0)$–circuit with $\mathcal{U}_{\rm En} = \mathcal{R}_y(\pi/2)$ and $\mathcal{U}_{\rm De} = \mathcal{R}_x(\pi/2)$. Here atoms are prepared initially in a product state, or coherent spin state (CSS), and remain in a product state during the evolution in interferometer followed by measurement of J_y. On the other hand, the interferometer with SSS as input, and GHZ interferometry emerge as $(1,0)$– and $(2,1)$–circuits, respectively.

In the presented entangler-decoder framework the performance of the interferometer is described, similar to Eq. (1), by the MSE

$$\epsilon(\phi) = \sum_m \big[\phi - \phi_{\rm est}(m)\big]^2 p_{\boldsymbol{\theta},\boldsymbol{\vartheta}}(m|\phi), \quad (8)$$

where the conditional probability is

$$p_{\boldsymbol{\theta},\boldsymbol{\vartheta}}(m|\phi) = |\langle m|\mathcal{U}_{\rm De}(\boldsymbol{\vartheta})e^{-i\phi J_z}\mathcal{U}_{\rm En}(\boldsymbol{\theta})|\psi_0\rangle|^2. \quad (9)$$

Therefore, the optimal interferometer found within the restricted set of available operations is described by the minimum of the BMSE

$$(\Delta\phi)^2 = \min_{\boldsymbol{\theta},\boldsymbol{\vartheta},a}\int_{-\infty}^{\infty}d\phi\sum_m(\phi - am)^2$$
$$\times\, p_{\boldsymbol{\theta},\boldsymbol{\vartheta}}(m|\phi)\mathcal{P}_{\delta\phi}(\phi). \quad (10)$$
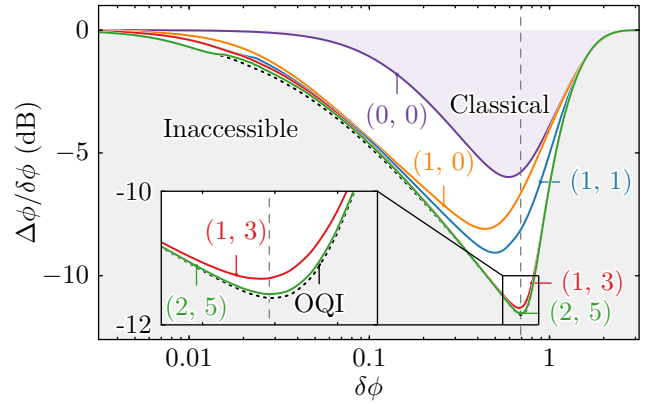

FIG. 2. Performance of the variationally enhanced interferometer with $N = 64$ particles. Performance is shown in terms of the posterior phase distribution width relative to the prior width, $\Delta\phi/\delta\phi$, for a given prior, that is, for a given dynamic range of the interferometer. Colored lines show the performance of variationally optimized circuits for the depth $(n_{\rm En}, n_{\rm De})$ of entangling and decoding layers as indicated. The number of variational parameters is given by $3(n_{\rm En}+n_{\rm De})$. The performance of the optimal quantum interferometer (OQI) [31] is indicated by the dotted line. The shaded areas indicate the classically accessible (purple) and the quantum mechanically forbidden (gray) regions (for $N = 64$). Related results applied to atomic clocks are shown in Fig. 10

To be specific, we assume for the prior a normal distribution $\mathcal{P}_{\delta\phi}(\phi)$ with standard deviation $\delta\phi$ [see Eq. (3)]. In addition, (10) assumes a linear estimator $\phi_{\rm est}(m) = am$ which is close to optimal, as shown below. We note that it is possible to use the optimal Bayesian estimator, which however is computationally demanding. We describe the corresponding iterative procedure in App. D for the case of a phase operator as observable.

C. Results of optimization

Results of interferometer optimizations [84] are shown in Fig. 2 for $N = 64$ atoms. The figure plots the ratio $\Delta\phi/\delta\phi$ of the root BMSE $\Delta\phi$ relative to the normal prior width $\delta\phi$. The more information we gain about the parameter ϕ in a single measurement, the smaller the value of this ratio.

The black dotted line shows the result of the unrestricted minimization of the cost function (2) with normal prior [31], which we refer to as optimal quantum interferometer (OQI). It defines the region (shaded area) inaccessible to any N-particle quantum interferometer. The purple line represents performance of the conventional Ramsey interferometer with CSS as input and a linear estimator, given by the $(0,0)$–circuit. Thus, the shaded area above the purple line roughly defines the classically achievable performance.
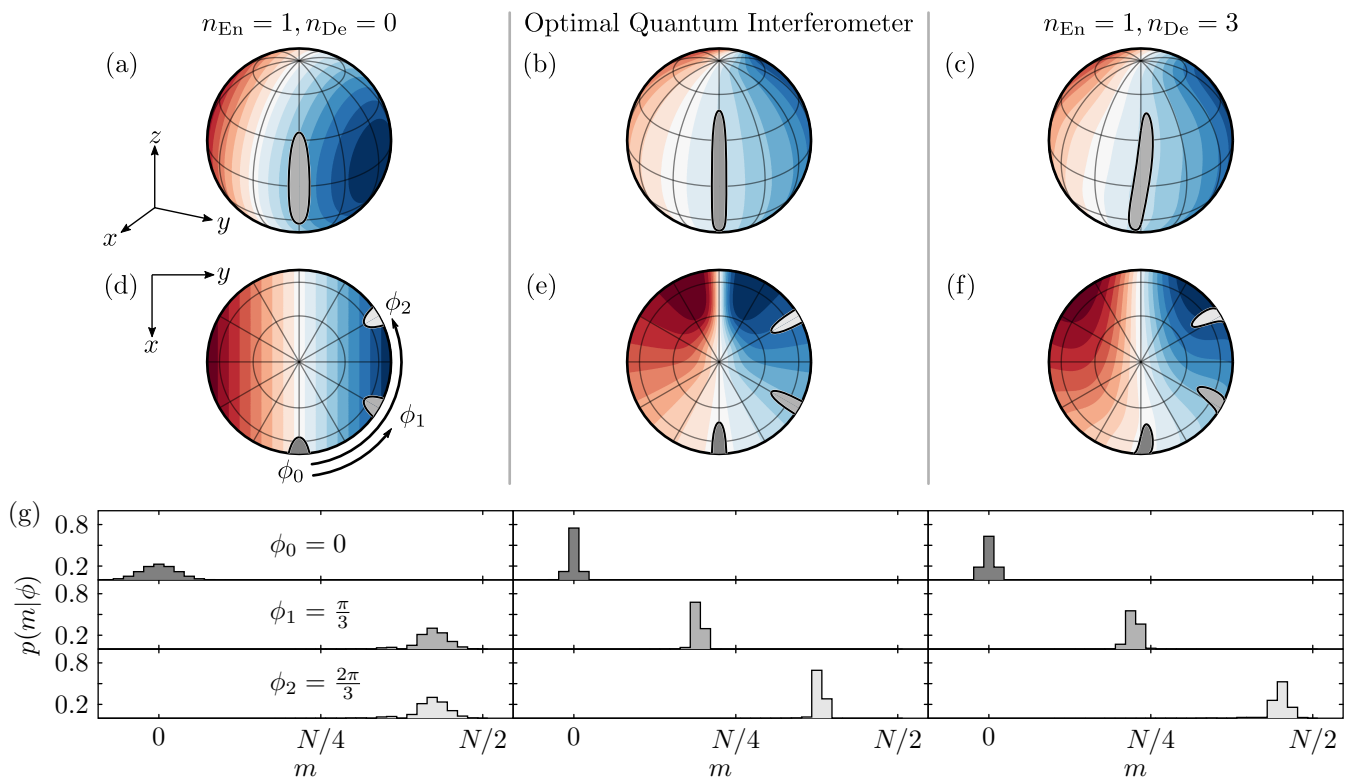
The performance of the entanglement enhanced inter-

FIG. 3. Visualization of quantum states $|\psi_\phi\rangle = \exp(-i\phi J_z)|\psi_{\rm in}\rangle$, and quantum measurement operators as Wigner distributions on the generalized Bloch sphere for $N = 64$ and $\delta\phi \approx 0.7$. The first (a,d), second (b,e), and third column (c,f) correspond to $(n_{\rm En}, n_{\rm De}) = (1, 0)$ (squeezed input state, and J_y measurement operator), the optimal quantum interferometer, and to a $(1, 3)$ quantum circuit, respectively. Measurement operators are visualized as colored contours on the Bloch sphere corresponding to different measurement outcomes. The corresponding optimized (optimal) states $|\psi_\phi\rangle$ are shown at various angles ϕ as gray shaded areas. (a,b,c) three dimensional view of the generalized Bloch sphere with a state rotated to $\phi = \pi/3$. (d, e, f) Top view of the Bloch sphere with the state rotated to angles $\phi = 0, \pi/3, 2\pi/3$. (g) Measurement probability $p(m|\phi)$ [see Eq. (9)] corresponding to the overlap between the contours of the measurement distribution and the respective state distribution, displayed in the same column. The three rows correspond to the above three angles ϕ. Note that for the J_y measurement the distributions at angles $\pi/3$ and $2\pi/3$ are indistinguishable in measurement statistics. In contrast, for the OQI and the $(1, 3)$ quantum circuit these angles are well resolved.

ferometer is shown with colored lines. The orange curve represents a $(1, 0)$–circuit corresponding to a squeezed spin state (SSS) interferometer [72, 73], employing the OAT interaction to generate an entangled initial state with suppressed fluctuations along the axis of the effective J_y measurement. The minimum of the orange line is located at smaller $\delta\phi$ values when compared to the minimum of the purple line corresponding to the SQL. This manifests the fact that SSS input state increases the sensitivity of the phase measurement at expense of dynamic range [76, 85, 86]. By adding a single layer of a decoding circuit we obtain the blue curve corresponding to the $(1, 1)$–interferometer with a slightly enhanced sensitivity and dynamic range. The red and green lines correspond to $(1, 3)$– and $(2, 5)$–circuits, respectively, and show striking improvement in sensitivity, providing an excellent approximation for the optimal interferometer (black dotted line). Remarkably, the minima of the red, green, and black curves are located at a wider dynamic range $\delta\phi$ than that of the CSS interferometer. Hence the

optimal entangled initial state and the effective nonlocal observable allows us to achieve both *a higher phase sensitivity* and *a wider dynamic range*.

To gain understanding of the physical meaning of the measurements and initial states emerging from the numerical optimization, we show their Wigner functions in Fig. 3. A formal definition of the Wigner distribution is provided in App. C. The three columns correspond, in consecutive order, to the $(1, 0)$–circuit (SSS interferometer), the optimal quantum interferometer of [31], and the $(1, 3)$–circuit. The chosen prior width $\delta\phi \approx 0.7$ is indicated in Fig. 2 by the vertical dashed line. The first row of panels 3(a-c) shows 3D views of the generalized Bloch sphere with Wigner functions of the measurement operators shown in shades of red and blue for J_y, the optimal observable, and $\mathcal{U}_{\rm De}^\dagger J_z \mathcal{U}_{\rm De}$ operators, respectively. A contour of constant color corresponds roughly to a certain measurement outcome which is obtained with the probability given by the overlap of the contour with the Wigner function of a quantum state. The states are shown
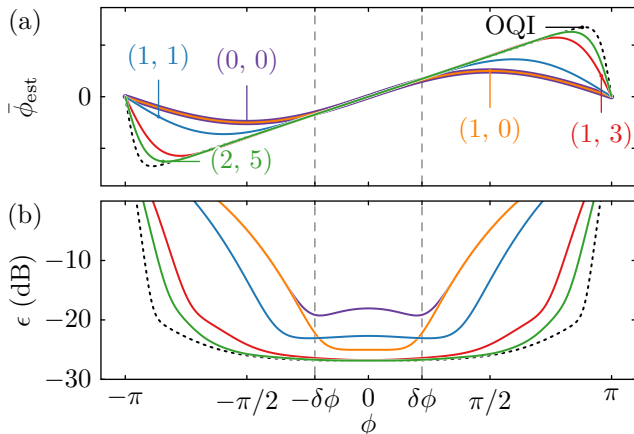
FIG. 4. (a) Phase ϕ dependence of the estimator expectation value [Eq. (11)] of the optimized $N = 64$ particle interferometer at different circuit depths $(n_{\mathrm{En}}, n_{\mathrm{De}})$ with $3(n_{\mathrm{En}} + n_{\mathrm{De}})$ variational parameters, in comparison to the optimal quantum interferometer (OQI). The optimization is performed for the prior distribution width $\delta\phi \approx 0.7$, indicated by the vertical lines. (b) Mean squared error [Eq. (8)] corresponding to the estimator expectation values curves above.

in 3(a-f) with the gray outlined areas.

Panel 3(a) shows clearly the non-optimality of the SSS interferometer with a measurement of the spin projection J_y. Optimization of the SSS results in a moderate level of squeezing (gray ellipse squeezed along the y-axis). More squeezing would produce stronger anti-squeezing along z-axis leading to overlap with more contours of the J_y Wigner function, thus increasing the variance of the measurement results for nonzero ϕ [76, 86]. Another limitation of the SSS interferometer, illustrated in panels (d) and (g), is the reduced dynamic range in the interval $-\pi/2$ and $\pi/2$. Panels (d) and (g) show that states rotated by the phase angle $\phi = 2\pi/3 > \pi/2$ have the same measurement statistics as states rotated by $\phi = \pi/3$. Thus, phases outside the $[-\pi/2, \pi/2]$ interval can not be reliably estimated.

The optimal quantum interferometer is explained in the central column of Fig. 3. Here panel (b) shows that the initial state is squeezed significantly stronger than in the SSS interferometer. This is possible because the corresponding optimal measurement is very similar to the phase operator of Pegg and Barnett [87], which has eigenstates with well defined phases (see Sec. II D below for detailed comparison). One can see that the color contours of the optimal measurement Wigner function in panels (b) and (e) are aligned with the meridians and thus overlap favorably with the strongly squeezed initial state rotated by a wide range of phase angles ϕ. Strikingly, the OQI can effectively use the full 2π dynamic range, as illustrated in panels (e) and (g).

Finally, the $(1,3)$–interferometer, presented in the third column of Fig. 3, exhibits properties similar to the OQI. Interestingly, the initial state in this case is not a con-

ventional squeezed state, as shown in panel (c), but a slightly twisted one. This, however, does not impair the performance of the interferometer as the effective measurement is also twisted such that it matches the initial state rotated by a wide range of phase angles. This peculiarity is a consequence of the restricted gate set available for the variational optimization in a realistic system. It is remarkable that the low depth $(1,3)$–circuit already provides an excellent approximation for the OQI.

The extended dynamic range of the variationally optimized interferometer is explored in Fig. 4. Panels (a) and (b) show, respectively, the estimator expectation value

$$\bar{\phi}_{\mathrm{est}} \equiv \sum_m \phi_{\mathrm{est}}(m)p(m|\phi) \tag{11}$$

and the estimator mean squared error (8) as functions of the actual phase ϕ for an interferometer optimized for prior width of $\delta\phi \approx 0.7$ (indicated with vertical dashed lines).

The estimator expectation value of the $(0,0)$– and $(1,0)$–circuit (CSS and SSS interferometer) is given by a sine function [purple and orange line in panel (a)], thus, it can unambiguously map the estimated phase to the actual phase in the range between $-\pi/2$ and $\pi/2$. However, the useful dynamic range of the interferometer is even narrower as shown by the estimator error in panel (b). The estimator error of the SSS state is suppressed below the CSS benchmark line only for phases between, roughly, $-\pi/4$ and $\pi/4$. The $(1,1)$–interferometer [blue line in (a) and (b)] starts to exploit the entangled measurement and achieves a bit wider linear regime of $\bar{\phi}_{\mathrm{est}}$ in (a) and a wider region of suppressed estimator error in (b). Although the minimum error of $(1,1)$–circuit is larger than that of $(1,0)$–circuit, it still has superior overall sensitivity as phases in the tails of the prior distribution are better resolved.

Finally, more complex decoding operations employed by the $(1,3)$– and $(2,5)$–circuit (red and green lines) allow to approach the performance of the optimal interferometer (black dotted lines). The linear regime of $\bar{\phi}_{\mathrm{est}}$ extends almost to the full 2π range, and the estimator error is well suppressed for phases deeply within the tails of the prior.

D. Comparison between variational and phase operator based interferometers

From a theory perspective it is interesting to compare the performance of the variationally optimized interferometer and the interferometer based on covariant measurement [33, 34]. Here covariant measurements represent the class of measurements optimal for phase estimation with no *a priori* knowledge and phase-shift symmetry, i.e. assuming a prior distribution $\mathcal{P}(\phi) = (2\pi)^{-1}$ and a 2π-periodic cost function, as opposed to the MSE (1).

In the case of clocks and magnetometry, the free evolution encoding the phase ϕ is the collective spin rotation, $e^{-i\phi J_z}$. The corresponding covariant measurement
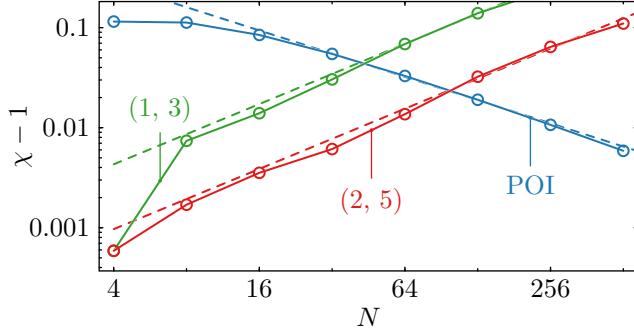
FIG. 5. Relative performance of the covariant, phase operator based interferometer (POI) (blue) and the variational $(1,3)$– and $(2,5)$–interferometers (green and red points, respectively) with respect to the OQI for a given system size N. The χ ratio is defined in Eq. (12), OQI corresponds to $\chi = 1$. The dashed lines represent empirical scalings, green and red one scale as $\sim N$, and blue is $\sim N^{-0.77}$.

optimal for estimation of the rotation angle ϕ can be represented by the von Neumann measurement [88] with phase operator $\hat{\Phi}$ [87], which we define in App. D.

In order to evaluate the performance of the phase operator based interferometer (POI), we minimize the cost function (2) for $\hat{\Phi}$ as the observable and the normal prior $\mathcal{P}_{\delta\phi}(\phi)$. To this end, we use the optimal Bayesian estimator known as the *Minimum Mean Squared Error* (MMSE) estimator [3] and find the corresponding optimal initial state $|\psi_{\hat{\Phi}}\rangle$ (see App. D for details). This results in the optimal posterior width $\Delta\phi_{\mathrm{POI}}$ as discussed in Sec. II C for the variationally optimized interferometer.

To compare different interferometers we consider their performance at the optimal prior width with respect to the OQI performance and define the ratio:

$$\chi = \frac{\min_{\delta\phi}(\Delta\phi/\delta\phi)}{\min_{\delta\phi}(\Delta\phi_{\mathrm{OQI}}/\delta\phi)}. \tag{12}$$

The χ value corresponds to the ratio of minima of an interferometer and the OQI curves in Fig. 2. The OQI corresponds to $\chi = 1$.

Figure 5 shows the $\chi - 1$ value for variationally optimized and $\hat{\Phi}$ based interferometers for various system sizes up to $N = 512$. The figure highlights sub-optimality of the POI (blue points) for the task of phase estimation with *non-periodic* cost function, as is relevant for frequency estimation in, e.g., optical clocks. For small systems, $N \lesssim 16$, the POI is up to $\sim 10\%$ less efficient than the OQI and the variational $(1,3)$– and $(2,5)$–interferometers (green and red points, respectively). $(1,3)$–circuit outperforms POI for systems of up to $N \sim 40$ atoms, whereas $(2,5)$–circuit is better for up to $N \sim 100$ atoms. In the limit of large number of atoms, $N \ggg 1$, the POI approaches the OQI performance. Empirical fitting indicates convergence rate $\chi_{\mathrm{POI}} - 1 \sim N^{-0.77}$, as N increases. On the other hand, the variationally optimized interferometers diverge from OQI linearly with N.

E. Variational Optimization in Presence of Imperfections and Noise

Variational optimization can be extended to include imperfections and decoherence. This optimization can also be carried out on the physical quantum sensor. This is particularly beneficial when the experimental characterization of imperfections and noise is incomplete.

There are various sources of imperfections and decoherence, which are relevant in our context. First, there are control errors in implementing variational quantum gates. These include offsets of control parameters and Hamiltonian design errors. The latter are deviations of the physically realized vs. the ideal Hamiltonian, e.g. in the implementation of one-axis twisting interaction. However, if these (unknown) control or design errors are static, i.e. do not fluctuate between experimental runs, a variational algorithm performed on the device will still optimize, and thus compensate in the best possible way for these errors in $\mathcal{U}_{\mathrm{En}}$ and $\mathcal{U}_{\mathrm{De}}$, i.e. find the best gate decomposition for given building blocks. In addition, there will be decoherence due to fluctuations of control parameters, or coupling to an environment as in spontaneous emission or dephasing.

To incorporate the latter we need to extend the formalism to density matrices instead of the previously discussed pure states. Below we illustrate this by an optimization of the Ramsey interferometer in the presence of single atom dephasing noise during the Ramsey interrogation time T, as one example of experimentally relevant decoherence. Local dephasing noise is described by the Lindbladian $\mathcal{L} \circ \rho = (1/4) \sum_{j=1}^{N} \left(\sigma_j^z \rho \sigma_j^z - \rho \right)$. Thus the density matrix after the Ramsey interrogation time,

$$\rho_{\boldsymbol{\theta}}^{\phi,\gamma T} = e^{-i\phi J_z} \left(e^{\gamma T \mathcal{L}\circ} \rho_{\boldsymbol{\theta}} \right) e^{i\phi J_z}, \tag{13}$$

can be expressed in terms of the dimensionless phase ϕ accumulated during the Ramsey interrogation time T and the effective exposure to the dephasing noise γT with dephasing rate γ. Here $\rho_{\boldsymbol{\theta}} = \mathcal{U}_{\mathrm{En}}(\boldsymbol{\theta}) |\psi_0\rangle \langle\psi_0| \mathcal{U}_{\mathrm{En}}^{\dagger}(\boldsymbol{\theta})$, where we used that the dephasing Lindbladian and the free evolution of the clock supercommute. The particle permutation symmetry of the Lindbladian enables us to simulate systems at a cubic cost in N [89, 90]. The conditional probability, required to determine the BMSE in Eq. (10) therefore reads

$$p_{\boldsymbol{\theta},\boldsymbol{\vartheta}}(m|\phi,\gamma T) = \langle m| \mathcal{U}_{\mathrm{De}}(\boldsymbol{\vartheta}) \rho_{\boldsymbol{\theta}}^{\phi,\gamma T} \mathcal{U}_{\mathrm{De}}^{\dagger}(\boldsymbol{\vartheta}) |m\rangle. \tag{14}$$

Figure 6 shows that the optimized $\Delta\phi/\delta\phi$ increases as the noise increases, as expected. For a small $\gamma T/\delta\phi = 0.01$ the variational $(1,3)$–interferometer is close to optimal without noise. Remarkably for all ratios $\gamma T/\delta\phi \lesssim 1$, the minimum of the $(1,3)$ interferometer remains well below the uncorrelated $(0,0)$– and the SSS $(1,0)$–interferometers. This ordering of the respective global minimum is independent of N, whereas for $\gamma T/\delta\phi = 10$ none of the entangling sequences improve significantly compared to SQL [91].
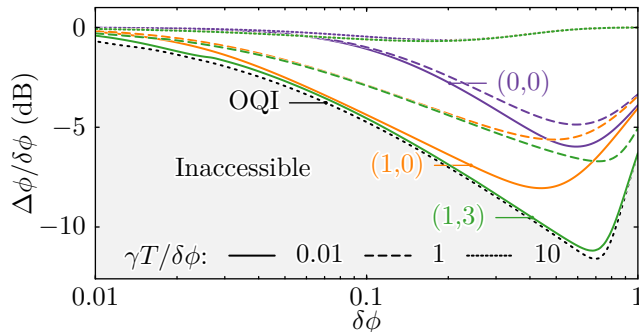
FIG. 6. Accuracy of an optimized $N = 64$ particle interferometer in the presence of single particle dephasing noise exposures $\gamma T/\delta\phi$ relative to the prior distribution width, indicated by different line styles. The results are displayed for different circuit complexities $(n_{\mathrm{En}}, n_{\mathrm{De}})$ with $3(n_{\mathrm{En}} + n_{\mathrm{De}})$ variational parameters. For comparison the accuracy of the noise free optimal quantum interferometer (OQI) is indicated by the black dotted line.
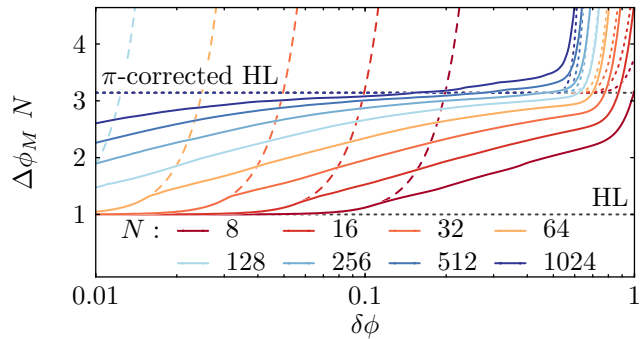
FIG. 7. Plot of $\Delta\phi_M N$, i.e. the standard deviation of an effective measurement rescaled by the ensemble size N, vs. prior width $\delta\phi$. Solid lines show results for the optimized interferometer with circuit depth $(n_{\mathrm{En}}, n_{\mathrm{De}}) = (2, 5)$ in comparison to the analytic expression describing a GHZ-state interferometer [Eq. (21)] shown with dashed lines and the π-corrected Heisenberg limit including phase slips [Eq. (22)] shown with dotted lines. The Heisenberg limit and the π-corrected Heisenberg limit are indicated with dotted horizontal lines.

F. Towards the Heisenberg limit

The variationally optimized interferometer with low-depth quantum circuits found within the Bayesian framework quickly approaches the accuracy of the optimal Ramsey interferometer. We will now discuss our results from the perspective of reaching the Heisenberg limit (HL).

The HL is a lower bound on the accuracy of an interferometer imposed by quantum mechanics. For an N-atom interferometer the HL and SQL are traditionally written as

$$\Delta\phi_{\mathrm{HL}}^2 \geq \frac{1}{N^2}, \quad \Delta\phi_{\mathrm{SQL}}^2 \geq \frac{1}{N}, \tag{15}$$

which must be understood in context of the quantum Fisher information [67, 69, 92] and quantum Cramér-Rao bound [33, 93] (implying $\delta\phi \to 0$). In contrast, in the present work we have adopted a Bayesian approach, which includes optimizing for a *finite dynamic range* $\delta\phi$. To evaluate the performance of our quantum variational results for a given circuit depth in comparison with HL, we will adopt below van Trees inequality [94, 95] as a bound for the BMSE.

In brief, for any given conditional probability distribution $p(m|\phi)$ the Cramér-Rao inequality

$$V(\phi) \geq \frac{1}{F_\phi} \tag{16}$$

provides a bound on the variance of an unbiased ($\bar{\phi}_{\mathrm{est}} = \phi$) estimator $V(\phi) \equiv \sum_m [\phi_{\mathrm{est}}(m) - \bar{\phi}_{\mathrm{est}}]^2 p(m|\phi)$ based on the Fisher information

$$F_\phi = \sum_m \left[\partial_\phi \log p(m|\phi)\right]^2 p(m|\phi). \tag{17}$$

For pure states, i.e. in the absence of decoherence, $F_\phi \leq N^2$ in correspondence to the HL above. We emphasize that the Cramér-Rao inequality seeks to identify optimal unbiased estimators, which can in general be achieved only *locally* in ϕ, i.e. in a small neighborhood of a given phase, and not for a finite dynamic range as is the goal in our Bayesian approach.

In the Bayesian framework, a bound on the BMSE is imposed by van Trees' inequality,

$$\left(\Delta\phi\right)^2 \geq \frac{1}{\overline{F_\phi} + \mathcal{I}}. \tag{18}$$

Here, the first term in the denominator is the Fisher information (17) averaged over the prior distribution, $\overline{F_\phi} = \int d\phi \mathcal{P}(\phi) F_\phi$. The second term is the Fisher information of the prior distribution, $\mathcal{I} = \int d\phi \mathcal{P}(\phi) \left[\partial_\phi \log \mathcal{P}(\phi)\right]^2$, representing the prior knowledge. To isolate the measurement contribution from the prior knowledge, we define an effective measurement variance $(\Delta\phi_M)^2$ via

$$\frac{1}{(\Delta\phi_M)^2} \equiv \frac{1}{(\Delta\phi)^2} - \mathcal{I}, \tag{19}$$

and obtain

$$(\Delta\phi_M)^2 \geq \frac{1}{\overline{F_\phi}} \geq \frac{1}{N^2}, \tag{20}$$

reminiscent of the Cramér-Rao inequality (16). In case of a normal prior distribution (3) we have $\mathcal{I} = (\delta\phi)^{-2}$, and the effective measurement variance (19) reads $(\Delta\phi_M)^{-2} \equiv (\Delta\phi)^{-2} - (\delta\phi)^{-2}$.

In Fig. 7 we plot $(\Delta\phi_M) \times N$, the measurement error scaled to the atom number, for the $(2, 5)$–variational interferometer (solid lines) as a function of the prior width $\delta\phi$ for a range of atom numbers N. In addition, we
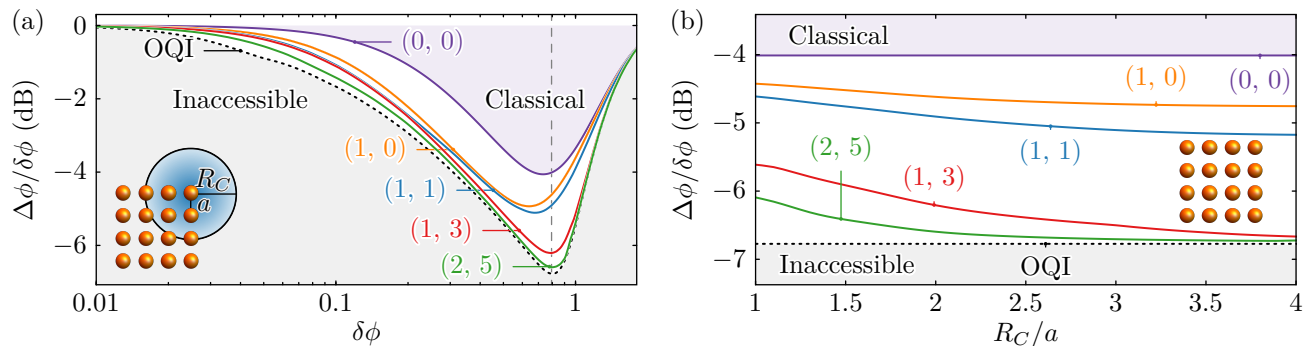
FIG. 8. Performance of the variationally optimized $N = 16$ Ramsey interferometer on a 4×4 lattice interacting via finite range Rydberg dressing interactions (23) with interaction radii R_C in units of the array spacing a. Colored lines show the reduction of the posterior phase distribution width $\Delta\phi$ relative to the prior distribution width $\delta\phi$ for variationally optimized circuits complexity $(n_{\text{En}}, n_{\text{De}})$ with $3(n_{\text{En}} + n_{\text{De}})$ variational parameters. The performance of the optimal quantum interferometer (OQI) [31] is indicated by dotted lines. The shaded areas indicate the classically accessible (purple) and the quantum mechanically inaccessible (gray) regions. (a) Prior width dependence of the optimized solution at $R_C/a = 2$. (b) Interaction range dependence of the optimal solution at a prior distribution width $\delta\phi \approx 0.8$ indicated by the vertical dashed line in (a).

indicate the HL and the π-corrected HL (see below) as dotted lines and show results for a GHZ interferometer with spin x-parity measurement [70] (dashed lines). In the case of the GHZ interferometer with a normal prior we have

$$(\Delta\phi_M^{\text{GHZ}})^2 = \frac{e^{(N\delta\phi)^2}}{N^2} - (\delta\phi)^2 \qquad (21)$$

showing that the GHZ interferometer attains the HL uncertainty $\Delta\phi_M \to 1/N$ for a given prior width $\delta\phi$ only for atom numbers $N \lesssim 1/\delta\phi$. This fact is illustrated in Fig. 7 by the dashed lines which diverge from the HL for smaller and smaller $\delta\phi$ as N grows. In contrast, the variational interferometer (solid lines) is of the order of the π-corrected HL [34, 35, 37, 96], $\Delta\phi_M \to \pi/N$, for a *wide range* of prior widths $\delta\phi$ as N increases.

Intuitively, the emergence of π-corrected HL can be understood as follows. The optimal N atom quantum interferometer can be described as a von Neumann measurement in the particle permutation symmetric subspace [31, 34]. Thus, there are $N+1$ possible measurement outcomes to distinguish at most $N+1$ phase values in the interval $[-\pi, \pi]$. The corresponding estimation error for evenly spread estimates reads $\Delta\phi \sim (1/2)\, 2\pi/(N+1) \to \pi/N$.

For large $\delta\phi$ the solid lines in Fig. 7 exhibit strong deviations from the asymptotic π-corrected HL behavior. The cusps are explained by phase slips outside the interval $[-\pi, \pi]$ which lead to a squared estimation error of $4\pi^2$. For a normal prior distribution, the performance of an interferometer limited by the π-corrected HL including the phase slips is given by

$$(\Delta\phi_M^{\pi\text{HL}})^2 = \frac{\pi^2}{N^2} + 4\pi^2\Big(1 - \text{erf}\frac{\pi}{\sqrt{2}\delta\phi}\Big). \qquad (22)$$

Results of this section are obtained in absence of decoherence.

G. Finite range interactions

Our previous discussion assumed infinite range interactions as entangling quantum resource, while e.g. neutral atoms stored in tweezer arrays feature finite range interactions. The variational optimization of the BMSE can be directly generalized to finite range interactions, which we illustrate by optimizing a sensor based on Rydberg dressing resources [97, 98] $\mathcal{D}_\mu(\theta) = \exp[-i\theta(H_\mu^D/V_0)]$, as is realized in alkaline earth tweezer clocks [46–48]. The effective interaction Hamiltonian we use for the optimization reads

$$H_\mu^D = \sum_{k,l=1}^{N} \frac{V_0 R_C^6/4}{|\boldsymbol{r}_k - \boldsymbol{r}_l|^6 + R_C^6} \sigma_k^\mu \sigma_l^\mu, \quad (\mu = x, y, z) \quad (23)$$

where $\boldsymbol{r}_k$ represents the position of particle k. The interaction strength at short distances V_0 and interaction radius R_C depend on the Rydberg level and the dressing laser used to let the particles interact [99].

Ref. [66] presented a study of variationally optimized spin-squeezed input states, and we refer to this work for the elementary gates we employ as building blocks for variationally optimizing entangling and decoding operations. In analogy to Eqs. (6) and (7), we write the entangler and decoder, effectively replacing the $\mathcal{T}_{x,z}$ by $\mathcal{D}_{x,z}$. In a similar way we can rewrite Eq. (9) to account for dynamics in full 2^N-dimensional Hilbert space.

Figure 8(a) shows the optimized $\Delta\phi/\delta\phi$ for a 4×4 square array for $R_C = a$ with a the lattice constant. We find variational solutions approximating the OQI, similarly to the OAT interactions in Fig. 2. In contrast to the infinite range OAT interaction we are not able to

exactly reproduce the optimal GHZ-state interferometer at $\delta\phi < 1/N$. Nonetheless, at any prior distribution width significant improvement beyond the uncorrelated interferometer is achieved and in particular around global minimum of the optimal interferometer (vertical dashed line), the decoder-enhanced circuits clearly surpass sensitivity of entangled input states only.

In Fig. 8(b) we further study the dependence on the scaled interaction radius R_C/a for a fixed prior distribution width $\delta\phi$ corresponding to the minima of variational and optimal interferometer curves in Fig. 8(a) (vertical dashed line). We see that even in the limit of an effective nearest neighbor interaction $R_C = a$ a clear improvement beyond the classical sensitivity limit is possible. As the interaction radius increases, the root BMSE of the variationally optimized interferometer decreases, ultimately reproducing the results of infinite-range interactions in the limit $R_C/a \to \infty$.

Theoretical treatment of the variational interferometry with finite range interactions involves solution of a quantum many-body problem. This, in general, is an exponentially hard problem representing the regime where variational optimization on the quantum sensor as a physical device provides a (relevant) quantum advantage, beyond the capabilities of classical computation.

III. APPLICATION TO ATOMIC CLOCKS

Atomic clocks realized with neutral atoms in optical trap arrays or trapped ions provide us with natural entanglement resources to implement variationally optimized Ramsey interferometry. Below we provide a study of a variationally optimized clock assuming as quantum resources global spin rotations and OAT, as realized, for example, with trapped ions as Mølmer-Sørensen gate, or in cavity setups with neutral atoms. This discussion is readily extended to other platforms and resources.

Optical atomic clocks operate by locking the fluctuating laser frequency $\omega_L(t)$ to an atomic transition frequency ω_A [1]. To this end, an atomic interferometer is used to repeatedly measure the phase $\phi_k = \int_{t_k}^{t_k+T} dt[\omega_L(t) - \omega_A]$ accumulated during interrogation time T at the k-th cycle of clock operation, i.e. $k = 1, 2, \ldots$. After each cycle, the measurement outcome m_k providing the phase estimate $\phi_{\text{est}}(m_k)$ is used to infer an estimated frequency deviation $\phi_{\text{est}}(m_k)/T$. In combination with previous measurement results this is used to correct the laser frequency fluctuations via a feedback loop yielding the corrected frequency of the clock $\omega(t)$. For further details on the actual clock operation we refer to App. G, where we also describe our numerical simulations of optical atomic clocks. We emphasize the importance of finite dynamic range in phase estimation in identifying the optimal clock operation, as provided by the Bayesian approach of Sec. II.

The relevant quantity characterizing the long-term clock instability is the Allan deviation $\sigma_y(\tau)$ for fluctuations of fractional frequency deviations $y \equiv [\omega(t) - \omega_A]/\omega_A$,

averaged over time $\tau \gg T$ [1]. To connect the Bayesian posterior phase variance of the optimized interferometer (10) of Sec. II, we follow the approach of [78] to obtain predictions for the clock instability in the limit of large averaging time τ. Our predictions are supported by numerical simulations of the closed servo-loop of the optical atomic clocks.

In the following we assume that interrogation cycles can be performed without dead times (Dick effect). This can be achieved using interleaved interrogation of two ensembles [100]. For interrogation of a single ensemble, Dick noise may pose limitations for interaction-enhanced protocols especially for larger ensembles, as was analyzed for squeezed states in Ref. [79]. In App. F and App. H we characterize in more detail the requirements regarding dead time for the class of variational protocols developed here.

A. Prediction of clock instability in the Bayesian framework

As shown in [78], the Allan deviation can be well approximated by means of the effective measurement uncertainty $\Delta\phi_M$ which isolates the measurement contribution from the prior knowledge, as in Eq. (19). Assuming no dark times between interrogation cycles, the Allan deviation reads

$$\sigma_y(\tau) = \frac{1}{\omega_A}\frac{\Delta\phi_M(T)}{T}\sqrt{\frac{T}{\tau}}. \tag{24}$$

Here τ/T is the number of cycles of clock operation and $\Delta\phi_M(T) \equiv [(\Delta\phi_T)^{-2} - (\delta\phi_T)^{-2}]^{-1/2}$ is the effective measurement uncertainty of one cycle. The posterior width $\Delta\phi_T$ is found according to (10) assuming a prior width $\delta\phi_T = (b_\alpha T)^{\alpha/2}$ corresponding to laser noise dominated spreading of the phase distribution within one interrogation cycle. The labels $\alpha = 1, 2, 3$ specify temporal correlations in the phase noise of the laser and correspond to atomic clocks with a white-, flicker-, or random-walk-frequency-noise-limited laser, respectively. The laser noise bandwidth b_α and the exponent α are related to the power spectral density $S_L(f) \propto f^{1-\alpha}$ of the free running laser (see App. A). Representative examples for $\sigma_y(\tau)$ when using variationally optimized protocols are shown in Fig. 9. The solid lines result from numerical simulations of the full feedback loop of an atomic clock in which an integrating servo corrects out frequency fluctuations over the course of multiple cycles, see App. G for details. For the simulations we assume the atoms as ideal frequency references without any systematic shift of ω_A.

In atomic clocks the simulated Allan deviations presented in Fig. 9 are larger at small averaging times $\tau/T \sim 1$, due to the delayed feedback, before reducing as $\sigma_y(\tau) \propto \tau^{-1/2}$ at long averaging times $\tau/T \gg 1$ when all correlated laser noise is corrected out. To determine long-term stability the Allan deviation is measured experimentally for a time τ long enough that clock instability
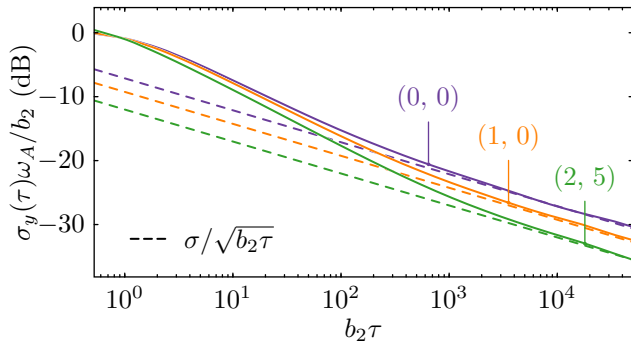
FIG. 9. Allan deviation representing a single run of a flicker frequency noise limited clock servo loop based on variationally optimized $N = 64$ particle interferometers at different circuit complexities (n_{En}, n_{De}) and a Ramsey interrogation time $b_2 T \approx 0.5$ (solid lines). For comparison the long time scaling predicted by Eq. (25) is shown by the dashed lines.

has reached this asymptotic scaling. Therefore, we introduce and consider below a dimensionless prefactor for the asymptotic scaling

$$\sigma = \frac{\Delta\phi_M(T)}{\sqrt{b_\alpha T}}, \qquad (25)$$

which gives the Allan deviation in units of $\omega_A^{-1}(b_\alpha/\tau)^{1/2}$, as shown by the dashed lines in Fig. 9. In the following, we use Eq. (25) to re-evaluate the performance of the optimized interferometers presented in Fig. 2 as the achievable long-term clock instability σ at an averaging time τ. In comparison to the framework of Sec. II the BMSE is replaced by the Allan deviation and the prior width by the interrogation time T. We note that the scaling of the Allan deviation with respect to T is more intricate than the one of the BMSE with the prior width: On the one hand, a large interrogation time means good accuracy in frequency estimation, but on the other hand, it also broadens the prior distribution and therefore degrades the phase estimation.

B. Results of the clock optimization

Figure 10(a,b) shows the achievable long-term clock instability σ as a function of the interrogation time T for clocks made of $N = 64$ atoms and the flicker-noise-limited laser. The purple line (in both panels) represents performance of the conventional clock exploiting Ramsey interferometer with CSS as input, collective spin projection measurement, and a linear estimator, given by the circuit $(0, 0)$. Thus, the shaded area above the purple line roughly defines the performance achievable by classical clocks. In the case of CSS based classical clocks the cost function (10) can be analytically minimized [78] yielding

the dimensionless Allan deviation

$$\sigma^{CSS} = \sqrt{\frac{1}{b_\alpha T}\left[\frac{e^\nu}{N} + \left(1 - \frac{1}{N}\right)\sinh\nu - \nu\right]}, \qquad (26)$$

where $\nu \equiv (\delta\phi_T)^2$. The expression (26) has two important limits. For small interrogation times and, consequently, small prior widths the performance of the clock is limited by the quantum projection noise of the uncorrelated atoms as $\sigma_{SQL} = (N b_\alpha T)^{-1/2}$. The SQL limited clock instability σ_{SQL} (dashed purple line) decreases as the interrogation time grows. For large interrogation times, $b_\alpha T \sim 1$, however, the laser noise becomes dominant and generates accumulated phase values exceeding the dynamic range of the atomic interferometer, thus, leading to the laser coherence time limit (CTL) [79] of the clock $\sigma_{CTL}^{CSS} = \{[\sinh(\delta\phi_T^2) - \delta\phi_T^2]/(b_\alpha T)\}^{1/2}$ (dotted purple line). Between these two limits there exists an optimal interrogation time delivering the minimum Allan deviation $\sigma_{opt} \equiv \min_T \sigma$ which defines the optimal clock performance.

The black dotted line in Fig. 10(a,b) shows the instability of the optimal quantum clock (OQC), σ^{OQC}, exploiting single-shot protocols with the optimal interferometer. The gray shaded region below the black dotted curve is inaccessible to any N-particle clock not using entanglement between different clock cycles for initial state preparations and/or measurements. The laser CTL for the optimal clock in the asymptotic limit of large N can be estimated from Eq. (2) by assuming zero phase estimation error within the $[-\pi, \pi]$ interval and $\epsilon(\phi) = 4\pi^2$ outside of the interval due to the phase slip

$$\sigma_{CTL}^{OQC} = \sqrt{\frac{4\pi^2}{b_\alpha T}\left(1 - \text{erf}\frac{\pi}{\sqrt{2}\,\delta\phi_T}\right)}. \qquad (27)$$

The green dotted line in panel (a) shows the laser CTL for the optimal clock, σ_{CTL}^{OQC}. The optimal clock instability at shorter interrogation times demonstrates two distinct scalings corresponding to the two Heisenberg limits discussed in Sec. II F. At very short times, $(b_\alpha T)^{\alpha/2} \lesssim N^{-1}$, the GHZ state based clock (red line) becomes optimal approaching the instability limit given by the conventional HL, $\sigma_{HL} = N^{-1}(b_\alpha T)^{-1/2}$ (red dashed line). Larger interrogation times correspond to wider prior phase distributions hence the π-corrected HL becomes the limiting factor, $\sigma_{\pi HL} = \pi N^{-1}(b_\alpha T)^{-1/2}$ (green dashed line). The optimal quantum clock instability in the limit of large number of atoms, $N \to \infty$, is fundamentally restricted by the interplay between the $\sigma_{\pi HL}$ and σ_{CTL}^{OQC} as we will discuss below.

The instabilities of clocks based on variationally optimized interferometers employing quantum circuits of various complexities are shown in Fig. 10(b) with solid color lines. In particular, the orange line corresponds to the SSS based clock, given by the circuit $(1, 0)$. As the circuits depth grows, the enhanced dynamic range of the variational interferometer shifts the laser CTL towards larger interrogation times which in combination
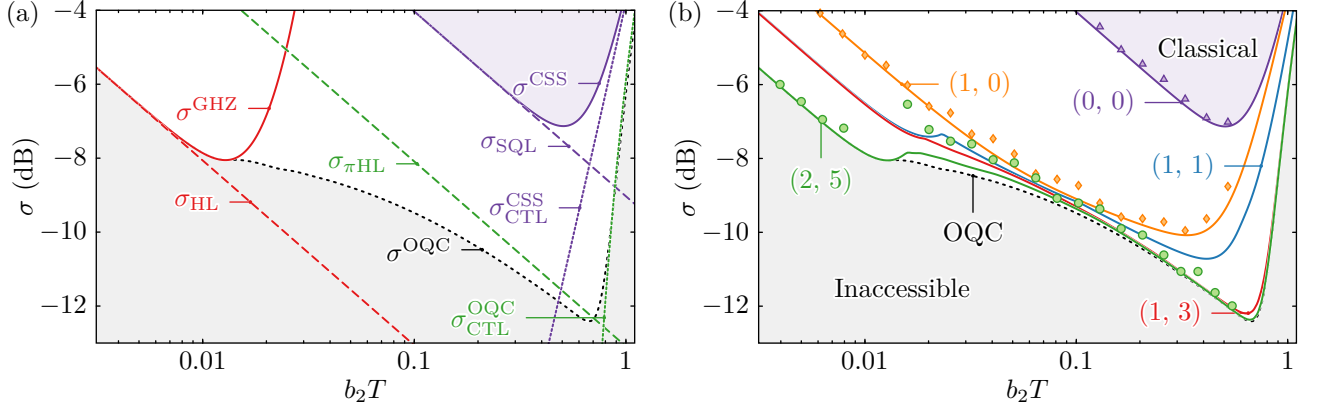
FIG. 10. Dimensionless Allan deviation σ [see Eq. (25)] of a $N = 64$ flicker-frequency-noise limited clock at constant averaging time τ as a function of the Ramsey interrogation time T rescaled to the bandwidth of the laser noise. The dotted black line indicates the instability of the optimal quantum clock (OQC). (a) Analytic expressions for the instability of coherent spin state (CSS) and GHZ-state clock (solid lines) in comparison to quantum projection noise limits, namely the standard quantum noise limit (SQL), π−corrected Heisenberg limit (πHL), HL (dashed) and the coherence time limits (CTL) of a CSS clock and the OQC (densely dotted). (b) Variational approximation of the OQC at increasing circuit complexities ($n_{\mathrm{En}}, n_{\mathrm{De}}$) with $3(n_{\mathrm{En}} + n_{\mathrm{De}})$ variational parameters. Markers show the numerically determined instability extracted from simulations of the full feedback loop of an atomic clock, described in detail in App. G, for a selection of optimized protocols. The numerical data is displayed only up to values of b_2T where no fringe-hop occurred within the 2×10^6 simulated clock cycles. Beyond this point an abrupt loss of stability was observed.

with suppressed shot noise reduces the clock instability. The figure shows that variational clocks of growing complexity quickly outperform the SSS clock and approach the optimal quantum clock instability. Beyond the model predictions this improvement is also observed in simulations of a full clock operation using variationally optimized protocols, as shown by the markers in Fig. 10(b). Deviations between theory and numerical results can arise due to a number of different effects. For one, the onset of fringe-hops for $b_2T \sim 1$ is not included explicitly in the models. Especially for small N a sudden loss of stability, resulting from fringe-hops, can occur before reaching the CTL due to stronger, non-Gaussian measurement noise [78, 79]. In contrast, for clocks with larger N and increasing complexity it is expected that the onset of fringe-hops and the minimum of CTL coincide. Another source of discrepancy is the assumption of a laser noise dominated prior width $\delta\phi_T = (b_\alpha T)^{\alpha/2}$. Propagation of the measurement uncertainty and delay within the feedback control can lead to a broadening of the true phase distribution. Especially protocols which are highly optimized to a particular prior width may thus not achieve their predicted stability in the simulations, e.g. around $b_2T \approx 0.02$ in Fig. 10(b).

Nevertheless, good agreement between the numerically determined instability and the theory prediction is found around the overall optimal protocols.

In Fig. 11 we study optimal instability of the variational clocks σ_{opt} (corresponds to minima in Fig. 10) as a function of the atomic ensemble size N. The CSS clock is represented by the purple line which scales asymptotically as $\sigma_{\mathrm{opt}}^{\mathrm{CSS}} \propto N^{-(3\alpha-1)/(6\alpha)}$. The scaling is a bit slower than the conventional SQL limit $\propto N^{-1/2}$ due to the laser
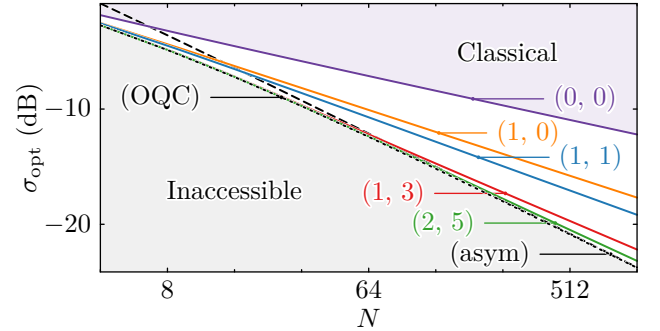

FIG. 11. Particle number N dependence of the dimensionless Allan deviation of a flicker-frequency-noise limited clock at a constant averaging time τ and the optimal Ramsey interrogation time for different circuit complexities ($n_{\mathrm{En}}, n_{\mathrm{De}}$) with $3(n_{\mathrm{En}} + n_{\mathrm{De}})$ variational parameters. For comparison we also show the performance of the optimal quantum clock (dotted line) and the asymptotic [see Eq. (28)] obtained from the π-corrected HL (dashed line).

CTL which reduces the optimal interrogation time as N grows. Any classical clock using one-shot protocols with collective spin measurements belongs to the shaded purple region above the CSS clock line.

The N-scaling of the optimal quantum clock is shown with the black dotted line for system sizes up to $N = 64$. For larger system sizes we show the asymptotic behavior (black dashed line) obtained by combining the noise contributions of the π-corrected HL and the laser CTL, $\sigma_{\mathrm{asym}} \equiv \min_T[\sigma_{\pi\mathrm{HL}}^2 + (\sigma_{\mathrm{CTL}}^{\mathrm{OQC}})^2]^{1/2}$. Similar to the classical clock scaling, the laser CTL prevents the optimal

quantum clock (OQC) from achieving the Heisenberg scaling $\propto N^{-1}$, instead, leading to a logarithmic correction in the large N limit as found in [101, 102]. The present approach allows obtaining tighter bounds on the asymptotic scaling for general α (see App. E). In particular, for the flicker-noise-limited laser, $\alpha = 2$, the OQC instability scales as

$$\sigma_{\text{opt}}^{\text{OQC}} \propto \frac{\sqrt{\pi}}{N} \left[\ln(z \ln z)^{-1/2} + \ln(z \ln z)^{1/2} \right]^{1/2}, \quad (28)$$

with $z \equiv 32N^4/\pi$ and the corresponding optimal interrogation time scaling as $T_{\text{opt}}^{\text{OQC}} \simeq \pi b_2^{-1} \ln(z \ln z)^{-1/2}$. The gray shaded area below the dashed and dotted black lines is inaccessible to quantum clocks without entangled clock cycles. Finally, the variationally optimized clocks of various circuit complexities are shown with solid color lines and demonstrate scalings approaching the optimal quantum clock as the circuit depth increases.

We have also studied performance of the variationally optimized clocks experiencing individual atomic dephasing during the interrogation period T. Similar to the results of Sec. II E, the optimized clocks perform well for decoherence rates small compared to the laser noise bandwidth, $\gamma/b_\alpha \ll 1$. For stronger noise, $\gamma/b_\alpha \gtrsim 1$, the optimized clock instability approaches the one of the classical clock, as expected. We also checked the performance of optimized clocks for other types of laser noise $\alpha = 1, 3$, and found no significant changes to the results presented above.

In summary, atomic clocks based on variational quantum interferometers with low-depth circuits can approach the performance of the optimal quantum clock in single-shot protocols. The variationally optimized clocks can be readily complemented with more sophisticated interrogation schemes [103, 104], eventually also approaching the ultimate quantum bound on the Allan deviation [105, 106].

IV. OUTLOOK AND CONCLUSIONS

In this work we have studied optimal Ramsey interferometry for phase estimation with entangled N-atom ensembles, and application of these optimal protocols to atomic clocks. We have considered a Bayesian approach to quantum interferometry, and have defined optimality via a cost function, which in the present study is the BMSE for a given prior distribution or, in the context of atomic clocks, the Allan deviation for a given Ramsey time. The key feature of the present work is that optimization is performed within the family of operational quantum resources provided by a particular programmable quantum sensor platform. Thus identifying the optimal quantum sensor is recast as a variational quantum optimization where the entangling circuits generating the optimal input state, and the decoding circuits implementing the optimal generalized measurement are variationally approximated with the given resource up to a certain circuit depth. We have presented two model studies: in our first model, we considered one-axis twisting as quantum resource; our second model uses finite range interactions as entangling operations. Our examples demonstrate that already low-depth circuits provide excellent approximations for optimal quantum interferometry. We emphasize that the familiar discussions of interferometry with spin-squeezing and GHZ states are included as special cases. Furthermore, advanced measurement strategies including adaptive measurement and quantum phase estimation are not advantageous for the present problem, as a von Neumann measurement has been proven optimal.

Given advances in building small atomic scale quantum computers, or programmable quantum simulators which can act also as quantum sensors, the variational approach to optimal quantum sensing provides a viable route to entanglement enhanced quantum measurements with existing experimental entangling, possibly non-universal resources, and optimizing in presence of noise. Indeed trapped ions with Mølmer-Sørensen entangling gates, and optical arrays interacting via Rydberg finite range interactions or cavity setups provide the necessary ingredients for implementing such variational protocols, and quantum sensors. While first generation experiments might demonstrate optimal Ramsey interferometry for a specified dynamic range of the phase, and optimization of quantum circuits 'on the quantum sensor' for various circuit depths (Sec. II), the present work also promises application of variational quantum sensing on existing quantum sensors, in particular atomic clocks (Sec. III). The guiding principle behind the present work of identifying for a sensing task the optimal sensing protocol given the quantum resources provided by a particular sensor and sensor platform, is, of course, general and generic, and applies beyond Ramsey interferometry, and beyond the BMSE as cost function.

As an outlook, we emphasize that the search for optimal sensing can also be run directly as a quantum–classical feedback loop on the physical quantum sensor. This offers the intriguing possibility of optimizing with given quantum resources and in presence of imperfections of the actual device, which might include control errors and noise. Further studies are needed to explore best optimization strategies of the cost function on the classical side of the optimization loop given the limited measurement budget on the programmable quantum sensor. This applies to both the initial global parameter search, supported by theoretical modeling, and small iterative readjustments of optimal operation points due to slow drifts of the quantum sensor.

Optimization *on* the (physical) quantum sensor can also be performed in the regime of large particle numbers N, which might be inaccessible to classical computations, i.e. in the regime of quantum advantage. Hybrid classical-quantum algorithms have been discussed previously as variational quantum eigensolvers for quantum chemistry and quantum simulation, where 'lowest energy' plays the role of the cost function which is evaluated on the quantum device. In contrast, in variational quantum sensing

we optimize quantum circuits in view of an 'optimal measurement' cost function, and it is the (potentially large scale) entanglement represented by the variational many-particle wavefunction in N-atom quantum memory, which provides the quantum resource and gain for the quantum measurement.

Note added. After submission of the present manuscript, Ref. [107] reported an experimental implementation of variationally optimized Ramsey interferometry in a systems of up to $N = 26$ trapped ions, in one-to-one correspondence to the present theoretical work. This includes demonstration of quantum enhancement in metrology beyond squeezing through low-depth, variational quantum circuits, and on-device quantum-classical feedback optimization to 'self-calibrate' the variational parameters. In both cases it is found that variational circuits outperform classical and direct spin squeezing strategies under realistic noise and imperfections.

Appendix A: Laser noise and prior distribution width

To present results in Sec. III in dimensionless units, we follow [78] and define an effective bandwidth $\tilde{b}$ via

$$\sigma_L(1/\tilde{b})\, \omega_A/\tilde{b} = 1, \qquad (A1)$$

where σ_L is the Allan deviation of the uncorrected reference laser. For a laser that is mainly limited by a single power spectral density component, i.e $S_L(f) = h_{1-\alpha}f^{1-\alpha}$ one can unambiguously express the bandwidth in terms of the prefactors $h_{1-\alpha}$ in the power spectral density and

the respective Allan deviation [108], so that

$$
\begin{aligned}
\tilde{b}_1 &= \frac{h_0}{2}\omega_A^2, \\
\tilde{b}_2 &= \sqrt{h_{-1}2\ln 2}\; \omega_A, \\
\tilde{b}_3 &= (h_{-2}/6)^{-1/3}(2\pi)^{-2/3}\omega_A^{-2/3}.
\end{aligned}
$$

Numerical simulation of the clock feedback loop [78] reveal that the dimensionless time $b_\alpha T$ is related to the prior distribution width of a stabilized clock by the relation $(\delta\phi)^2 = (b_\alpha T)^\alpha$, where $b_\alpha = \chi(\alpha)^{1/\alpha}\tilde{b}_\alpha$ is a rescaled bandwidth, differing from $\tilde{b}_\alpha$ only by an empirically determined prefactor $\chi \approx 1, 1.8, 2$ for $\alpha = 1, 2, 3$. For a laser spectrum containing all three contributions Eq. (A1) can still be used to determine an effective bandwidth, and servo loop simulations of the clock can reveal the modified time dependence of the prior distribution width enabling one to extend the clock model to realistic laser noise parameters.

Appendix B: Spin x-parity in entangling and decoding circuits

We consider global rotations $\mathcal{R}_\mu$, OAT interactions $\mathcal{T}_\mu$ (see Sec. II B) and finite range dressing interactions $\mathcal{D}_\mu$ (see Sec. II G) with $\mu = x, y, x$ as resources for the variational optimization. Within is this set of resources we are able to ensure an anti-symmetric estimator by imposing invariance under the spin x-parity P_x on the Entangler and Decoder, i.e. $P_x\mathcal{U}_{\mathrm{En}}\mathcal{R}_y(-\pi/2)P_x = \mathcal{U}_{\mathrm{En}}\mathcal{R}_y(-\pi/2)$ and $P_x\mathcal{U}_{\mathrm{De}}P_x = \mathcal{U}_{\mathrm{De}}$ under the spin x-parity $P_x = \mathcal{R}_x(\pi/2)$, since this implies

$$
\begin{aligned}
\bar{\phi}_{\mathrm{est}}(\phi) &= \langle\psi_0|\mathcal{U}_{\mathrm{En}}^\dagger e^{i\phi J_z}\mathcal{U}_{\mathrm{De}}^\dagger J_y\mathcal{U}_{\mathrm{De}}e^{-i\phi J_z}\mathcal{U}_{\mathrm{En}}|\psi_0\rangle \\
&= -\langle\psi_0|\mathcal{U}_{\mathrm{En}}^\dagger e^{-i\phi J_z}\mathcal{U}_{\mathrm{De}}^\dagger J_y\mathcal{U}_{\mathrm{De}}e^{i\phi J_z}\mathcal{U}_{\mathrm{En}}|\psi_0\rangle \\
&= -\bar{\phi}_{\mathrm{est}}(-\phi) \qquad (B1)
\end{aligned}
$$

where we use that $P_x J_x P_x = J_x$, $P_x J_{y,z} P_x = -J_{y,z}$, $P_x^\dagger = P_x$ and $P_x\mathcal{R}_y(\pi/2)|\psi_0\rangle = \mathcal{R}_y(\pi/2)|\psi_0\rangle$. The most general entangling and decoding sequences satisfying these constraints are used in Eq. (6),(7) and displayed in Fig. 1.

Appendix C: Wigner distribution

In Secs. II C, II E we visualize collective spin operators $\mathcal{O}$ like the density matrix $\rho = |\psi_{\mathrm{in}}\rangle\langle\psi_{\mathrm{in}}|$ of the initial state of the interferometer [Eq. (4)], or the variationally decoded measurement operator $M = \sum_m \phi_{\mathrm{est}}(m)\mathcal{U}_{\mathrm{De}}^\dagger|m\rangle\langle m|\mathcal{U}_{\mathrm{De}}$ [decomposed in terms of the projection in Eq. (5)] by means of the Wigner distribution [109].

To obtain the Wigner distribution, the operator is

expanded in terms of spherical tensors

$$T_{k,q} = \sum_{m,m'=-N/2}^{N/2} (-1)^{j-m}\sqrt{2k+1}$$
$$\times \begin{pmatrix} j & k & j \\ -m & q & m' \end{pmatrix} |m\rangle \langle m'|, \qquad (C1)$$

where $\begin{pmatrix} j & k & j \\ -m & q & m' \end{pmatrix}$ denotes the Wigner $3j$ symbol. $\mathcal{O}$ can be represented in the spherical tensor basis

$$\mathcal{O} = \sum_{k=0}^{N} \sum_{q=-k}^{k} c_{k,q} T_{k,q} \qquad (C2)$$

where $c_{k,q} = \mathrm{Tr}\left(\mathcal{O}T_{k,q}\right)$. Replacing $T_{k,q}$ in this representation by spherical harmonics $Y_{k,q}(\theta,\phi)$, one arrives at the Wigner distribution,

$$W_{\mathcal{O}}(\theta,\phi) = \sum_{k=0}^{N} \sum_{q=-k}^{k} c_{k,q} Y_{k,q}(\theta,\phi) \qquad (C3)$$

as a quasi-probability distribution on a generalized Bloch sphere.

The Wigner function can be used to calculate the expectation value

$$\mathrm{Tr}(\rho M) = \int_0^{\pi} d\theta \int_0^{2\pi} d\varphi\, W_M(\theta,\varphi) W_\rho(\theta,\varphi) \qquad (C4)$$

by integrating the overlap of the respective Wigner functions over the generalized Bloch sphere. This implies that we can interpret contours of the measurement distribution with the different eigenvalues of the measurement operator while the amplitude of the state distribution indicates how much the state overlaps with the respective projection of the measurement projection.

Appendix D: Numerical optimization of the phase-operator based interferometer

Here we define the phase operator and describe an iterative optimization procedure allowing us to minimize the cost function (2) for a given observable using the Minimal Mean Squared Error (MMSE) estimator [3]. The phase operator $\hat{\Phi}$ reads [87, 88]:

$$\hat{\Phi} = \sum_{s=-J}^{J} \phi_s |s\rangle \langle s|, \qquad (D1)$$

$$\phi_s = \frac{2\pi s}{2J+1}, \qquad (D2)$$

$$|s\rangle = \frac{1}{\sqrt{2J+1}} \sum_{m=-J}^{J} e^{-i\phi_s m} |m\rangle, \qquad (D3)$$

where $J_z |m\rangle = m|m\rangle$.

Our goal is to minimize the cost function Eq. (2) for the observable $\hat{\Phi}$ and the MMSE estimator by finding the optimal initial state $|\psi_{\hat{\Phi}}\rangle$. The MMSE estimator reads [3]:

$$\phi_{\mathrm{est}}^{\mathrm{MMSE}}(s) = \int \phi\, p(\phi|s)d\phi, \qquad (D4)$$

where the conditional probability is $p(\phi|s) \propto p(s|\phi)\mathcal{P}(\phi)$ with $p(s|\phi) = |\langle s|e^{-i\phi \hat{J}_z}|\psi_{\mathrm{in}}\rangle|^2$ and the observable eigenstate $|s\rangle$ defined in Eq. (D3).

The optimization is performed iteratively. Initially we start with $s = 0$ eigenstate of $\hat{\Phi}$ as the input state $|\psi_{\mathrm{in}}^{(0)}\rangle = |s=0\rangle$, which is a good approximation for a state highly sensitive to phases around $\phi = 0$. The state defines the corresponding MMSE estimator $\phi_{\mathrm{est}(0)}^{\mathrm{MMSE}}(s)$ as given by Eq. (D4). In the next iteration we find the state $|\psi_{\mathrm{in}}^{(1)}\rangle$ minimizing the cost function (2) for the given $\phi_{\mathrm{est}(0)}^{\mathrm{MMSE}}(s)$ estimator by solving a corresponding eigenproblem, as described in [31]. The iterative procedure converges quickly yielding the optimal initial state for the POI $|\psi_{\mathrm{in}}^{(k)}\rangle \to_{k\to\infty} |\psi_{\hat{\Phi}}\rangle$ which, in turn, defines the optimal estimator via Eq. (D4) and the corresponding posterior width $\Delta\phi_{\mathrm{POI}}$. This result is used in Sec. II D.

Appendix E: N-scaling of the optimal quantum clock instability

Here we derive asymptotic scaling of the optimal interrogation time and the corresponding minimal instability of the optimal quantum clock. As discussed in Sec. III, the instability of clocks exploiting single-shot protocols is fundamentally limited by the measurement shot noise given by the π-corrected HL for short interrogation times T, and the laser CTL for large T. For the dimensionless Allan variance we write

$$(\sigma^{\mathrm{OQC}})^2 = \sigma_{\pi\mathrm{HL}}^2 + (\sigma_{\mathrm{CTL}}^{\mathrm{OQC}})^2,$$
$$= \frac{1}{s\pi^{\frac{2}{\alpha}}} \left\{ \frac{\pi^2}{N^2} + 4\pi^2 \left[1 - \mathrm{erf}\left(\frac{1}{\sqrt{2s^\alpha}}\right) \right] \right\}. \qquad (E1)$$

where $s \equiv \pi^{-\frac{2}{\alpha}} b_\alpha T$ is the dimensionless Ramsey time. The goal is to minimize Eq. (E1) with respect to s in the limit of large number of atoms, $N \to \infty$. The derivative with respect to s reads

$$\frac{d}{ds}(\sigma^{\mathrm{OQC}})^2 = -\frac{1}{s^2\pi^{\frac{4}{\alpha}}} \left\{ \frac{\pi^2}{N^2} + 4\pi^2 \left[1 - \mathrm{erf}\left(\frac{1}{\sqrt{2s^\alpha}}\right) \right] \right.$$
$$\left. - 2\sqrt{\frac{2\pi^3\alpha^2}{s^\alpha}} e^{-\frac{1}{2s^\alpha}} \right\},$$

and, using a self-consistent assumption for optimal time $s_* \ll 1$, results in the following equation for s_*,

$$e^{s_*^{-\alpha}} s_*^\alpha = \frac{8\alpha^2 N^4}{\pi}. \qquad (E2)$$

Here we used the error function asymptotic $1 - \mathrm{erf}(x) \to e^{-x^2}/(\sqrt{\pi}x)$ for $x \to \infty$. Taking the logarithm of the

expression (E2) (s_*, α, and N are positive) we obtain an equation for $w \equiv s_*^{-\alpha}$,

$$w - \ln w = \ln z,$$

with $z \equiv 8\alpha^2 N^4/\pi$. For $z > e$, the solution can be written as the infinitely nested logarithm, $w(z) = \ln(z \ln(z \ln(z(\ln \ldots) \ldots)))$, and can be checked by direct substitution. Using the $w(z)$ function we can express the optimal Ramsey time for $N \gg 1$ as follows

$$b_\alpha T_{\text{opt}} = \pi^{\frac{2}{\alpha}} s_* = \left[\frac{1}{\pi^2} w(z) \right]^{-\frac{1}{\alpha}} \qquad \text{(E3)}$$

Finally, we substitute the optimal Ramsey time into Eq. (E1)

$$(\sigma_{\text{opt}}^{\text{OQC}})^2 \simeq \frac{\pi^2}{N^2} \left[\frac{w(z)}{\pi^2} \right]^{\frac{1}{\alpha}} \left[1 + \frac{2}{\alpha w(z)} \right]. \qquad \text{(E4)}$$

We use Eqs. (E3) and (E4) and keep only the first two logarithms in the definition of $w(z)$ to obtain expressions for the optimal interrogation time and minimal instability of the optimal quantum clocks in Sec. III for $\alpha = 2$.

Appendix F: Finite dead time in the atomic clock protocol

Here we discuss upper limits to the dead times of atomic clocks, which are required to reach the variationally optimized stability presented in Sec. III. When each interrogation cycle of duration $T_C = T_D + T$ is composed of a dead time $T_D > 0$, and Ramsey free evolution time T, the stability is reduced compared to the ideal case at $T_D = 0$ discussed in the main text.

Let us consider $S_L(f) = h_{-1}f^{-1}$ as the power spectral density of the free running laser. In addition, we assume that the protocols are sensitive to phase shifts during T only and that all entangling and decoding operations are included in the dead time where we assume no sensitivity. Given these assumptions, the instability contribution of the Dick effect is [110]

$$\sigma_{\text{Dick}}^2(\tau) = \frac{b_2}{\omega_A^2 \tau} \frac{b_2 T}{\chi(2) 2 \ln 2} \frac{1}{d^3} \sum_{n=1}^{\infty} \frac{\sin^2(\pi n d)}{\pi^2 n^3} \qquad \text{(F1)}$$

with χ given in App. A and the duty cycle $d = T/T_C$. In addition, the instability predicted in the Bayesian framework, Eq. (24), becomes

$$\sigma_{\text{Bay}}^2(\tau) = \frac{b_2}{\omega_A^2 \tau} \frac{\sigma^2}{d} \qquad \text{(F2)}$$

with σ as defined in Eq. (25).

In the following we want to estimate below which level of dead time the combined instability $\sigma_y(\tau) = \sqrt{\sigma_{\text{Bay}}^2(\tau) + \sigma_{\text{Dick}}^2(\tau)}$ is no longer dominated by the contribution of the Dick effect. The minimal required duty
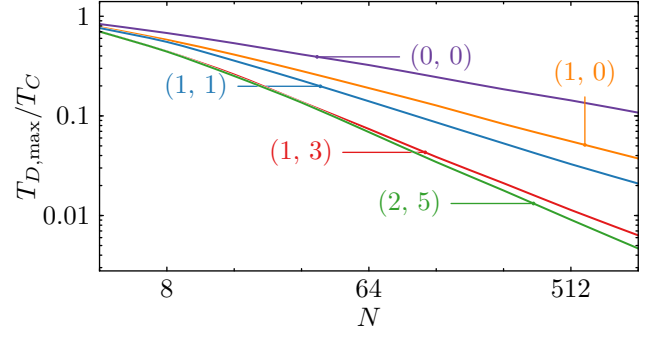

FIG. 12. Largest fraction $T_{D,\text{max}}/T_C$ of dead time compared to cycle time for which the clock operation is limited predominantly by the variationally optimized instability displayed in Fig. 11. The overall instabilities at dead time fractions above the lines are limited by the Dick effect instead.

cycle d_{min} where the value for $\sigma_{\text{Dick}}^2(\tau)$ at optimal Ramsey time $b_2 T_{\text{opt}}$ dives below the lowest variational instability is

$$d_{\text{min}} = \min \left\{ d \Big| \frac{b_2 T_{\text{opt}}}{\chi(2) 2 \ln 2} \frac{1}{d^2} \sum_{n=1}^{\infty} \frac{\sin^2(\pi n d)}{\pi^2 n^3} \leq \sigma_{\text{opt}}^2 \right\}. \qquad \text{(F3)}$$

From d_{min} one can directly infer the maximum fraction $R = T_{D,\text{max}}/T_C = 1 - T_{\text{opt}}/T_C = 1 - d_{\text{min}}$ of dead time in the clock cycle, where $T_C = T_{\text{opt}} + T_{D,\text{max}}$. In the limit $R \ll 1$ it can be shown that $-\ln(R)R^2/(1-R)^2 \propto (b_2 T_{\text{opt}})\sigma_{\text{opt}}^2$, so it is expected that for $N \gg 1$ this ratio will eventually follow a similar scaling as σ_{opt}^2. The exact relation is shown in Fig. 12. It is worth noting that $R \ll 1$ is still recommended for small ensemble sizes, even though this condition is not required based on d_{min}, to prevent unnecessarily increasing the clock instability. A more complete model for the influence of dead time and the Dick effect requires to include the full spectral density $S_L(f)$ of the laser and evaluating the sensitivity function during the entangling and decoding dynamics.

Appendix G: Numerical simulation of the variational clock operation

In order to see how well σ [Eq. (25)] reflects an achievable instability we perform numerical simulations of all essential parts involved in the closed feedback loop of an optical atomic clock when operating with the variationally optimized Ramsey protocols.

Building up the simulations proceeds as follows:
(i) The free-running laser is simulated. Given a particular spectral density $S_L(f) = h_{1-\alpha}f^{1-\alpha}$ and the Ramsey time T we generate a sequence of random numbers $\bar{y}_k = \frac{1}{T} \int_{t_k}^{t_k+T} dt[\omega_L(t) - \omega_A]/\omega_A$ which gives the average frequency fluctuations of the laser without any feedback in each cycle k. Correlations between different cycles, required when $\alpha \neq 1$, can e.g. be obtained in the time

domain by implementing $\bar{y}_k$ as a random walk or a sum of multiple damped random walks [78].

(ii) To stabilise the laser frequency for long averaging times $\tau \gg T$ a feedback correction is applied to the laser frequency at the end of each cycle. In the simulations, the estimated frequency deviation $\bar{y}_{\mathrm{est,k}} = m_k/(2\pi\omega_A T\partial_\phi\bar{m}(\phi)_{|\phi=0})$ obtained from measurement result m_k at t_k is multiplied by a gain factor $0 < g \leq 1$ and subtracted from the true laser frequency. This integrating servo corrects frequency errors over $\sim 1/g$ cycles and is sufficient to achieve a robust stabilization at $\tau/T \gg 1/g$ for flicker noise limited lasers [79]. However, to simulate the quantum probabilities $p(m|\phi_k)$ at t_k the phase $\phi_k = \omega_A T\bar{y}'_k$ based on the actual laser noise $\bar{y}'_k$ is needed. Thus, later measurements are affected not only by the noise of the free-running laser but also by the measurement results and corrections from earlier cycles. To implement this efficiently, the simulation runs sequentially: At the beginning the phase ϕ_1 is calculated for the first cycle only. Then the probabilities $p(m|\phi_1)$ with this particular phase are calculated and a single measurement result m_1 is sampled according to this distribution. The estimator $y_{\mathrm{est,1}}$ is calculated and the servo corrects the laser frequency so that $\bar{y}'_2 = \bar{y}_2 - g\bar{y}_{\mathrm{est,1}}$ is the actual noise in the second cycle. This procedure is repeated in each cycle with the corrected frequencies, meaning e.g. $\phi_2 = \omega_A T\bar{y}'_2$.

(iii) The clock stability is evaluated, based on the simulated sequence of stabilized frequency deviations $\bar{y}'_k$. The overlapping Allan deviation $\sigma_y(\tau = nT)$ is calculated numerically from averages over n cycles. Statistical averaging is performed over many intervals of length n in a single run with $n_{\mathrm{tot}} \gg n$ cycles and then averaging again over multiple runs. Finally, the long term instability is extracted by fitting the prefactor to the asymptotic scaling $\sigma_y(\tau) \propto \tau^{1/2}$ reached typically after $n \sim 10^4$ cycles in simulations of $n_{\mathrm{tot}} = 2 \times 10^6$ cycles.

To compare numerical results to theory predictions, as in Fig. 10(b), the values for T and $h_{1-\alpha}$ in the simulations are matched to reproduce the same laser induced prior width $(\delta\phi)^2 = (b_\alpha T)^\alpha$.

Appendix H: Cumulative interaction angle

A relevant question regarding the Dick effect is the time it takes to perform the entangling and decoding sequence. The slowest time scale on a quantum simulator is usually the interaction strength. Results presented in Fig. 2, 10 were obtained for interaction angles $\leq \pi/2$. From a practical point of view, however, it might be beneficial to consider smaller interaction angles.

Here we show that, close to the respective minima in Fig. 2, 10, the displayed results of the variationally optimized interferometers can be well approximated by quantum circuits with small cumulative interaction angles $\theta_{\mathrm{OAT}} = \sum_{k=1}^{n_{\mathrm{En}}}\left(\theta_k^{(1)} + \theta_k^{(2)}\right) + \sum_{k=1}^{n_{\mathrm{De}}}\left(\vartheta_k^{(1)} + \vartheta_k^{(2)}\right)$. In Fig. 13 we constrain each interaction angle to be positive and smaller than a threshold that decreases with
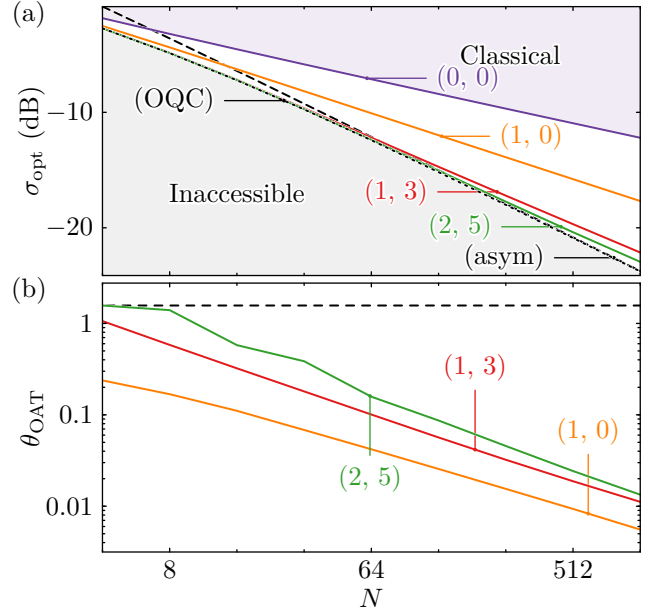

FIG. 13. (a) Recalculation of the particle number dependence of the dimensionless Allan deviation in Fig. 11 for constrained interaction angles. (b) The cumulative angle of all one axis twisting gates $\mathcal{T}_{x,y}$ required to obtain the dimensionless Allan deviations displayed above. The vertical dashed line indicated the interaction angle of $\pi/2$ required to prepare a GHZ-state.

the depth of the circuit. In addition we require that the cumulative interaction angle θ_{OAT} is always smaller or equal than $\pi/2$, the interaction angle required to prepare a GHZ-state. Similarly to the OAT squeezing [73], the variational sequences can also work with a cumulative interaction that decrease rapidly with N while the resulting Allan deviation remains a good approximation of the unconstrained optimization in Fig. 11.

[1] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt, Optical atomic clocks, Rev. Mod. Phys. **87**, 637 (2015).

[2] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, Optics and interferometry with atoms and molecules, Rev. Mod. Phys. **81**, 1051 (2009).

[3] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, *Quantum Limits in Optical Interferometry*, Progress in Optics, Vol. 60 (Elsevier, 2015) pp. 345 – 435.

[4] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[5] J. Grotti, S. Koller, S. Vogt, S. Häfner, U. Sterr, C. Lisdat, H. Denker, C. Voigt, L. Timmen, A. Rolland, F. N.

Baynes, H. S. Margolis, M. Zampaolo, P. Thoumany, M. Pizzocaro, B. Rauf, F. Bregolin, A. Tampellini, P. Barbieri, M. Zucco, G. A. Costanzo, C. Clivati, F. Levi, and D. Calonico, Geodesy and metrology with a transportable optical clock, Nature Physics 14, 437 (2018).

[6] T. E. Mehlstäubler, G. Grosche, C. Lisdat, P. O. Schmidt, and H. Denker, Atomic clocks for geodesy, Reports on Progress in Physics 81, 064401 (2018).

[7] T. Bothwell, C. J. Kennedy, A. Aeppli, D. Kedar, J. M. Robinson, E. Oelker, A. Staron, and J. Ye, Resolving the gravitational redshift within a millimeter atomic sample (2021), arXiv:2109.12238 [physics.atom-ph].

[8] X. Zheng, J. Dolde, V. Lochab, B. N. Merriman, H. Li, and S. Kolkowitz, High precision differential clock comparisons with a multiplexed optical lattice clock (2021), arXiv:2109.12237 [physics.atom-ph].

[9] S. Kolkowitz, I. Pikovski, N. Langellier, M. D. Lukin, R. L. Walsworth, and J. Ye, Gravitational wave detection with optical lattice atomic clocks, Phys. Rev. D 94, 124043 (2016).

[10] M. S. Safronova, D. Budker, D. DeMille, D. F. J. Kimball, A. Derevianko, and C. W. Clark, Search for new physics with atoms and molecules, Rev. Mod. Phys. 90, 025008 (2018).

[11] C. Sanner, N. Huntemann, R. Lange, C. Tamm, E. Peik, M. S. Safronova, and S. G. Porsev, Optical clock comparison for lorentz symmetry testing, Nature 567, 204 (2019).

[12] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. 90, 035005 (2018).

[13] M. Tse, H. Yu, N. Kijbunchoo, A. Fernandez-Galiana, P. Dupej, L. Barsotti, C. D. Blair, D. D. Brown, S. E. Dwyer, A. Effler, M. Evans, P. Fritschel, V. V. Frolov, A. C. Green, G. L. Mansell, F. Matichard, N. Mavalvala, D. E. McClelland, L. McCuller, T. McRae, J. Miller, A. Mullavey, E. Oelker, I. Y. Phinney, D. Sigg, B. J. J. Slagmolen, T. Vo, R. L. Ward, C. Whittle, R. Abbott, C. Adams, R. X. Adhikari, A. Ananyeva, S. Appert, K. Arai, J. S. Areeda, Y. Asali, S. M. Aston, C. Austin, A. M. Baer, M. Ball, S. W. Ballmer, S. Banagiri, D. Barker, J. Bartlett, B. K. Berger, J. Betzwieser, D. Bhattacharjee, G. Billingsley, S. Biscans, R. M. Blair, N. Bode, P. Booker, R. Bork, A. Bramley, A. F. Brooks, A. Buikema, C. Cahillane, K. C. Cannon, X. Chen, A. A. Ciobanu, F. Clara, S. J. Cooper, K. R. Corley, S. T. Countryman, P. B. Covas, D. C. Coyne, L. E. H. Datrier, D. Davis, C. Di Fronzo, J. C. Driggers, T. Etzel, T. M. Evans, J. Feicht, P. Fulda, M. Fyffe, J. A. Giaime, K. D. Giardina, P. Godwin, E. Goetz, S. Gras, C. Gray, R. Gray, A. Gupta, E. K. Gustafson, R. Gustafson, J. Hanks, J. Hanson, T. Hardwick, R. K. Hasskew, M. C. Heintze, A. F. Helmling-Cornell, N. A. Holland, J. D. Jones, S. Kandhasamy, S. Karki, M. Kasprzack, K. Kawabe, P. J. King, J. S. Kissel, R. Kumar, M. Landry, B. B. Lane, B. Lantz, M. Laxen, Y. K. Lecoeuche, J. Leviton, J. Liu, M. Lormand, A. P. Lundgren, R. Macas, M. MacInnis, D. M. Macleod, S. Márka, Z. Márka, D. V. Martynov, K. Mason, T. J. Massinger, R. McCarthy, S. McCormick, J. McIver, G. Mendell, K. Merfeld, E. L. Merilh, F. Meylahn, T. Mistry, R. Mittleman, G. Moreno, C. M. Mow-Lowry, S. Mozzon, T. J. N. Nelson, P. Nguyen, L. K. Nut-

tall, J. Oberling, R. J. Oram, B. O'Reilly, C. Osthelder, D. J. Ottaway, H. Overmier, J. R. Palamos, W. Parker, E. Payne, A. Pele, C. J. Perez, M. Pirello, H. Radkins, K. E. Ramirez, J. W. Richardson, K. Riles, N. A. Robertson, J. G. Rollins, C. L. Romel, J. H. Romie, M. P. Ross, K. Ryan, T. Sadecki, E. J. Sanchez, L. E. Sanchez, T. R. Saravanan, R. L. Savage, D. Schaetzl, R. Schnabel, R. M. S. Schofield, E. Schwartz, D. Sellers, T. J. Shaffer, J. R. Smith, S. Soni, B. Sorazu, A. P. Spencer, K. A. Strain, L. Sun, M. J. Szczepańczyk, M. Thomas, P. Thomas, K. A. Thorne, K. Toland, C. I. Torrie, G. Traylor, A. L. Urban, G. Vajente, G. Valdes, D. C. Vander-Hyde, P. J. Veitch, K. Venkateswara, G. Venugopalan, A. D. Viets, C. Vorvick, M. Wade, J. Warner, B. Weaver, R. Weiss, B. Willke, C. C. Wipf, L. Xiao, H. Yamamoto, M. J. Yap, H. Yu, L. Zhang, M. E. Zucker, and J. Zweizig, Quantum-enhanced advanced ligo detectors in the era of gravitational-wave astronomy, Phys. Rev. Lett. 123, 231107 (2019).

[14] C. A. Casacio, L. S. Madsen, A. Terrasson, M. Waleed, K. Barnscheidt, B. Hage, M. A. Taylor, and W. P. Bowen, Quantum-enhanced nonlinear microscopy, Nature 594, 201 (2021).

[15] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, R. Reichle, and D. J. Wineland, Creation of a six-atom 'schrödinger cat' state, Nature 438, 639 (2005).

[16] J. Appel, P. J. Windpassinger, D. Oblak, U. B. Hoff, N. Kjærgaard, and E. S. Polzik, Mesoscopic atomic entanglement for precision measurements beyond the standard quantum limit, Proceedings of the National Academy of Sciences 106, 10960 (2009), https://www.pnas.org/content/106/27/10960.full.pdf.

[17] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler, Nonlinear atom interferometer surpasses classical precision limit, Nature 464, 1165 (2010).

[18] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, Atom-chip-based generation of entanglement for quantum metrology, Nature 464, 1170 (2010).

[19] I. D. Leroux, M. H. Schleier-Smith, and V. Vuletić, Implementation of cavity squeezing of a collective atomic spin, Phys. Rev. Lett. 104, 073602 (2010).

[20] B. Lücke, M. Scherer, J. Kruse, L. Pezzé, F. Deuretzbacher, P. Hyllus, O. Topic, J. Peise, W. Ertmer, J. Arlt, L. Santos, A. Smerzi, and C. Klempt, Twin matter waves for interferometry beyond the classical limit, Science 334, 773 (2011), https://www.science.org/doi/pdf/10.1126/science.1208798.

[21] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell, Magnetic sensitivity beyond the projection noise limit by spin squeezing, Phys. Rev. Lett. 109, 253605 (2012).

[22] C. D. Hamley, C. S. Gerving, T. M. Hoang, E. M. Bookjans, and M. S. Chapman, Spin-nematic squeezed vacuum in a quantum gas, Nature Physics 8, 305 (2012).

[23] T. Berrada, S. van Frank, R. Bücker, T. Schumm, J. F. Schaff, and J. Schmiedmayer, Integrated mach–zehnder interferometer for bose–einstein condensates, Nature Communications 4, 2077 (2013).

[24] G. Barontini, L. Hohmann, F. Haas, J. Estève, and J. Reichel, Deterministic generation of multiparticle entanglement by quantum

zeno dynamics, Science **349**, 1317 (2015), https://www.science.org/doi/pdf/10.1126/science.aaa0754.

[25] Y.-L. Zhang, C.-L. Zou, X.-B. Zou, L. Jiang, and G.-C. Guo, Detuning-enhanced cavity spin squeezing, Phys. Rev. A **91**, 033625 (2015).

[26] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, Quantum spin dynamics and entanglement generation with hundreds of trapped ions, Science **352**, 1297 (2016).

[27] O. Hosten, N. J. Engelsen, R. Krishnakumar, and M. A. Kasevich, Measurement noise 100 times lower than the quantum-projection limit using entangled atoms, Nature **529**, 505 (2016).

[28] M. A. Norcia, R. J. Lewis-Swan, J. R. K. Cline, B. Zhu, A. M. Rey, and J. K. Thompson, Cavity-mediated collective spin-exchange interactions in a strontium superradiant laser, Science **361**, 259 (2018), https://www.science.org/doi/pdf/10.1126/science.aar3102.

[29] Y.-Q. Zou, L.-N. Wu, Q. Liu, X.-Y. Luo, S.-F. Guo, J.-H. Cao, M. K. Tey, and L. You, Beating the classical precision limit with spin-1 dicke states of more than 10,000 atoms, Proceedings of the National Academy of Sciences **115**, 6381 (2018), https://www.pnas.org/content/115/25/6381.full.pdf.

[30] E. Pedrozo-Peñafiel, S. Colombo, C. Shu, A. F. Adiyatullin, Z. Li, E. Mendez, B. Braverman, A. Kawasaki, D. Akamatsu, Y. Xiao, and V. Vuletić, Entanglement on an optical atomic-clock transition, Nature **588**, 414 (2020).

[31] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New Journal of Physics **16**, 113002 (2014).

[32] S. Personick, Application of quantum estimation theory to analog communication over quantum channels, IEEE Transactions on Information Theory **17**, 240 (1971).

[33] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland series in statistics and probability (North-Holland Publishing Company, 1982).

[34] V. Bužek, R. Derka, and S. Massar, Optimal quantum clocks, Phys. Rev. Lett. **82**, 2207 (1999).

[35] M. Jarzyna and R. Demkowicz-Dobrzański, True precision limits in quantum metrology, New Journal of Physics **17**, 013010 (2015).

[36] E. Martínez-Vargas, C. Pineda, F. Leyvraz, and P. Barberis-Blostein, Quantum estimation of unknown parameters, Phys. Rev. A **95**, 012136 (2017).

[37] W. Górecki, R. Demkowicz-Dobrzański, H. M. Wiseman, and D. W. Berry, π-corrected heisenberg limit, Phys. Rev. Lett. **124**, 030501 (2020).

[38] B. J. Bloom, T. L. Nicholson, J. R. Williams, S. L. Campbell, M. Bishof, X. Zhang, W. Zhang, S. L. Bromley, and J. Ye, An optical lattice clock with accuracy and stability at the 10^{-18} level, Nature **506**, 71 (2014).

[39] I. Ushijima, M. Takamoto, M. Das, T. Ohkubo, and H. Katori, Cryogenic optical lattice clocks, Nature Photonics **9**, 185 (2015).

[40] W. F. McGrew, X. Zhang, R. J. Fasano, S. A. Schäffer, K. Beloy, D. Nicolodi, R. C. Brown, N. Hinkley, G. Milani, M. Schioppa, T. H. Yoon, and A. D. Ludlow, Atomic clock performance enabling geodesy below the centimetre level, Nature **564**, 87 (2018).

[41] S. Origlia, M. S. Pramod, S. Schiller, Y. Singh, K. Bongs, R. Schwarz, A. Al-Masoudi, S. Dörscher, S. Herbers, S. Häfner, U. Sterr, and C. Lisdat, Towards an optical clock for space: Compact, high-performance optical lattice clock based on bosonic atoms, Phys. Rev. A **98**, 053443 (2018).

[42] E. Oelker, R. B. Hutson, C. J. Kennedy, L. Sonderhouse, T. Bothwell, A. Goban, D. Kedar, C. Sanner, J. M. Robinson, G. E. Marti, D. G. Matei, T. Legero, M. Giunta, R. Holzwarth, F. Riehle, U. Sterr, and J. Ye, Demonstration of 4.8×10^{-17} stability at 1 s for two independent optical clocks, Nature Photonics **13**, 714 (2019).

[43] C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband, Frequency comparison of two high-accuracy al^{+} optical clocks, Phys. Rev. Lett. **104**, 070802 (2010).

[44] N. Huntemann, C. Sanner, B. Lipphardt, C. Tamm, and E. Peik, Single-ion atomic clock with 3×10^{-18} systematic uncertainty, Phys. Rev. Lett. **116**, 063001 (2016).

[45] S. M. Brewer, J.-S. Chen, A. M. Hankin, E. R. Clements, C. W. Chou, D. J. Wineland, D. B. Hume, and D. R. Leibrandt, ^{27}al^{+} quantum-logic clock with a systematic uncertainty below 10^{-18}, Phys. Rev. Lett. **123**, 033201 (2019).

[46] J. P. Covey, I. S. Madjarov, A. Cooper, and M. Endres, 2000-times repeated imaging of strontium atoms in clock-magic tweezer arrays, Phys. Rev. Lett. **122**, 173201 (2019).

[47] J. Wilson, S. Saskin, Y. Meng, S. Ma, R. Dilip, A. Burgers, and J. Thompson, Trapped arrays of alkaline earth rydberg atoms in optical tweezers (2019), arXiv:1912.08754 [quant-ph].

[48] A. W. Young, W. J. Eckner, W. R. Milner, D. Kedar, M. A. Norcia, E. Oelker, N. Schine, J. Ye, and A. M. Kaufman, Half-minute-scale atomic coherence and high relative stability in a tweezer clock, Nature **588**, 408 (2020).

[49] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), arXiv:1411.4028 [quant-ph].

[50] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).

[51] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New Journal of Physics **18**, 023023 (2016).

[52] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Scalable quantum simulation of molecular energies, Phys. Rev. X **6**, 031007 (2016).

[53] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller, Self-verifying variational quantum simulation of lattice models, Nature **569**, 355 (2019).

[54] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[55] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum approximate optimization algorithm:

Performance, mechanism, and implementation on near-term devices, Phys. Rev. X **10**, 021067 (2020).

[56] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nature Reviews Physics **3**, 625 (2021).

[57] O. Hosten, R. Krishnakumar, N. J. Engelsen, and M. A. Kasevich, Quantum phase magnification, Science **352**, 1552 (2016), https://www.science.org/doi/pdf/10.1126/science.aaf3397.

[58] E. Davis, G. Bentsen, and M. Schleier-Smith, Approaching the heisenberg limit without single-particle detection, Phys. Rev. Lett. **116**, 053601 (2016).

[59] F. Fröwis, P. Sekatski, and W. Dür, Detecting large quantum fisher information with finite measurement precision, Phys. Rev. Lett. **116**, 090801 (2016).

[60] T. Macrì, A. Smerzi, and L. Pezzè, Loschmidt echo for quantum metrology, Phys. Rev. A **94**, 010102(R) (2016).

[61] S. P. Nolan, S. S. Szigeti, and S. A. Haine, Optimal and robust quantum metrology using interaction-based readouts, Phys. Rev. Lett. **119**, 193601 (2017).

[62] F. Anders, L. Pezzè, A. Smerzi, and C. Klempt, Phase magnification by two-axis countertwisting for detection-noise robust interferometry, Phys. Rev. A **97**, 043813 (2018).

[63] M. Schulte, V. J. Martínez-Lahuerta, M. S. Scharnagl, and K. Hammerer, Ramsey interferometry with generalized one-axis twisting echoes, Quantum **4**, 268 (2020).

[64] I. H. Deutsch, Harnessing the power of the second quantum revolution, PRX Quantum **1**, 020101 (2020).

[65] National Academies of Sciences and Medicine, *Manipulating Quantum Systems: An Assessment of Atomic, Molecular, and Optical Physics in the United States* (The National Academies Press, Washington, DC, 2020).

[66] R. Kaubruegger, P. Silvi, C. Kokail, R. van Bijnen, A. M. Rey, J. Ye, A. M. Kaufman, and P. Zoller, Variational spin-squeezing algorithms on programmable quantum sensors, Phys. Rev. Lett. **123**, 260505 (2019).

[67] S. L. Braunstein, C. M. Caves, and G. Milburn, Generalized uncertainty relations: Theory, examples, and lorentz invariance, Annals of Physics **247**, 135 (1996).

[68] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. **96**, 010401 (2006).

[69] M. G. A. Paris, Quantum estimation for quantum technology, International Journal of Quantum Information **07**, 125 (2009).

[70] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Optimal frequency measurements with maximally correlated states, Phys. Rev. A **54**, R4649 (1996).

[71] M. Hayashi, S. Vinjanampathy, and L. C. Kwek, Resolving unattainable cramer–rao bounds for quantum sensors, Journal of Physics B: Atomic, Molecular and Optical Physics **52**, 015503 (2018).

[72] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen, Spin squeezing and reduced quantum noise in spectroscopy, Phys. Rev. A **46**, R6797 (1992).

[73] M. Kitagawa and M. Ueda, Squeezed spin states, Phys. Rev. A **47**, 5138 (1993).

[74] Variational algorithms for quantum metrology optimizing local phase sensitivity, i.e. maximizing the quantum Fisher information, are considered in Refs. [111–113].

[75] While we consider optimal non-adaptive measurement protocols, adaptive protocols for phase estimation have been studied in [102–105, 114–118].

[76] A. André, A. S. Sørensen, and M. D. Lukin, Stability of atomic clocks based on entangled atoms, Phys. Rev. Lett. **92**, 230801 (2004).

[77] M. Fraas, An analysis of the stationary operation of atomic clocks, Communications in Mathematical Physics **348**, 363 (2016).

[78] I. D. Leroux, N. Scharnhorst, S. Hannig, J. Kramer, L. Pelzer, M. Stepanova, and P. O. Schmidt, On-line estimation of local oscillator noise and optimisation of servo parameters in atomic clocks, Metrologia **54**, 307 (2017).

[79] M. Schulte, C. Lisdat, P. O. Schmidt, U. Sterr, and K. Hammerer, Prospects and challenges for squeezing-enhanced optical atomic clocks, Nature Communications **11**, 5955 (2020).

[80] For simplicity of notation we assume pure states. This discussion is readily extended to mixed states with density operator ρ_{in}.

[81] A phase $-\infty < \phi < \infty$ is consistent with a laser phase in atomic clocks. A diffusion of the laser phase as a phase winding outside of $-\pi < \phi < \pi$ corresponds to a phase slip which must be avoided in clock operation.

[82] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, 14-qubit entanglement: Creation and coherence, Phys. Rev. Lett. **106**, 130506 (2011).

[83] I. Pogorelov, T. Feldker, C. D. Marciniak, L. Postler, G. Jacob, O. Krieglsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskiy, R. Blatt, P. Schindler, and T. Monz, Compact ion-trap quantum computing demonstrator, PRX Quantum **2**, 020343 (2021).

[84] For the optimization results presented throughout this manuscript, we calculate exact gradients and use a combination of sequential quadratic programming and the Nelder-Mead method, as implemented in MATLAB, to determine optimal solutions.

[85] Q. Y. He, S.-G. Peng, P. D. Drummond, and M. D. Reid, Planar quantum squeezing and atom interferometry, Phys. Rev. A **84**, 022107 (2011).

[86] B. Braverman, A. Kawasaki, and V. Vuletić, Impact of non-unitary spin squeezing on atomic clock performance, New Journal of Physics **20**, 103019 (2018).

[87] D. T. Pegg and S. M. Barnett, Unitary phase operator in quantum mechanics, Europhysics Letters (EPL) **6**, 483 (1988).

[88] R. Derka, V. Bužek, and A. K. Ekert, Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement, Phys. Rev. Lett. **80**, 1571 (1998).

[89] B. A. Chase and J. M. Geremia, Collective processes of an ensemble of spin-$\frac{1}{2}$ particles, Phys. Rev. A **78**, 052101 (2008).

[90] N. Shammah, S. Ahmed, N. Lambert, S. De Liberato, and F. Nori, Open quantum systems with local and collective incoherent processes: Efficient numerical simulations using permutational invariance, Phys. Rev. A **98**, 063815 (2018).

[91] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Improvement of frequency standards with quantum entanglement, Phys. Rev. Lett. **79**, 3865 (1997).

[92] S. L. Braunstein and C. M. Caves, Statistical distance

and the geometry of quantum states, Phys. Rev. Lett. **72**, 3439 (1994).

[93] C. W. Helstrom, *Quantum detection and estimation theory*, Math. Sci. Eng. (Academic Press, New York, NY, 1976).

[94] H. L. V. Trees, *Detection, Estimation and Modulation* (Wiley, New York, 1968).

[95] R. D. Gill and B. Y. Levit, Applications of the van trees inequality: a bayesian cramér-rao bound, Bernoulli **1**, 59 (1995).

[96] D. W. Berry and H. M. Wiseman, Optimal states and almost optimal adaptive measurements for quantum interferometry, Phys. Rev. Lett. **85**, 5098 (2000).

[97] N. Henkel, R. Nath, and T. Pohl, Three-dimensional roton excitations and supersolid formation in rydberg-excited bose-einstein condensates, Phys. Rev. Lett. **104**, 195302 (2010).

[98] G. Pupillo, A. Micheli, M. Boninsegni, I. Lesanovsky, and P. Zoller, Strongly correlated gases of rydberg-dressed atoms: Quantum and classical dynamics, Phys. Rev. Lett. **104**, 223002 (2010).

[99] L. I. R. Gil, R. Mukherjee, E. M. Bridge, M. P. A. Jones, and T. Pohl, Spin squeezing in a rydberg lattice clock, Phys. Rev. Lett. **112**, 103601 (2014).

[100] M. Schioppo, R. C. Brown, W. F. McGrew, N. Hinkley, R. J. Fasano, K. Beloy, T. H. Yoon, G. Milani, D. Nicolodi, J. A. Sherman, N. B. Phillips, C. W. Oates, and A. D. Ludlow, Ultrastable optical clock with two cold-atom ensembles, Nature Photonics **11**, 48 (2017).

[101] J. Borregaard and A. S. Sørensen, Near-heisenberg-limited atomic clocks in the presence of decoherence, Phys. Rev. Lett. **111**, 090801 (2013).

[102] E. M. Kessler, P. Kómár, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin, Heisenberg-limited atom clocks based on entangled qubits, Phys. Rev. Lett. **112**, 190403 (2014).

[103] J. Borregaard and A. S. Sørensen, Efficient atomic clocks operated with several atomic ensembles, Phys. Rev. Lett. **111**, 090802 (2013).

[104] T. Rosenband and D. R. Leibrandt, Exponential scaling of clock stability with atom number (2013), arXiv:1303.6357 [quant-ph].

[105] K. Chabuda, I. D. Leroux, and R. Demkowicz-Dobrzański, The quantum allan variance, New Journal of Physics **18**, 083035 (2016).

[106] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański, Tensor-network approach for

quantum metrology in many-body quantum systems, Nature Communications **11**, 250 (2020).

[107] C. D. Marciniak, T. Feldker, I. Pogorelov, R. Kaubruegger, D. V. Vasilyev, R. van Bijnen, P. Schindler, P. Zoller, R. Blatt, and T. Monz, Optimal metrology with variational quantum circuits on trapped ions (2021), arXiv:2107.01860 [quant-ph].

[108] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. F. C. Vessot, and G. M. R. Winkler, Characterization of frequency stability, IEEE Transactions on Instrumentation and Measurement **IM-20**, 105 (1971).

[109] J. P. Dowling, G. S. Agarwal, and W. P. Schleich, Wigner distribution of a general angular-momentum state: Applications to a collection of two-level atoms, Phys. Rev. A **49**, 4101 (1994).

[110] G. J. Dick, Local oscillator induced instabilities in trapped ion frequency standards, in *Proceedings of the 19th Annu. Precise Time and Time Inverval Meeting, Redendo Beach, 1987* (U.S. Naval Observatory, 1888) pp. 133–147, http://tycho.usno.navy.mil/ptti/1987/Vol%2019_13.pdf.

[111] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, Variational-state quantum metrology, New Journal of Physics **22**, 083038 (2020).

[112] J. L. Beckey, M. Cerezo, A. Sone, and P. J. Coles, Variational quantum algorithm for estimating the quantum fisher information (2020), arXiv:2010.10488 [quant-ph].

[113] Z. Ma, P. Gokhale, T.-X. Zheng, S. Zhou, X. Yu, L. Jiang, P. Maurer, and F. T. Chong, Adaptive circuit learning for quantum metrology (2020), arXiv:2010.08702 [quant-ph].

[114] M. Mullan and E. Knill, Optimizing passive quantum clocks, Phys. Rev. A **90**, 042310 (2014).

[115] D. B. Hume and D. R. Leibrandt, Probing beyond the laser coherence time in optical clock comparisons, Phys. Rev. A **93**, 032138 (2016).

[116] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general markovian noise, Phys. Rev. X **7**, 041009 (2017).

[117] L. Pezzè and A. Smerzi, Heisenberg-limited noisy atomic clock using a hybrid coherent and squeezed state protocol, Phys. Rev. Lett. **125**, 210503 (2020).

[118] L. Pezzè and A. Smerzi, Quantum phase estimation algorithm with gaussian spin states, PRX Quantum **2**, 040301 (2021).