



PAPER

Noisy atomic magnetometry in real time

OPEN ACCESS

RECEIVED
27 May 2021REVISED
14 November 2021ACCEPTED FOR PUBLICATION
19 November 2021PUBLISHED
17 December 2021Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).Any further distribution
of this work must
maintain attribution to
the author(s) and the
title of the work, journal
citation and DOI.

Júlia Amorós-Binefa* and Jan Kołodnyński*

Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland
* Authors to whom any correspondence should be addressed.E-mail: j.amoros-binefa@cent.uw.edu.pl and j.kolodynski@cent.uw.edu.pl**Keywords:** quantum metrology, atomic magnetometry, continuous quantum measurement, spin-squeezing, Kalman filtering

Abstract

Continuously monitored atomic spin-ensembles allow, in principle, for real-time sensing of external magnetic fields beyond classical limits. Within the linear-Gaussian regime, thanks to the phenomenon of measurement-induced spin-squeezing, they attain a quantum-enhanced scaling of sensitivity both as a function of time, t , and the number of atoms involved, N . In our work, we rigorously study how such conclusions based on Kalman filtering methods change when inevitable imperfections are taken into account: in the form of collective noise, as well as stochastic fluctuations of the field in time. We prove that even an infinitesimal amount of noise disallows the error to be arbitrarily diminished by simply increasing N , and forces it to eventually follow a classical-like behaviour in t . However, we also demonstrate that, ‘thanks’ to the presence of noise, in most regimes the model based on a homodyne-like continuous measurement actually achieves the ultimate sensitivity allowed by the decoherence, yielding then the optimal quantum-enhancement. We are able to do so by constructing a noise-induced lower bound on the error that stems from a general method of classically simulating a noisy quantum evolution, during which the stochastic parameter to be estimated—here, the magnetic field—is encoded. The method naturally extends to schemes beyond the linear-Gaussian regime, in particular, also to ones involving feedback or active control.

1. Introduction

Optical magnetometers based on atomic spin-ensembles [1] are considered today as state-of-the-art magnetic-field sensors competing head to head in sensitivity with SQUID-based devices [2] without need of cryogenic cooling, while being already miniaturised to chip scales [3]. On one hand, they have been demonstrated to be capable of revolutionising medical applications [4–7], on the other, when interconnected into global networks, they are used in searches of new exotic physics [8, 9].

Despite constituting a prominent example of quantum sensors [10], the nominal sensitivity of atomic magnetometers is commonly described within their ‘classical’ regime of operation as follows [1]—also referred to as the ‘Equation One’ [11]:

$$\Delta^2 B \approx \frac{1}{(\gamma T_{\text{coh}})^2} \cdot \frac{T_{\text{coh}}}{T} \cdot \frac{1}{N} \quad (1)$$

where the (squared) error in sensing the value of the magnetic field, B , simply decreases (quadratically) with respect to the time over which the sensor gathers information about the field while undergoing Larmor precession—with the constant of proportionality given then by the effective gyromagnetic ratio γ . The maximal value of such probing time, T_{coh} , is dictated by the timescale within which the sensor can preserve its coherence, with T_{coh} being determined by dominant noise-mechanisms exhibited by the atomic ensemble, e.g. spin-relaxation $T_{\text{coh}} = T_1^*$ and/or spin-decoherence $T_{\text{coh}} = T_2^*$.¹ Moreover, as ‘Equation One’

¹ The *-notation is typically used to emphasise that these phenomenological quantities refer to the whole ensemble rather than each individual atom [12].

applies (only) in the limit when a sufficient number of measurements are performed over the total sensing time T , it contains a division over the number of repetitions, T/T_{coh} , in accordance with the central limit theorem. For the same reason, $\Delta^2 B$ is further divided by the number of atoms N , as in the ‘classical’ regime of operation the atoms within the ensemble can be treated as independent probes.

Although the derivation of ‘Equation One’ may be considered ‘hand-waving’, it can be formalised by following the so-called *frequentist approach* to estimation theory [13], which, indeed, applies in the limit of *large number of repetitions* (asymptotic statistics), here $T \gg T_{\text{coh}}$. In particular, the lhs of equation (1) formally corresponds to the *mean squared error* (MSE) of an (unbiased) estimator of B built basing on all the collected measurement data [10], while its minimum can then be generally associated with the *Cramér–Rao bound* [13], which effectively represents the rhs of equation (1). As a result, one can then answer fundamental questions using techniques of quantum metrology [14, 15], in particular, by how much can the sensitivity (1) be improved by allowing for arbitrary quantum states of the atomic ensemble and measurements more general than the natural light-probing scheme based on the Faraday effect [16]. This has lead to the seminal observation that ‘Equation One’ can be breached—in particular, its $1/N$ -behaviour commonly referred to as the *standard quantum limit* (SQL)—by preparing the atomic ensemble in an entangled state [17], so that the MSE can in principle attain the ultimate *Heisenberg limit* (HL) $\sim 1/N^2$ [18].

Unfortunately, such claims about the attainable scaling of precision with N have been proven to be overoptimistic in the asymptotic N limit, whenever one accounts for decoherence—noise—whose strength does not effectively decrease with the number of atoms² [15, 20]. In particular, the noise when affecting each atom in an uncorrelated manner constrains the quantum improvement to a constant factor beyond the SQL [21, 22], while when disturbing the whole ensemble in a collective way enforces a positive lower bound on $\Delta^2 B$ that cannot be overcome also when letting $N \rightarrow \infty$ [15, 23, 24]. Nonetheless, for finite but very large atomic ensembles ($N \gtrsim 10^5$), multiple optical-magnetometry experiments have spectacularly demonstrated that sensitivities beyond SQL can be reached [25–27]. In such experiments, the spin-ensemble is prepared in an (entangled) *spin-squeezed state* [28] every time before using it to sense the external magnetic field, B , which importantly cannot vary over the process of performing the necessary number of experimental repetitions.

The above paradigm, however, must be abandoned if the sensing task considered requires tracking of the magnetic field in *real time*, and one cannot assure the same magnitude of the field to be probed sufficiently many times. This may occur whenever the field follows a predetermined time-varying waveform subject to stochastic fluctuations, or its behaviour in time is simply not known at all [29]. Still, the *continuous quantum measurement theory* [30] allows one, in principle, to describe then the dynamics of an optical magnetometer conditioned on the data collected in real time [31, 32]—also beyond the ‘classical’ regime—while the ability to perform quantum non-demolition measurements [33, 34] of the atomic ensemble continuously in time [35–37] appears to be natural. Indeed, first experiments in this direction have been conducted demonstrating the capability to preserve quantum enhancement when waveforms of a known shape are being probed [38], or a fluctuating signal is sensed by a magnetometer operating in the ‘classical’ regime [39]. Moreover, the *Bayesian approach* to estimation theory [13] provides analogous tools to constrain the attainable ‘single-shot’ precision—via the *Bayesian Cramér–Rao bound* (BCRB) [40] and its variations [41]—that have been also generalised to the quantum setting [42].

In case of atomic magnetometry schemes in which spin-squeezing is realised continuously in time by light-probing based on the Faraday effect [43], the Bayesian inference techniques strikingly predict the *average mean squared error* (aMSE)³ to follow at short timescales the HL in absence of decoherence [44, 45]:

$$\Delta^2 \tilde{B} \propto \frac{1}{\gamma^2} \cdot \frac{1}{t^3} \cdot \frac{1}{N^2}, \quad (2)$$

where in comparison to the (classical) ‘Equation One’ (1) also the scaling in time is improved from $1/T$ to $1/t^3$ —we use above small t instead to emphasise that it now represents the time along a single experimental trajectory. Equation (2) has been subsequently adapted to account for stochastic fluctuations of the field [46, 47]. However, it has never been rigorously generalised to take into account the impact of decoherence beyond numerical considerations [45] and models that do not incorporate measurement back-action in real time [48], also when allowing for arbitrary quantum measurements to be performed at the end of the sensing procedure [49, 50].

In our work we tackle this problem by including the collective noise—general anisotropic decoherence affecting the atomic spin-ensemble as a whole—into the analysis, while similarly considering the short

² For complementary considerations of N -dependent decoherence models see e.g. [19].

³ Averaged in the Bayesian single-shot scenario over the prior knowledge about the B -field dynamics.

timescales, i.e. the *linear-Gaussian regime*, for which equation (2) was derived [44, 45]. In particular, we focus on the canonical optical magnetometry setup, in which all the atoms are initially polarised along one direction before being continuously spin-squeezed in time via the mechanism of light-probing in a perpendicular direction [31, 32, 44]. We explicitly account for the presence of collective noise, as well as stochastic fluctuations of the field, and construct the optimal Bayesian estimator, i.e. the *Kalman filter* (KF) [51, 52], whose minimal error we are able to resolve in time. Furthermore, we demonstrate how to generalise and adapt the theoretical techniques previously developed to deal with noise within the frequentist approach to quantum metrology [22, 53], in particular, the *classical simulation* (CS) method [54], which allows us to establish fundamental upper bounds on the attainable precision dictated by the decoherence. As a result, we are able to prove that, although at short timescales, when the collective decoherence can be effectively ignored, the error $\Delta^2 \hat{B}$ follows the $1/t^3$ -behaviour of equation (2), it is subsequently degraded to $1/t$ once the impact of decoherence ‘kicks-in’. Moreover, at large times at which the KF attains its optimal performance in tracking the fluctuating field—its *steady state* (SS)—the error reaches the minimal value that is determined by both the decoherence rate and the strength of field fluctuations. Focussing instead on the dependence of error on the ensemble size, we show that the Heisenberg-like behaviour $1/N^2$ is similarly lost as N grows and the impact of collective noise becomes significant. In particular, the error always approaches a constant value dictated by the decoherence rate. However, our method allows us to crucially demonstrate that the precision achieved by the magnetometry setup here considered saturates the ultimate bound for long times and large ensembles, and hence should be considered as the *optimal* realistic scheme to track fluctuating magnetic fields in presence of collective noise.

The manuscript is organised as follows. In section 2 we describe the atomic magnetometer of interest, in particular, the corresponding experimental setup as well as the resulting dynamical model of the spin-ensemble being probed continuously in time by light. This allows us to explain in detail in subsection 2.3 the linear-Gaussian regime being considered, and discuss the impact of noise on the evolution of the spin-squeezing parameter in subsection 2.4. The section 3 is then devoted to the explanation and derivation of the KF as the optimal estimator, as well as the precision it achieves in real time. In section 4 we present how the CS method can be utilised to derive the ultimate limit on precision dictated by the decoherence, stemming from the BCRB. Consequently, in subsection 4.3 we discuss the different regimes the error follows both in time and N , and give an intuitive explanation of their origin. Moreover, we demonstrate that the continuous estimation scheme saturates the ultimate limit derived in section 4 and, hence, may be regarded as optimal. Finally, we conclude our findings in section 5.

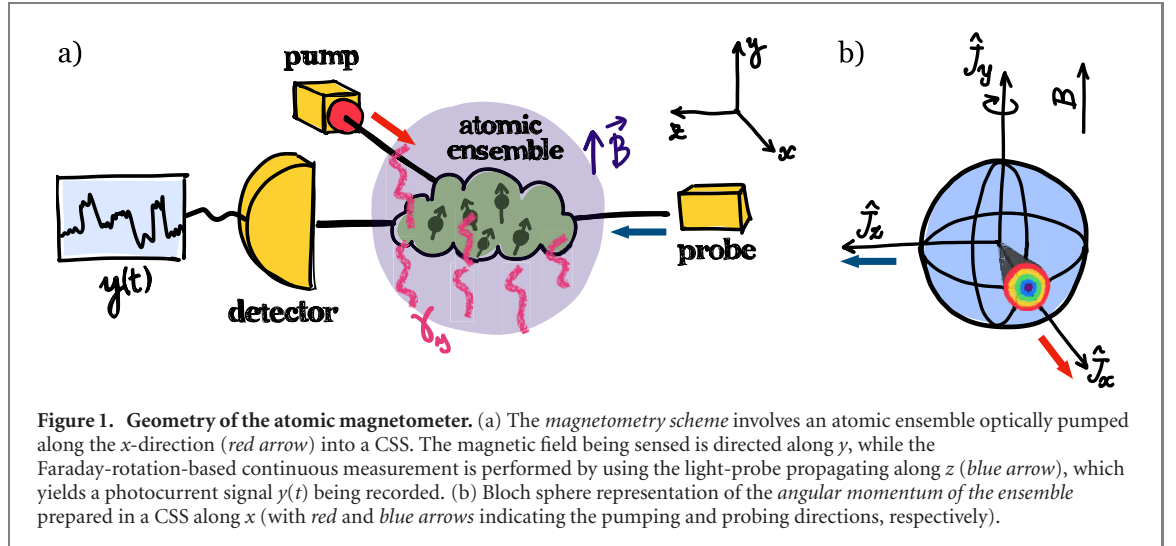
2. Atomic magnetometer model

2.1. Setup

The optical magnetometry scheme we consider, see figure 1(a), consists of an ensemble of N spin-1/2 atoms prepared in a *coherent spin state* (CSS) polarised along the x -direction [28], so that the initial mean and variance of the ensemble angular momentum operators, denoted by the vector $\hat{\mathbf{J}}(t) = (\hat{J}_x(t), \hat{J}_y(t), \hat{J}_z(t))^T$ in the Heisenberg picture, read $\langle \hat{\mathbf{J}}(0) \rangle_{\text{CSS}} = (J, 0, 0)^T$ and $\langle \Delta^2 \hat{\mathbf{J}}(0) \rangle_{\text{CSS}} = (0, J/2, J/2)^T$, respectively, where $J = N/2$. As shown in figure 1(b), one may then naturally visualise the distribution of the angular momentum with help of the Bloch sphere representation. The aim of the scheme is to measure and estimate in real-time the magnetic field B_t being directed along y , which fluctuates according to the *Ornstein–Uhlenbeck* (OU) stochastic process [55]:

$$dB_t = -\chi B_t dt + dW_B, \quad (3)$$

where by dW_B we denote the Wiener noise of zero mean, $\mathbb{E}[dW_B] = 0$, variance $\mathbb{E}[dW_B^2] = q_B dt$. The magnetic field B_t induces a Larmor precession around the y -axis at the frequency $\omega_L(t) = \gamma B_t$, with γ being the effective (constant) gyromagnetic ratio. However, as discussed in more detail below, in our analysis we restrict to dynamics at small times in the presence of small magnetic fields, under which the angular momentum operator $\hat{\mathbf{J}}(t)$ deviates only slightly from pointing along the x -direction—the direction along which the atoms are initially pumped. In principle, this regime could also be achieved over larger timescales in experiments involving stroboscopic probing [56–59]. However, such implementations would then require the atoms to be initially pumped in the direction perpendicular to the magnetic field, as depicted in figure 1.



2.2. System and measurement dynamics

The atomic ensemble is continuously monitored by exploiting the paramagnetic Faraday rotation effect [16], which twists the linear polarisation of the (off-resonant) light propagating through the ensemble in the z -direction—the light probe, see figure 1(a). As a result, a quantum non-demolition measurement is realised [33, 34] that is undertaken continuously in time [35–37], which can be effectively described by means of the homodyne-like continuous measurement [60] with the shot-noise in the measured signal taking the form of a Wiener process [43, 61]. Note that, even if higher-spin atoms are considered, the following measurement model still applies as long as the contribution from the atomic polarisability tensor component can be suppressed [43, 62, 63]. In particular, the photocurrent being measured at time t is proportional to the mean value of the collective atomic spin-component along the probe, $\langle \hat{J}_z(t) \rangle_{(c)}$, i.e.:

$$y(t)dt = 2\eta\sqrt{M}\langle \hat{J}_z(t) \rangle_{(c)} dt + \sqrt{\eta}dW, \quad (4)$$

where η is the detection efficiency, M is the measurement strength, and dW is the Wiener differential fulfilling $\mathbb{E}[dW^2] = dt$ according to the Itô lemma [55].

Importantly, due to the active influence the continuous measurement (4) exerts on the atoms, the ensemble dynamics becomes conditional—denoted by the subscript (c)—and the atoms evolve differently depending on a particular trajectory of the measurement outcomes registered up to (and including) time t , $y_{\leq t} = \{y(\tau)\}_{0 \leq \tau \leq t}$ [61]. In particular, the ensemble is characterized at time t by the quantum state conditioned on the past measurement record, $\rho_{(c)}(t) \equiv \rho(t|y_{\leq t})$, which undergoes further *conditional dynamics* described by a stochastic master equation:

$$d\rho_{(c)}(t) = -i\gamma B_t[\hat{J}_y, \rho_{(c)}(t)] dt + \sum_{\alpha=x,y,z} \gamma_\alpha \mathcal{D}[\hat{J}_\alpha]\rho_{(c)}(t) dt \quad (5)$$

$$+ M\mathcal{D}[\hat{J}_z]\rho_{(c)}(t) dt + \sqrt{M\eta}\mathcal{H}[\hat{J}_z]\rho_{(c)}(t) dW, \quad (6)$$

where the superoperators \mathcal{D} and \mathcal{H} are defined as $\mathcal{D}[\mathcal{O}]\rho = \mathcal{O}\rho\mathcal{O}^\dagger - \frac{1}{2}(\mathcal{O}^\dagger\mathcal{O}\rho + \rho\mathcal{O}^\dagger\mathcal{O})$ and $\mathcal{H}[\mathcal{O}]\rho = \mathcal{O}\rho + \rho\mathcal{O}^\dagger - \text{Tr}[(\mathcal{O} + \mathcal{O}^\dagger)\rho]\rho$ for any (also non-Hermitian) operator \mathcal{O} [30].

The first term in (5) arises simply from the free Hamiltonian, $\hat{H} = \gamma B_t \hat{J}_y$, responsible for the Larmor precession of the ensemble spin. In contrast, both terms in (6) describe the continuous measurement of \hat{J}_z introduced above [31, 32], of which the former is responsible for measurement-induced decoherence, or the backaction, while the latter constitutes the non-linear information gain provided when observing a particular measurement signal in (4) [30]. Furthermore, in our analysis we model all possible, e.g. environment-induced, decoherence mechanisms affecting the ensemble (as a whole) by introducing the second term in (5) with three dissipative components being parametrised by distinct effective rates, γ_α , in the three directions $\alpha \in \{x, y, z\}$. Let us also emphasise that the Wiener differential, dW , appearing in (6) is the same as in the measurement dynamics (4)—a particular fluctuation of the photocurrent in (4) after being registered drives the conditional state of the ensemble according to (6).

Finally, let us highlight the difference between the fluctuations of the estimated magnetic field B_t , as dictated by the OU process (3), and the collective decoherence introduced above in (5), of which *only* the latter should be interpreted as a form of ‘noise’—in the statistical sense, forcing the quantum state of

ensemble to lose its purity over time. In particular, due to its presence, the expectation value of \hat{J}_α ($\alpha = x, y, z$) exponentially decreases with a rate of γ_α , whose inverse can be identified as the phenomenological (ensemble) spin-decoherence time T_2^* [12]. Within the theory of open quantum systems [64], this may be associated with tracing out some environmental degrees of freedom or the environment itself monitoring continuously each component \hat{J}_α , whose stochastic trajectory of outcomes is inaccessible and, hence, must be averaged out. On the contrary, the above dynamical model (3)–(6) describes the evolution along a *single* trajectory of the fluctuating magnetic field. As a result, the impact of field fluctuations on the performance in magnetometry should not be treated on equal grounds with the decoherence of the atomic ensemble, but rather associated with the inability to perfectly estimate the field in real time due to its stochastic nature. In fact, if one on purpose decided to ignore the field fluctuations and rather consider the effective dynamics averaged over all possible field trajectories (see appendix F), one would recover the above model of collective decoherence in the direction of the field (here, \hat{J}_y), but with a decoherence rate that accumulates as $\gamma_y \sim t^2$ over time.

2.3. Linear-Gaussian regime

In order to define the regime of interest in which the atomic sensor operates, let us first consider the *unconditional dynamics* of the atomic ensemble,

$$d\rho(t) = -i\gamma B_t[\hat{J}_y, \rho(t)] dt + \sum_{\alpha=x,y,z} \gamma_\alpha \mathcal{D}[\hat{J}_\alpha]\rho(t) dt + M\mathcal{D}[\hat{J}_z]\rho(t) dt, \quad (7)$$

which can be simply obtained by dropping the stochastic term in (5) and (6), so that it now describes the evolution of the atomic state at time t , averaged over all the stochastic trajectories of the measured outcomes, i.e. [55]:

$$\rho(t) = \mathbb{E}_{p(\mathbf{y}_{\leq t})}[\rho_{(c)}(t)] = \int d\mathbf{y}_{\leq t} p(\mathbf{y}_{\leq t}) \rho(t|\mathbf{y}_{\leq t}), \quad (8)$$

where by $\int d\mathbf{y}_{\leq t}$ we denote the integral over all possible measurement records $\mathbf{y}_{\leq t} = \{y(\tau)\}_{0 \leq \tau \leq t}$. Using equation (7) to define the dynamics of any observable $\hat{O}(t)$ in the Heisenberg picture, whose mean must then obey $\langle \hat{O}(t) \rangle = \text{Tr}\{\hat{O}(t)\rho\} = \text{Tr}\{\hat{O}\rho(t)\}$, we observe that the (unconditional) evolution of the mean spin-component along the direction of pumping, $\langle \hat{J}_x(t) \rangle$, is governed by the solution to the following set of coupled differential equations:

$$d\langle \hat{J}_x(t) \rangle = \gamma B_t \langle \hat{J}_z(t) \rangle dt - \frac{1}{2}(M + \gamma_y + \gamma_z) \langle \hat{J}_x(t) \rangle dt, \quad (9)$$

$$d\langle \hat{J}_z(t) \rangle = -\gamma B_t \langle \hat{J}_x(t) \rangle dt - \frac{1}{2}(\gamma_x + \gamma_y) \langle \hat{J}_z(t) \rangle dt, \quad (10)$$

$$dB_t = -\chi B_t dt + dW_B, \quad (11)$$

of which the last (stochastic) one just corresponds to the OU process with $\mathbb{E}[dW_B^2] = q_B$, describing the magnetic-field fluctuations in (3).

Although this implies that the dynamics of $\langle \hat{J}_x(t) \rangle$ is stochastic, we show in appendix A that, if one focuses on short timescales such that $\overline{\omega_L(t)} t \ll 1$, where $\overline{\omega_L(t)} = \gamma |\overline{B_t}|$ is the time-averaged Larmor frequency, then the mean spin-component along the direction of the pump must decay exponentially as follows:

$$\langle \hat{J}_x(t) \rangle \approx J e^{-(M+\gamma_y+\gamma_z)t/2} = J e^{-r t/2} \quad (12)$$

with an effective *decay rate*: $r = M + \gamma_y + \gamma_z$. For this to be true, not only the time-average value of the magnetic field, $\overline{B_t} = \frac{1}{t} \int_0^t d\tau B_\tau$, must be small ($|\overline{B_t}| \ll \frac{1}{\gamma_t}$), but also the distribution of $\overline{B_t}$ should be narrow enough, which we ensure by verifying that $|\mathbb{E}[\overline{B_t}] \pm 2\sqrt{\text{Var}[\overline{B_t}]}| \lesssim \frac{1}{\gamma_t}$ according to the 68–95–99.7 rule for normal distributions (see also figure 3(a)). We explicitly prove in appendix A that for approximation (12) to hold it is sufficient to consider short timescales such that $rt \lesssim 1$, while also requiring the parameters of the OU process in (11) (or (3)) to obey:

$$\chi \ll \frac{4r}{3} \quad \text{and} \quad q_B \lesssim \frac{3r^3}{4\gamma^2}. \quad (13)$$

Furthermore, by restricting to short timescales at which the approximation (12) is valid, we effectively deal with spin dynamics in the regime in which $\hat{J}(t)$ only slightly deviates from its initial x-polarisation.

This allows us then to perform the *Gaussian approximation* [45, 65] and introduce the effective (time-dependent) quadrature operators:

$$\hat{X}(t) = \frac{\hat{J}_y}{\sqrt{|\langle \hat{J}_x(t) \rangle|}} \approx \frac{\hat{J}_y}{\sqrt{J e^{-rt/2}}} \quad \text{and} \quad \hat{P}(t) = \frac{\hat{J}_z}{\sqrt{|\langle \hat{J}_x(t) \rangle|}} \approx \frac{\hat{J}_z}{\sqrt{J e^{-rt/2}}}, \quad (14)$$

which satisfy the canonical commutation relationship $[\hat{X}(t), \hat{P}(t)] = i \frac{\hat{J}_x}{|\langle \hat{J}_x(t) \rangle|} \approx i$, as long as $|\langle \hat{J}_x(t) \rangle| \gg 1$ [45, 65], which is assured within the approximation (12) whenever $J \gg 1$ —recall $J = N/2 \gtrsim 10^5$ in optical-magnetometry experiments [25–27]. As a consequence, it is enough to consider the evolution of the first and second moments of the quadratures, so that when focussing on the probed spin-component in the z -direction that fulfils $\hat{J}_z \approx \sqrt{J e^{-rt/2}} \hat{P}(t)$, its conditional dynamics (5) and (6) is completely described by the following set of equations:

$$y(t)dt = 2\eta\sqrt{M}\langle \hat{J}_z(t) \rangle_{(c)}dt + \sqrt{\eta}dW, \quad (15)$$

$$dB_t = -\chi B_t dt + dW_B \quad (\text{with } \mathbb{E}[dW_B^2] = q_B) \quad (16)$$

$$d\langle \hat{J}_z(t) \rangle_{(c)} = -\gamma B_t J e^{-rt/2} dt + 2\sqrt{\eta M} \langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} dW, \quad (17)$$

$$d\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} = -4M\eta \langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}^2 dt + \gamma_y J^2 e^{-rt} dt, \quad (18)$$

which assume $[\hat{X}(t), \hat{P}(t)] \approx i$ to hold—satisfied whenever $J \gg 1$, $rt \lesssim 1$ and the conditions (13) are valid. Note that equations (15)–(17) are linear with respect to one another, while equation (18) can be solved independently. Moreover, they involve only Gaussian (in particular, Wiener) stochastic-noise terms, and are derived based on the Gaussian approximation (14). Hence, we refer in short to the evolution described by (15)–(18) as the (conditional) dynamics of \hat{J}_z in the *linear-Gaussian regime*.

2.4. Continuous spin-squeezing

Firstly, let us note that within the linear-Gaussian regime, the equations of motion (15)–(17) are independent of the decoherence rate γ_x , which is thus redundant. This is a consequence of the approximation (12) (see also appendix A) and can be intuitively explained—deviations of $\hat{J}(t)$ from the x -direction are then too small for the collective noise generated by the $\mathcal{D}[\hat{J}_x]$ -term in equation (5) to have any effect on the quadratures (14). Furthermore, the decoherence rate γ_z is also unnecessary, since by transforming the parameters of the continuous measurement (4) as follows: $M \rightarrow M - \gamma_z$, $\eta \rightarrow \eta M / (M - \gamma_z)$ and $y \rightarrow y \sqrt{M / (M - \gamma_z)}$; we retrieve the conditional dynamics (15)–(18) with $\gamma_z = 0$. Hence, the impact of the collective noise generated by the $\mathcal{D}[\hat{J}_z]$ -term in equation (5) instead, can always be interpreted and incorporated into a modified form of the continuous measurement (4).

As a result, without loss of generality, we restrict the effective decay rate of $\langle \hat{J}_z(t) \rangle$ introduced in (12) to take the form $r = M + \gamma_y$, when considering the (conditional) dynamics of \hat{J}_z in (15)–(17). We then solve for the dynamics of $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ in (18), which constitutes a fully decoupled differential equation. For the complete analytical solution we refer the reader to appendix B, where we also show that the time-dependence of the (conditional) \hat{J}_z -variance (satisfying $\langle \Delta^2 \hat{J}_z(0) \rangle_{(c)} = \langle \Delta^2 \hat{J}_z(0) \rangle_{\text{CSS}} = J/2$ at $t = 0$) can be described by two distinct short-time and long-time behaviours:

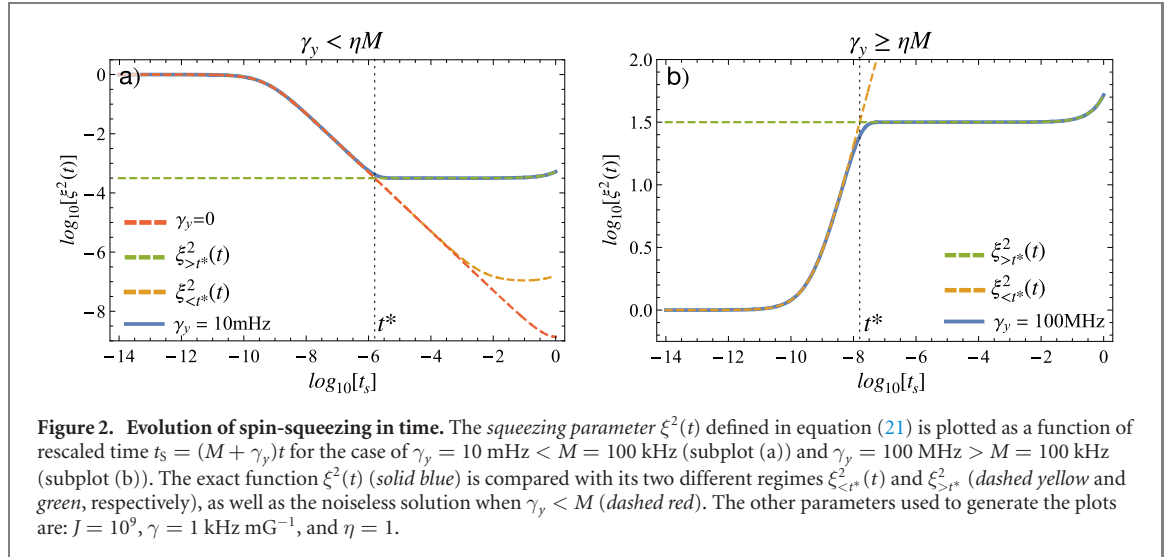
$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} = \begin{cases} V_{<t^*}(t) = \frac{J}{2} \frac{1 + 2Jt\gamma_y}{1 + 2JtM\eta} e^{-(M+\gamma_y)t/2}, & \text{if } 0 \leq t \ll t^* \quad (19a) \\ V_{>t^*}(t) = \frac{J}{2} \sqrt{\frac{\gamma_y}{\eta M}} e^{-(M+\gamma_y)t/2}, & \text{if } t \gg t^* \quad (19b) \end{cases}$$

where $t^* = (2J\sqrt{M\gamma_y\eta})^{-1}$ is the transition time between the two regimes.

Importantly, note that $t \ll t^*$ implies $2Jt\gamma_y \ll \sqrt{\gamma_y/(\eta M)}$. Then, if also $\gamma_y < \eta M$, we may infer $2Jt\gamma_y \ll 1$ and approximate $1 + 2Jt\gamma_y \approx 1$ in (19a). As a result, we then recover the noiseless ($\gamma_y = 0$) solution for the variance within the short-time regime, previously found by Geremia *et al* [44], i.e.:

$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} = \frac{J}{2} \frac{1 + 2Jt\gamma_y}{1 + 2JtM\eta} e^{-(M+\gamma_y)t/2} \approx \frac{J}{2 + 4JtM\eta}, \quad (20)$$

despite the actual presence of the noise ($\gamma_y > 0$). In fact, we prove also in appendix B that $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ is a non-decreasing function at $t \approx 0$ if $\gamma_y \geq \eta M$. Hence, the noise may be considered negligible at small times *only if* $\gamma_y < \eta M$.



The dynamics of $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ enables us to study the phenomenon of spin-squeezing of the atomic ensemble [28] by directly computing the (conditional) squeezing parameter introduced by Wineland *et al* [66] along the pumping x -direction:

$$\xi^2(t) = \frac{\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}}{\langle \hat{J}_x(t) \rangle^2} \left[\frac{\langle \Delta^2 \hat{J}_z(0) \rangle_{\text{CSS}}}{\langle \hat{J}_x(0) \rangle_{\text{CSS}}^2} \right]^{-1} = \frac{2J \langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}}{\langle \hat{J}_x(t) \rangle^2} \quad (21)$$

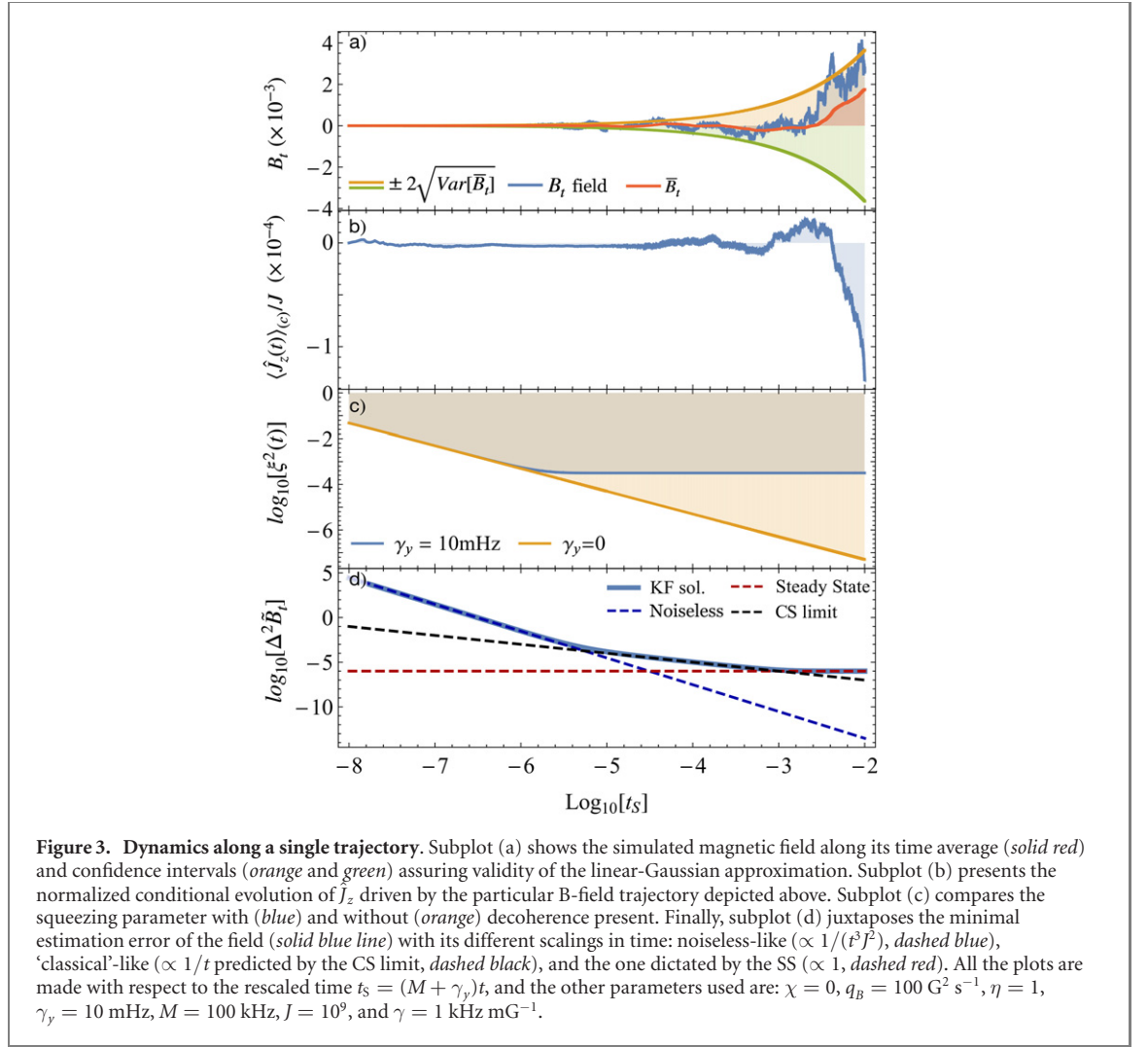
which when satisfying $\xi^2(t) < 1$ indicates a gain in interferometric precision over the CSS [66]. In particular, just like we decomposed $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ according to (19), we can similarly split the behaviour of the squeezing parameter in the linear-Gaussian regime, in which $\langle \hat{J}_x(t) \rangle \approx J e^{-r t/2}$, as follows

$$\xi^2(t) \approx \frac{2 \langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}}{J e^{-(M+\gamma_y)t}} = \begin{cases} \xi^2_{<t^*}(t) = \frac{1 + 2Jt\gamma_y}{1 + 2JtM\eta} e^{(M+\gamma_y)t/2}, & \text{if } 0 \leq t \ll t^*, \quad (22a) \\ \xi^2_{>t^*}(t) = \sqrt{\frac{\gamma_y}{\eta M}} e^{(M+\gamma_y)t/2}, & \text{if } t \gg t^*. \quad (22b) \end{cases}$$

From now on in all plots, in order to confirm that apart from (13) the condition $rt \lesssim 1$ (with now $r = M + \gamma_y$) is satisfied and, hence, the linear-Gaussian approximation (12) is valid, we introduce the rescaled time $t_s = (M + \gamma_y)t$ and limit its range to $t_s \lesssim 1$, what assures the timescale of the linear-Gaussian regime to always be consistently considered.

In figure 2 we present explicitly the exact dynamics of the squeezing parameter (21) for two important cases: (a) when $\gamma_y < \eta M$ and the spin-squeezing ($\xi^2(t) < 1$) occurs within a finite-time window (see figure 2(a)); and (b) when $\gamma_y \geq \eta M$ for which spin-squeezing is forbidden (see figure 2(b)). In both cases, it is evident that the exact solution for $\xi^2(t)$ very closely follows the two-regime behaviour in (22), with a clear transition at $t \approx t^*$. Moreover, as seen explicitly from the two-regime solution (22)—and the exact solution that can be straightforwardly derived from the exact evolution of the \hat{J}_z -variance (19) described in appendix B—the dynamics of the squeezing parameter is specified solely by the properties of the continuous measurement (η and M), collective decoherence (γ_y), and the number of atoms ($J = N/2$).

In what follows, we turn to the problem of sensing the magnetic field, B_t , in real time by constructing the optimal estimator, \tilde{B}_t , of its fluctuating value in the form of a KF [51, 52]. However, we have already included in figure 3 the evolution of the corresponding minimal estimation error, $\Delta^2 \tilde{B}_t$, in order to emphasise the fact that the stochasticity of the B_t -signal affects the estimation procedure, but *not* the spin-squeezing phenomenon. In particular, for a given magnetic-field trajectory generated according to the OU process (3)—of a magnitude that does not invalidate the linear-Gaussian regime, see figure 3(a)—we obtain a measurement signal $y(t)$ from which we can infer via (4) the conditional dynamics of the mean spin-component $\langle \hat{J}_z(t) \rangle_{(c)}$, see figure 3(b). Now, the evolution of its variance, $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$, allows us to compute the squeezing parameter $\xi^2(t)$, which, as described above, evolves in the linear-Gaussian regime in the same manner for any particular stochastic trajectory of the magnetic field—see figure 3(c). In contrast, the stochastic fluctuations of B_t affect the estimation procedure, so that the KF reaches a constant error (the SS) after a certain time despite the atomic ensemble still being spin-squeezed ($\xi^2(t) < 1$). This can be



directly seen by comparing the evolution of $\Delta^2 \tilde{B}_t$ in figure 3(d) against the one of $\xi^2(t)$ in figure 3(c) for $t_s \gtrsim 10^{-3}$. The transition time when this occurs depends on the parameters of the OU process (3)—as later shown in section 4.2.

3. Field estimation with Kalman filtering

The goal within the magnetometry scheme considered here is to most accurately infer the true value of the magnetic field at a time t , given a particular measurement signal $y_{\leq t}$ recorded up to this time instance. In order to do so, we seek the optimal estimator $\tilde{B}_t(y_{\leq t})$ of B_t that minimises the *average (Bayesian) mean squared error* (aMSE) [13]:

$$\Delta^2 \tilde{B}_t = \mathbb{E}_{p(B_t, y_{\leq t})} [(B_t - \tilde{B}_t(y_{\leq t}))^2] = \int dB_t p(B_t) \mathbb{E}_{p(y_{\leq t}|B_t)} [(B_t - \tilde{B}_t(y_{\leq t}))^2], \quad (23)$$

where by $\int dB_t$ we denote the integral over the values that the magnetic field may take at time t (only), i.e. the random variable B_t with probability distribution:

$$p(B_t) = \int dB_0 p(B_t|B_0) p(B_0), \quad (24)$$

which for the OU process (3) can be determined once the *a priori* distribution $p(B_0)$ is specified, e.g. a normal distribution of variance σ_0^2 (see appendix E.1). From the Bayesian perspective, $p(B_0)$ should most accurately describe the knowledge an experimentalist possesses about the field at $t = 0$, prior to taking any measurements. Note that, as in our work we are interested in estimation of B in real time, we will not consider the setting in which one seeks the optimal estimator for some time $t' < t$ and accounts also for ‘future’ measurement-data collected in the time-window $[t', t]$. In that case, one should resort to the

inference methods of Bayesian smoothing [67–69], rather than to the filtering ones that are relevant for our purpose.

The optimal estimator minimising the aMSE is always given by the mean of the posterior distribution [40], i.e.:

$$\tilde{B}_t(\mathbf{y}_{\leq t}) = \mathbb{E}_{p(B_t|\mathbf{y}_{\leq t})} [B_t] = \int dB_t p(B_t|\mathbf{y}_{\leq t}) B_t = \mathbb{E}_{p(\mathbf{B}_{\leq t}|\mathbf{y}_{\leq t})} [B_t], \quad (25)$$

where in the last step we have used $p(B_t|\mathbf{y}_{\leq t}) = \int d\mathbf{B}_{<t} p(\mathbf{B}_{\leq t}|\mathbf{y}_{\leq t})$ —the fact that the probability of the magnetic field taking at time t the value B_t , given a particular measurement trajectory $\mathbf{y}_{\leq t}$, can be obtained by averaging over all past B -field stochastic trajectories up to (but not including) time t . However, since the problem of interest is described by a set of equations (15)–(17) that are linear and Gaussian, the posterior distribution $p(B_t|\mathbf{y}_{\leq t})$ does not have to be explicitly reconstructed. Instead, the optimal estimator (25) can be determined by solving the so-called Kalman–Bucy equation [40] and is commonly referred to as the KF [51, 52]. The name ‘filter’ originates from the fact that the optimal estimator (25) at time t can be constructed basing solely on the previous-step estimate and the current measurement outcome y_t , so there is, in principle, no need to store all the measurement data.

In order to formulate the problem within the KF-framework [70, 71], we first define the *state* vector of the variables $\langle \hat{J}_z(t) \rangle_{(c)}$ and B_t to be estimated⁴ $\mathbf{x}_t = (\langle \hat{J}_z(t) \rangle_{(c)}, B_t)^T$, as well as the *measurement* vector $\mathbf{y}_t = \int_0^t y(\tau) d\tau$, which takes, however, a scalar form with only one d.o.f. being continuously measured. Consequently, we then identify the corresponding noise vectors as $d\mathbf{w}_t = (dW, dW_B)^T$ and $d\mathbf{v}_t = \sqrt{\eta} dW$, respectively, so that the system of equations (15)–(17) can be compactly rewritten as:

$$d\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_t dt + \mathbf{B}_t d\mathbf{w}_t, \quad (26)$$

$$d\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t dt + d\mathbf{v}_t, \quad (27)$$

where

$$\mathbf{F}_t = \begin{pmatrix} 0 & -\gamma J e^{-rt/2} \\ 0 & -\chi \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} 2\sqrt{\eta M} \langle \Delta^2 J_z(t) \rangle_{(c)} & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{H}_t = 2\sqrt{\eta M} \begin{pmatrix} 1 & 0 \end{pmatrix}. \quad (28)$$

The self- and cross-correlations between stochastic noise-terms are then given by

$$\mathbb{E}[d\mathbf{w}_t d\mathbf{w}_s^T] = \mathbf{Q}_t \delta(t-s) dt, \quad (29)$$

$$\mathbb{E}[d\mathbf{v}_t d\mathbf{v}_s^T] = \mathbf{R}_t \delta(t-s) dt, \quad (30)$$

$$\mathbb{E}[d\mathbf{w}_t d\mathbf{v}_s^T] = \mathbf{S}_t \delta(t-s) dt, \quad (31)$$

where

$$\mathbf{Q}_t = \begin{pmatrix} 1 & 0 \\ 0 & q_B \end{pmatrix}, \quad \mathbf{R}_t = \eta \quad \text{and} \quad \mathbf{S}_t = \begin{pmatrix} \sqrt{\eta} \\ 0 \end{pmatrix}, \quad (32)$$

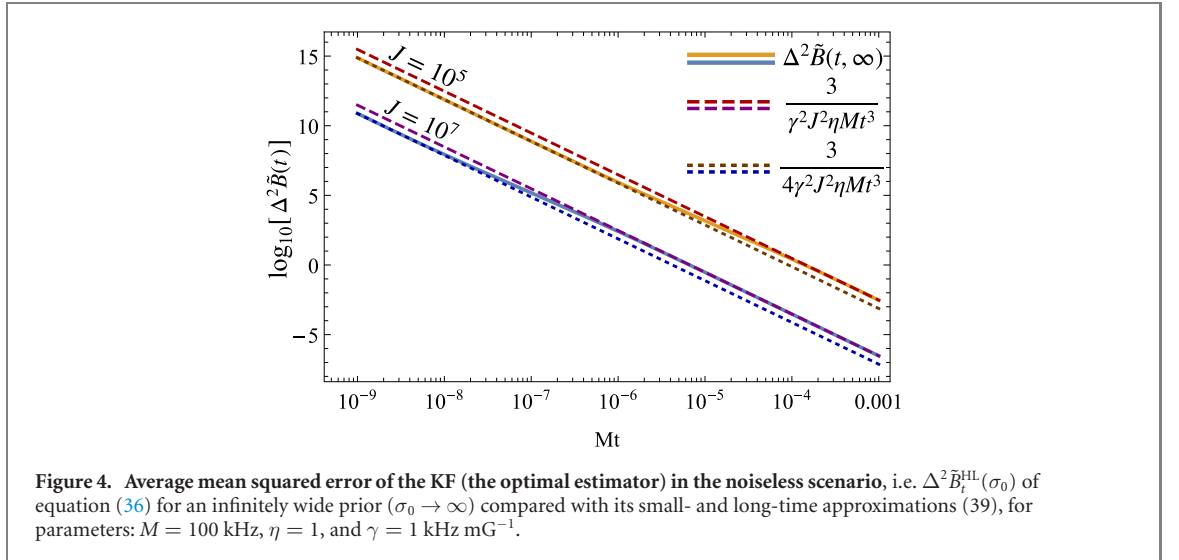
are the covariance and cross-correlation matrices (or scalars) of the process and measurement noises, respectively. Let us note that we have used $\mathbb{E}[dW_B dW] = \mathbb{E}[dW dW_B] = 0$ in (31), as no cross-correlations are present between the B -field fluctuations and the measurement noise; whereas, because $q_B \geq 0$ and $\eta \geq 0$, the covariances of process and measurement noises consistently satisfy $\mathbf{Q}_t, \mathbf{R}_t \geq 0$.

As discussed in more detail in appendix C, the optimal estimator $\tilde{\mathbf{x}}_t = (\langle \hat{J}_z(t) \rangle_{(c)}, \tilde{B}_t)^T$ that minimises the overall aMSE—formally corresponding to the trace of the covariance matrix $\Sigma_t = \mathbb{E}[(\mathbf{x}_t - \tilde{\mathbf{x}}_t)(\mathbf{x}_t - \tilde{\mathbf{x}}_t)^T]$, i.e. $\text{Tr}\{\Sigma_t\} = \mathbb{E}[\|\mathbf{x}_t - \tilde{\mathbf{x}}_t\|^2]$ —is given by the *correlated* KF [70]. In particular, focussing on the optimal covariance matrix that the filter yields, its dynamics is then determined by a non-linear differential equation of the Riccati form [70]:

$$\begin{aligned} \frac{d\Sigma_t}{dt} \equiv \dot{\Sigma}_t &= (\mathbf{F}_t - \mathbf{B}_t \mathbf{S}_t \mathbf{R}_t^{-1} \mathbf{H}_t) \Sigma_t + \Sigma_t (\mathbf{F}_t - \mathbf{B}_t \mathbf{S}_t \mathbf{R}_t^{-1} \mathbf{H}_t)^T \\ &\quad - \Sigma_t \mathbf{H}_t^T \mathbf{R}_t^{-1} \mathbf{H}_t \Sigma_t + \mathbf{B}_t (\mathbf{Q}_t - \mathbf{S}_t \mathbf{R}_t^{-1} \mathbf{S}_t^T) \mathbf{B}_t^T, \end{aligned} \quad (33)$$

where the initial condition reads $\Sigma_0 = \text{diag}(0, \sigma_0^2)$ with σ_0^2 being the variance of the Gaussian prior distribution $p(B_0)$ in equation (24). In order to allow for analytic solutions in some parameter regimes, we

⁴ The KF-framework then naturally provides also the optimal estimate $\langle \hat{J}_z(t) \rangle_{(c)}$ of the mean $\langle \hat{J}_z(t) \rangle_{(c)}$.



redefine the covariance matrix as $\Sigma_t = Y_t X_t^{-1}$, so that we may rewrite (33) as a system of coupled but linear differential equations,

$$\dot{X}_t = - (F_t - B_t S_t R_t^{-1} H_t)^T X_t + H_t^T R_t^{-1} H_t Y_t, \quad (34)$$

$$\dot{Y}_t = (F_t - B_t S_t R_t^{-1} H_t) Y_t + B_t (Q_t - S_t R_t^{-1} S_t^T) B_t^T X_t, \quad (35)$$

with the initial condition now corresponding to $X_0 = 1$ and $Y_0 = \Sigma_0$.

3.1. Solution in the absence of decoherence and field fluctuations

In the absence of collective noise and field fluctuations ($\gamma_y = 0$ and $q_B = \chi = 0$), the Riccati equation (33) can be explicitly solved [44]. Since $\gamma_y = 0$, the conditional dynamics of the \hat{J}_z -variance is then given by equation (20). Moreover, since no fluctuations of the field are being considered, the noise correlation matrix just reads $Q_t = \text{diag}(1, 0)$. As a result, the decoupled system of differential equations introduced in (34) and (35) can be analytically solved, providing the solution for the minimal aMSE in estimating the B -field (23) as

$$\Delta^2 \tilde{B}_t^{\text{HL}}(\sigma_0) = [\Sigma_t]_{22} = \frac{M^2}{16\eta\gamma^2 J^2} \frac{(1 + 2JMt\eta)}{a(t)e^{-Mt} + 4(1 + 4J\eta)e^{-Mt/2} + b(t)}, \quad (36)$$

where we have emphasised its dependence on the width of the prior, σ_0 , and defined:

$$a(t) = -(1 + 2\eta J(4 + Mt)), \quad (37)$$

$$b(t) = \frac{M^2}{16\eta\gamma^2 J^2 \sigma_0^2} + \frac{M^3 t}{8\gamma^2 J \sigma_0^2} + c(t). \quad (38)$$

with $c(t) = (Mt - 3) + 2\eta J(Mt - 4)$. Note that for an infinitely wide prior, $\sigma_0 \rightarrow \infty$, $b(t) = c(t)$ and the aMSE (36) matches consistently the solution obtained in [44] with $\Delta^2 \tilde{B}_t^{\text{HL}}(\infty)$ exhibiting the non-classical scaling in both t and J —following the HL $\sim 1/N^2$ with $N = J/2$ —as mentioned already in equation (2). This becomes clear after making the approximation summarised in figure 4, i.e.:

$$\Delta^2 \tilde{B}_t^{\text{HL}}(\infty) \approx \begin{cases} \frac{3}{\eta M t^3} \frac{1}{4\gamma^2 J^2}, & \text{for } t \ll (JM)^{-1}, \\ \frac{3}{\eta M t^3} \frac{1}{\gamma^2 J^2}, & \text{for } (JM)^{-1} \ll t \quad (< M^{-1}), \end{cases} \quad (39a)$$

$$(39b)$$

where the terms (39a) and (39b) are obtained by Taylor-expanding $\Delta^2 \tilde{B}_t^{\text{HL}}(\infty)$ in equation (36) to leading order in time t and $(JM)^{-1}$, respectively.

4. Impact of noise

In figure 3(d) we have already presented (solid blue line) the behaviour of the minimal aMSE, $\Delta^2 \tilde{B}_t$, in presence of noise ($\gamma_y > 0$), which we have determined by numerically solving the Riccati equation (33) after

assuming no initial knowledge about the field ($\sigma_0 \rightarrow \infty$), and setting $\gamma_y < \eta M$. The latter condition is necessary for the minimal aMSE to initially follow the noiseless solution introduced in (39) and, hence, exhibit a ‘supraclassical’ scaling in time at short timescales, $\sim 1/t^3$, as only then the spin-squeezing induced by the continuous measurement can occur—the conditional variance of \hat{J}_z decreases with time, see (20) and appendix B, so that $\xi^2(t) < 1$ in equation (22), as depicted in figure 2(a).

Importantly, although the spin-squeezing is maintained at larger timescales (see figure 3(c)), the noiseless behaviour of aMSE in figure 3(d) is quickly lost due to noise. In particular, the minimal aMSE firstly follows a classical-like scaling, $\sim 1/t$, before attaining a constant value at timescales $t_S \sim 1$, at which still $\xi^2(t) < 1$. In the upcoming sections we explicitly show how the presence of collective decoherence and field fluctuations causes the ‘supraclassical’ scaling to be lost both in time t and the atom number $N(J = N/2)$. We achieve this by giving analytical solutions to what in figure 3(d) we refer to as the CS *limit* and the SS.

4.1. No-go theorem for the Heisenberg limit: the classical simulation method

Within the Bayesian approach to estimation theory there exist various general lower bounds on the minimal aMSE, which are commonly referred to as *Bayesian* (or global) *Cramér–Rao Bounds* (BCRBs) [72]. Still, in the case of linear and Gaussian stochastic processes many of these simplify to a single form [40, 41]. When interested in estimating the process at its last step—here, the magnetic field B_t at the final time t —the applicable bound is the so-called *marginal unconditional* BCRB [41]:

$$\Delta^2 \tilde{B}_t \geq (J_B)^{-1}, \quad (40)$$

where J_B is the *Bayesian information* (BI) [40] evaluated at time t :

$$J_B = \mathbb{E}_{p(B_t, \mathbf{y}_{\leq t})} \left[\left(\partial_{B_t} \log p(B_t, \mathbf{y}_{\leq t}) \right)^2 \right]. \quad (41)$$

The BI can be conveniently split into two terms:

$$J_B = J_P + J_M, \quad (42)$$

where J_P represents our knowledge about B_t prior to any estimation, and J_M accounts for the contribution of the measurement record $\mathbf{y}_{\leq t}$. Formally, the BI associated with our prior knowledge, J_P , corresponds to the *Fisher information* (FI) [13] evaluated with respect to the estimated field value B_t for the distribution $p(B_t)$ defined in (24), i.e.:

$$J_P = F[p(B_t)] = \mathbb{E}_{p(B_t)} \left[\left(\partial_{B_t} \log p(B_t) \right)^2 \right]. \quad (43)$$

In contrast, J_M reads

$$J_M = \mathbb{E}_{p(B_t, \mathbf{y}_{\leq t})} \left[\left(\partial_{B_t} \log p(\mathbf{y}_{\leq t}, B_t) \right)^2 \right] = \int dB_t p(B_t) F[p(\mathbf{y}_{\leq t} | B_t)], \quad (44)$$

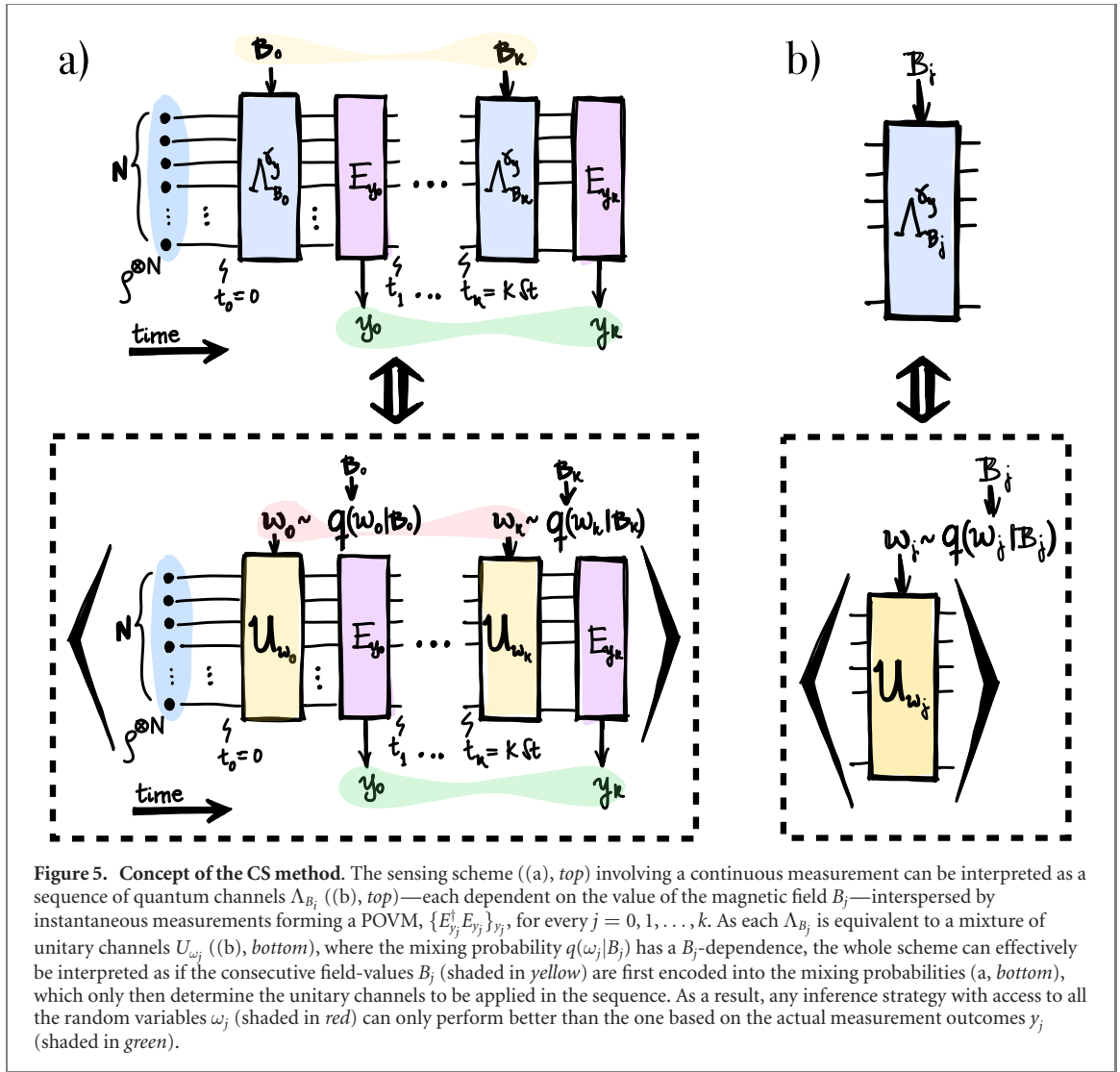
where

$$F[p(\mathbf{y}_{\leq t} | B_t)] = \mathbb{E}_{p(\mathbf{y}_{\leq t} | B_t)} \left[\left(\partial_{B_t} \log p(\mathbf{y}_{\leq t} | B_t) \right)^2 \right] \quad (45)$$

is now the FI of the distribution $p(\mathbf{y}_{\leq t} | B_t)$ evaluated again with respect to B_t .

The prior contribution to the BI (43) can always be ignored by letting the prior distribution describing our knowledge about the field at $t = 0$, $p(B_0)$, be of infinite width [40]—see appendix E.1 where we demonstrate that $J_P = 0$ for $\sigma_0^2 \rightarrow \infty$. In contrast, the contribution of the measurement record to the BI (44) in general requires one to explicitly evaluate the FI of $p(\mathbf{y}_{\leq t} | B_t)$ for every value B_t may take. In this work, we avoid doing so by constructing an upper-bound on $F[p(\mathbf{y}_{\leq t} | B_t)]$ that is determined by the noise in the atomic magnetometry scenario, which turns out to be independent of the form of the continuous measurement and the initial state of the atoms.

In order to do so, we discretize the time evolution such that $t = k\delta t$ —this does not prevent us from obtaining the results in the continuous-time limit, which can be recovered by finally letting $\delta t \rightarrow 0$ [30]. As a result, the continuous-measurement record $\mathbf{y}_{\leq t}$ can be described by a finite set of (time-ordered) outcomes, $\mathbf{y}_k = \{y_0, y_1, \dots, y_k\}$, collected at the end of each time interval indexed by $j = 0, 1, \dots, k$. Similarly, the evolution of the magnetic field corresponds now to a discrete-time stochastic process, $\mathbf{B}_k = \{B_0, B_1, \dots, B_k\}$, with B_j representing the value taken by the B -field throughout the j th time interval. Moreover, the conditional quantum state of the ensemble at time $t = k\delta t$, previously described by



$\rho_{(c)}(t) \equiv \rho(t|y_{\leq t})$ as the solution to the conditional dynamics (6), can now be explicitly written as

$$\rho_k \equiv \rho(t|y_k) = \frac{E_{y_k} \Lambda_{B_k} [\dots E_{y_1} \Lambda_{B_1} [E_{y_0} \Lambda_{B_0} [\rho_0] E_{y_0}^\dagger] E_{y_1}^\dagger \dots] E_{y_k}^\dagger}{\text{Tr} \left\{ E_{y_k} \Lambda_{B_k} [\dots E_{y_1} \Lambda_{B_1} [E_{y_0} \Lambda_{B_0} [\rho_0] E_{y_0}^\dagger] E_{y_1}^\dagger \dots] E_{y_k}^\dagger \right\}}, \quad (46)$$

where the continuous measurement is formally described by a set of measurement operators⁵ E_{y_j} that form a positive-operator valued measure (POVM), i.e. $\{E_{y_j}^\dagger E_{y_j}\}_{y_j}$ with $\sum_{y_j} E_{y_j}^\dagger E_{y_j} = 1$ for every j . In the expression (46), every Λ_{B_j} denotes the *quantum channel* (a completely positive trace-preserving map) describing the evolution of the atomic ensemble during the j th interval in between measurements. Such evolution will then solely incorporate the Larmor precession under the magnetic field B_j (constant in that time interval), and the collective decoherence. A scheme describing this sequence of channels interspersed with measurements is depicted in figure 5(a) (top).

The denominator of the expression (46) should be interpreted as the probability of registering the measurement record y_k given a particular (discrete) trajectory of the B -field B_k , i.e.:

$$p(y_k|B_k) = \text{Tr} \left\{ E_{y_k} \Lambda_{B_k} [\dots E_{y_1} \Lambda_{B_1} [E_{y_0} \Lambda_{B_0} [\rho_0] E_{y_0}^\dagger] E_{y_1}^\dagger \dots] E_{y_k}^\dagger \right\}. \quad (47)$$

Moreover, its *marginal* distribution can then be determined using the Bayes' rule as

$$p(y_k|B_k) = \frac{p(y_k, B_k)}{p(B_k)} = \frac{1}{p(B_k)} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) p(y_k|\mathbf{B}_k), \quad (48)$$

⁵ However, the following construction is also valid for schemes in which the form of the continuous measurement (and, hence, operators E_{y_j}) changes between time intervals, i.e. $E_{y_j}^{(j)} \neq E_{y_j}^{(j')}$ for $j \neq j'$.

constituting the discrete-time equivalent of $p(\mathbf{y}_{\leq t}|B_t)$ appearing in equation (44).

The crucial step in our construction is the observation that the dynamics of the atomic ensemble in between measurements can be CS [22, 53, 54]. In particular, each quantum channel Λ_{B_j} in equation (46) can be equivalently replaced by a probabilistic mixture of unitary evolutions such that only the *classical* mixing probability depends on the instantaneous B -field value B_j —see figure 5(b). In other words, we ‘shift’ the encoding of the parameter B_t from a quantum channel to a classical probability, which we show to be Gaussian. Namely, for any $j = 0, 1, \dots, k$:

$$\Lambda_{B_j}[\cdot] = \int d\omega_j q(\omega_j|B_j) \mathcal{U}_{\omega_j, \delta t}[\cdot], \quad (49)$$

where ω_j is an auxiliary frequency-like random variable distributed according to the Gaussian distribution:

$$\omega_j \sim q(\omega_j|B_j) = \mathcal{N}\left(\gamma B_j, \frac{\gamma y}{\delta t}\right), \quad (50)$$

which parametrises now a *noiseless* Larmor precession of the ensemble within the j th time interval described by the unitary transformation:

$$\mathcal{U}_{\omega_j, \delta t}[\cdot] = e^{-i\omega_j \hat{H} \delta t} \cdot e^{i\omega_j \hat{H} \delta t} \quad (51)$$

with the Hamiltonian $\hat{H} = \hat{J}_y$ consistent with the dynamics (5). A step by step demonstration of this statement is presented in appendix E.2. Moreover, within the setting here considered, the random variable B_j follows a (discrete-time) OU process (3), while the initial B_0 is drawn from a Gaussian prior distribution of variance σ_0^2 .

As a consequence, by now defining the set of particular frequencies ω_j chosen in each time interval, $\boldsymbol{\omega}_k = \{\omega_0, \omega_1, \dots, \omega_k\}$, we can interpret the conditional distribution of the measurement record (47) as a probabilistic average over the frequency-like parameters:

$$p(\mathbf{y}_k|\mathbf{B}_k) = \mathbb{E}_{q(\boldsymbol{\omega}_k|\mathbf{B}_k)}[p(\mathbf{y}_k|\boldsymbol{\omega}_k)] = \int d\boldsymbol{\omega}_k q(\boldsymbol{\omega}_k|\mathbf{B}_k) p(\mathbf{y}_k|\boldsymbol{\omega}_k), \quad (52)$$

where $q(\boldsymbol{\omega}_k|\mathbf{B}_k) = \prod_{j=0}^k q(\omega_j|B_j)$ is just a product distribution and

$$p(\mathbf{y}_k|\boldsymbol{\omega}_k) = \text{Tr}\left\{E_{y_k} \mathcal{U}_{\omega_k, \delta t} [\dots E_{y_1} \mathcal{U}_{\omega_1, \delta t} [E_{y_0} \mathcal{U}_{\omega_0, \delta t} [\rho_0] E_{y_0}^\dagger] E_{y_1}^\dagger \dots] E_{y_k}^\dagger\right\} \quad (53)$$

is now the conditional probability of detecting a set of measurement records \mathbf{y}_k , given a particular set of frequencies $\boldsymbol{\omega}_k$ dictating subsequent unitary evolutions of the atomic ensemble within each time interval.

The decomposition (52) proves the equivalence between a circuit of non-unitary quantum channels describing the noisy dynamics of the atomic ensemble, and the average ($\mathbb{E}_{q(\boldsymbol{\omega}_k|\mathbf{B}_k)}[\dots]$ denoted in figure 5 with $\langle \dots \rangle$) of circuits involving noiseless unitary evolutions, whose frequencies $\boldsymbol{\omega}_k$ are drawn from a distribution $q(\boldsymbol{\omega}_k|\mathbf{B}_k)$ containing all the information about the field trajectory \mathbf{B}_k . We represent this equivalence schematically in figure 5(a).

Focussing on the marginal distribution of interest (48), we substitute the decomposition (52) into it, in order to obtain

$$p(\mathbf{y}_k|\mathbf{B}_k) = \int d\boldsymbol{\omega}_k p(\mathbf{y}_k|\boldsymbol{\omega}_k) \left[\frac{1}{p(\mathbf{B}_k)} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\boldsymbol{\omega}_k|\mathbf{B}_k) \right] = \mathcal{S}_{\boldsymbol{\omega}_k \rightarrow \mathbf{y}_k}[\mathbb{P}_{\mathbf{B}_k}(\boldsymbol{\omega}_k)], \quad (54)$$

where we can now identify $\mathcal{S}_{\boldsymbol{\omega}_k \rightarrow \mathbf{y}_k}[\cdot] = \int d\boldsymbol{\omega}_k p(\mathbf{y}_k|\boldsymbol{\omega}_k) \cdot$ as a stochastic map independent of the estimated \mathbf{B}_k , and define the probability distribution function:

$$\mathbb{P}_{\mathbf{B}_k}(\boldsymbol{\omega}_k) = \frac{1}{p(\mathbf{B}_k)} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\boldsymbol{\omega}_k|\mathbf{B}_k), \quad (55)$$

which effectively describes the information about \mathbf{B}_k contained within $q(\boldsymbol{\omega}_k|\mathbf{B}_k)$.

Crucially, as the FI is generally contractive (monotonic) under the action of stochastic maps [13, 73], we can now upper-bound $F[p(\mathbf{y}_{\leq t}|B_t)]$ in (45) as follows

$$F[p(\mathbf{y}_k|\mathbf{B}_k)] = F[\mathcal{S}_{\boldsymbol{\omega}_k \rightarrow \mathbf{y}_k}[\mathbb{P}_{\mathbf{B}_k}(\boldsymbol{\omega}_k)]] \leq F[\mathbb{P}_{\mathbf{B}_k}(\boldsymbol{\omega}_k)], \quad (56)$$

and evaluate the FI of $\mathbb{P}_{\mathbf{B}_k}(\boldsymbol{\omega}_k)$, which we denote as $\mathbb{P}_{B_t}(\boldsymbol{\omega}_{<t})$ upon letting $\delta t \rightarrow 0$ (with $t = k\delta t$) to recover the continuous-time limit. In appendix E.3, we explicitly show that for the case of infinitely wide ($\sigma_0 \rightarrow \infty$)

Gaussian prior distribution $p(B_0)$, the FI reads

$$F[\mathbb{P}_{B_t}(\omega_{<t})] = \sqrt{\frac{\gamma^2}{\gamma_y q_B}} \tanh\left(t \sqrt{\frac{q_B \gamma^2}{\gamma_y}}\right), \quad (57)$$

and being independent of the estimated B_t directly sets an upper bound in J_M , i.e.:

$$J_M = \int dB_t p(B_t) F[p(\mathbf{y}_{\leq t}|B_t)] \quad (58)$$

$$\leq \int dB_t p(B_t) F[\mathbb{P}_{B_t}(\omega_{<t})] = \sqrt{\frac{\gamma^2}{\gamma_y q_B}} \tanh\left(t \sqrt{\frac{q_B \gamma^2}{\gamma_y}}\right). \quad (59)$$

Hence, by using also the fact that $J_P = 0$ when $\sigma_0 \rightarrow \infty$, we conclude that

$$\Delta^2 \tilde{B}_t \geq (J_P + J_M)^{-1} = J_M^{-1} \quad (60)$$

$$\geq \sqrt{\frac{\gamma_y q_B}{\gamma^2}} \coth\left(t \sqrt{\frac{q_B \gamma^2}{\gamma_y}}\right) \equiv \Delta^2 \tilde{B}_t^{\text{CS}}(q_B) \quad (61)$$

$$\geq \frac{1}{\gamma^2} \frac{\gamma_y}{t} = \Delta^2 \tilde{B}_t^{\text{CS}}(0), \quad (62)$$

where the last inequality follows from $x \coth(x) \geq 1$ for any $x > 0$ with the CS limit, $\Delta^2 \tilde{B}_t^{\text{CS}}(q_B)$, monotonically increasing with q_B —as expected, the error is predicted to deteriorate with an increase in the strength of the field fluctuations.

Although we defer the full proof of the upper bound (59) to appendix E.3, let us note that in the simplest scenario with the magnetic field being time-invariant, the CS limit can be straightforwardly derived and takes the form $\Delta^2 \tilde{B}_t^{\text{CS}}(0)$ as in equation (62). For a constant B -field, the probability distribution (55) simplifies to the product $\mathbb{P}_B(\omega_k) = q(\omega_k|B) = \prod_{j=0}^k q(\omega_j|B)$, whose FI is now evaluated with respect to B and reads:

$$F[\mathbb{P}_B(\omega_k)] = F\left[\prod_{j=0}^k q(\omega_j|B)\right] = \sum_{j=0}^k F[q(\omega_j|B)] = k F[q(\omega_j|B)] = \gamma^2 \frac{k\delta t}{\gamma_y} = \gamma^2 \frac{t}{\gamma_y}, \quad (63)$$

so that the CS limit (62) then directly follows.

In general, the CS limit $\Delta^2 \tilde{B}_t^{\text{CS}}(q_B)$ defined in (61) can always be approximated independently of the value of q_B by splitting the timescales into two regimes, as follows:

$$\Delta^2 \tilde{B}_t^{\text{CS}}(q_B) \approx \begin{cases} \Delta^2 \tilde{B}_t^{\text{CS}}(0) = \frac{1}{\gamma^2} \frac{\gamma_y}{t}, & \text{if } t \lesssim \frac{1}{\gamma} \sqrt{\frac{\gamma_y}{q_B}}, \\ \Delta^2 \tilde{B}_\infty^{\text{CS}}(q_B) = \frac{\sqrt{\gamma_y q_B}}{\gamma}, & \text{if } t \gtrsim \frac{1}{\gamma} \sqrt{\frac{\gamma_y}{q_B}} \quad (\lesssim (M + \gamma_y)^{-1}), \end{cases} \quad (64a)$$

$$\Delta^2 \tilde{B}_\infty^{\text{CS}}(q_B) = \frac{\sqrt{\gamma_y q_B}}{\gamma}, \quad \text{if } t \gtrsim \frac{1}{\gamma} \sqrt{\frac{\gamma_y}{q_B}} \quad (\lesssim (M + \gamma_y)^{-1}), \quad (64b)$$

where the expression (64a) implies that the CS limit (61) always simplifies at short times to its form applicable in absence of field fluctuations, i.e. $\Delta^2 \tilde{B}_t^{\text{CS}}(0)$ in (62).

Finally, let us highlight that the derivation of the CS limit is independent of the measurement dynamics and the initial state, and depends only on the form of the noisy quantum channel describing the evolution of the system in each discretised step between subsequent measurements.

4.2. Steady-state solution of the Kalman filter

Returning to the particular measurement model and the estimation strategy considered, we determine the SS solution of the KF derived in section 3, as it may then be directly compared with the fundamental CS limit determined above. Still, for simplicity, we focus here only on the SS solution when the magnetic field fluctuates according to the Wiener process [55], i.e. the special case of the OU process (3) with $\chi = 0$. The complete solution for $\chi > 0$ is presented in appendix D.

In general, the SS is attained by the KF when its covariance matrix Σ_t no longer changes in time, so that $\frac{d\Sigma_t}{dt} = 0$ in equation (33). This corresponds to the solution of equating the rhs of (33) to zero—solving the corresponding (non-linear) algebraic Riccati equation [70]. In our case, finding the SS solution becomes

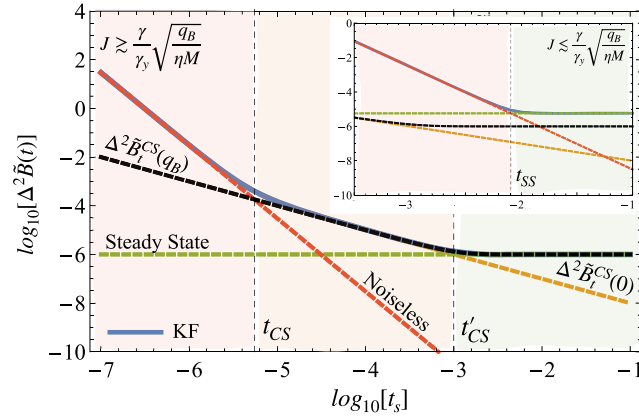


Figure 6. Estimation error as a function of time for ‘large’ (main plot) and ‘small’ (inset) atomic ensembles: here for $J = 10^9$ and $J = 10^3$, respectively. Solid blue lines represent the aMSE of the KF, while the dashed lines of different colours denote: the noiseless solution (red), the exact CS limit (black), the CS limit in the absence of field fluctuations (orange), and the SS solution of the KF (green). Within the main plot the dashed vertical black lines mark the transition times, t_{CS} and t'_{CS} defined in (68), between the noiseless-like ($\propto 1/t^3$, rosé), ‘classical’-like ($\propto 1/t$, orange) and steady-state ($= \sqrt{q_B \gamma_y} / \gamma$, green) regimes. Within the latter two, the CS limit (61) is importantly saturated. In contrast, for ‘small’ ensembles of $J \lesssim \frac{\gamma}{\gamma_y} \sqrt{q_B / (\eta M)}$ (see the inset) the aMSE of the KF saturates, directly after exhibiting the noiseless-like behaviour, the SS solution (65) at t_{SS} defined in (68), without attaining the CS limit at all. Other parameters are set to: $M = 100$ kHz, $\gamma = 1$ kHz mG $^{-1}$, $q_B = 100$ G 2 s $^{-1}$, $\gamma_y = 100$ mHz, $\chi = 0$ and $\eta = 1$, while the rescaled time is defined as $t_s = (M + \gamma_y)t$.

much easier after noting that for $t \gg t^*$ (guaranteed as $t \rightarrow \infty$ for SS) $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} \approx V_{>t^*}(t)$, as in (19b). Then, the SS solution for $\Delta^2 \tilde{B}_t = [\Sigma_t]_{22}$ can be shown (see appendix D) to read

$$\Delta^2 \tilde{B}_t^{SS} = \left(\frac{q_B \gamma_y}{\gamma^2} + \frac{1}{\gamma J} \sqrt{\frac{q_B^3}{M \eta}} e^{(M + \gamma_y)t/2} \right)^{1/2}, \quad (65)$$

whose second term in the parenthesis—the one surviving in the absence of noise ($\gamma_y = 0$)—has been derived previously in [46]. In contrast, it is the first term arising due to the decoherence considered here ($\gamma_y > 0$) that always dominates for large J .

In particular, at the relevant timescales $t \lesssim (M + \gamma_y)^{-1}$ whenever $J \gtrsim \frac{\gamma}{\gamma_y} \sqrt{q_B / (\eta M)}$, the second term within the brackets in (65) becomes negligible and the SS solution (65) coincides with the CS limit (64b) valid at long timescales, i.e.:

$$\Delta^2 \tilde{B}_t^{SS} \Big|_{J \gg 1} \approx \Delta^2 \tilde{B}_\infty^{CS}(q_B) = \frac{\sqrt{q_B \gamma_y}}{\gamma}, \quad (66)$$

which moreover implies that—see also the main plot of figure 6— $\Delta^2 \tilde{B}_t^{SS} \approx \Delta^2 \tilde{B}_t^{CS}(q_B)$, as long as further $t \gtrsim \frac{1}{\gamma} \sqrt{\gamma_y / q_B}$ for which the CS limit takes the form (64b).

Note that this proves that in presence of decoherence ($\gamma_y > 0$) for large enough ensemble size ($J = N/2 \gg 1$) the chosen type of continuous measurement (homodyne-like model) and the initial state of the atomic ensemble (CSS) *always* give the best possible precision in estimating the fluctuating magnetic field ($q_B > 0$) in the SS regime—bearing in mind that the linear-Gaussian approximation made throughout our work must hold. Crucially, this conclusion holds for large atomic ensembles also at short timescales at which the SS solution does not apply, as we will now explicitly show.

4.3. Different regimes exhibited by the estimation error

Once the CS limit and the SS solution of the KF have been explicitly derived, we finally have all the information we need to fully understand the evolution of the estimation error in figure 3(d), which we supplement now with additional plots of the aMSE with respect to time and ensemble size ($J = N/2$) in figures 6 and 7, respectively.

4.3.1. Estimation error as a function of time

We depict the time evolution of the estimation error, i.e. the aMSE of the KF, in figure 6. Importantly, as shown in the main plot, for ‘large’ ensembles (we quantify ‘large’ below) the aMSE attains the fundamental CS limit (61) that holds for any potential measurement and estimation strategy, hence, proving that the magnetometry scheme achieves then the ultimate precision dictated by the decoherence.

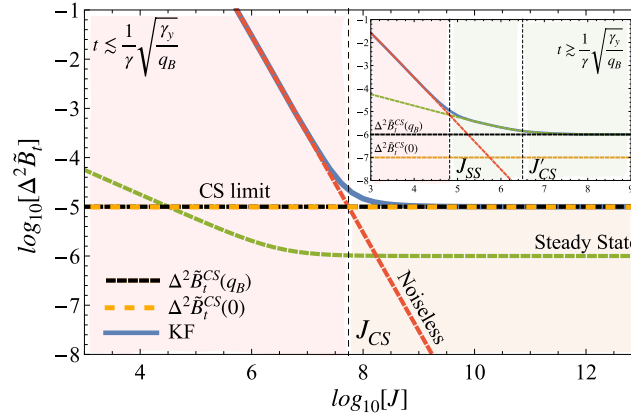


Figure 7. Estimation error as a function of the ensemble size at ‘short’ (main plot) and ‘long’ (inset) timescales: here at $t_s = 10^{-4}$ and $t_s = 10^{-2}$ rescaled times, respectively. The solid blue lines represent the aMSE of the KF, while the dashed lines of different colours denote: the noiseless solution (red), the CS limit (black), the CS limit in the absence of field fluctuations (orange), and the SS solution of the KF (green). At ‘short’ timescales of $t \lesssim \sqrt{\gamma_y/(\gamma^2 q_B)}$ (main plot) the SS solution does not apply even for $J \rightarrow \infty$. The aMSE, after initially following with J the Heisenberg-like behaviour ($\propto 1/J^2$, shaded in rose), saturates around J_{CS} (defined in (69) and marked with a black dashed vertical line) at a constant value—given by the CS limit $\gamma_y/(\gamma^2 t)$ derived in (64a). In contrast, at ‘long’ timescales (inset) the aMSE of the KF is explicitly given by its SS solution (65) for all J above J_{SS} defined in (69) (green-shaded area). As a result, the aMSE, after following with J initially the Heisenberg-like behaviour up to J_{SS} , firstly scales as $1/\sqrt{J}$ before saturating around J'_{CS} defined in (67) again at the CS limit, which is now given by the expression (64b) instead. As in figure 6, we set the other parameters to: $M = 100$ kHz, $\gamma = 1$ kHz mG^{-1} , $q_B = 100$ $\text{G}^2 \text{s}^{-1}$, $\gamma_y = 100$ mHz, $\chi = 0$ and $\eta = 1$, while the rescaled time is defined as $t_s = (M + \gamma_y)t$.

In such a setting, the behaviour of error as a function of time can be intuitively explained. Firstly, at very short timescales for which the atomic decoherence can be ignored, the aMSE follows the *noiseless-like* (or supra-classical) scaling of $1/t^3$. Once the impact of decoherence ‘kicks in’, its behaviour transitions to a ‘classical’-like regime exhibiting $1/t$ scaling, as determined by the CS limit in (64a). At long times, at which the field fluctuations are optimally compensated by the KF, it reaches its SS and the error saturates at a constant value $\sqrt{q_B \gamma_y}/\gamma$. However, let us emphasise again that this value arises due to the presence of decoherence ($\gamma_y > 0$), which in this case dominates the SS solution of the KF in equation (66) and, in fact, assures the error to attain the ultimate CS limit, in particular, its long-time behaviour (64b).

In contrast, for ‘small’ ensembles, as presented in the inset, the aMSE does not reach the CS limit. It is so, as the impact of the atomic decoherence can then be actually neglected in comparison to the field fluctuations. As a consequence, the KF reaches quickly with time its SS solution, which is now effectively independent of the decoherence, i.e. $\Delta^2 \tilde{B}_t^{SS} \approx \sqrt[4]{q_B^3/(M\eta\gamma^2 J^2)}$ in equation (65). One can then interpret the magnetometer to operate in a ‘perfect’ manner, with its precision being dictated fully by just the characteristics of the field.

Note that the above two settings are formally distinguished by whether the SS solution in (65) attains or not the long-time CS limit (64b)—or, in other words, whether the green line can cross the black line in figure 6. Hence, this condition defines naturally what constitutes a ‘large’ or ‘small’ ensemble above, i.e. $J \gtrsim J'_{CS}$ or $J \lesssim J'_{CS}$, respectively, where

$$J'_{CS} = \frac{\gamma}{\gamma_y} \sqrt{\frac{q_B}{\eta M}} \quad (67)$$

follows from the derivation of equation (66).

Moreover, by evaluating the times at which the dominant behaviours of the error appearing in figure 6 cross one another, we explicitly determine the transition times between all the different regimes:

$$t_{CS} = \frac{1}{J} \sqrt{\frac{3}{\eta M \gamma_y}}, \quad t'_{CS} = \frac{1}{\gamma} \sqrt{\frac{\gamma_y}{q_B}}, \quad t_{SS} = \frac{3^{1/3}}{\gamma^2 J^2 \eta M q_B}, \quad (68)$$

which are valid as long as $t \lesssim (M + \gamma_y)^{-1}$ (and the linear-Gaussian approximation holds). In particular, t_{CS} marks the transition time of the aMSE from the noiseless-like to the ‘classical’-like regime, whose later transition to the steady-state regime occurs then at t'_{CS} . Similarly, for ‘small’ ensembles t_{SS} indicates when the aMSE reaches its SS.

4.3.2. Estimation error as a function of the number of atoms

The aMSE as a function of the ensemble size ($J = N/2$) is presented in figure 7, where we show two distinct scenarios differentiated by the timescales considered. Still, we observe that the ultimate CS limit dictated by

the decoherence is attained in both cases as long as a sufficiently large ensemble is considered—proving then the measurement scheme and the estimation procedure to be optimal.

In particular, for ‘short’ timescales (main plot), the error as a function of J initially follows the Heisenberg-like scaling $1/J^2$ before attaining the short-time CS limit (64a) once the effect of the collective noise becomes significant. For ‘long’ timescales (inset), the KF optimally compensates for the field fluctuations and its SS solution applies as long as J is sufficiently large. Still, the long-time CS limit (64b) is achieved as $J \rightarrow \infty$, as it dictates in this case the SS solution (66) affected by the collective noise ($\gamma_y > 0$). The time threshold differentiating between these two cases, t'_{CS} , is then given by (64)—coinciding consistently with the definition in (68).

Importantly, both when the aMSE (solid blue line) attains the CS limit at ‘short’ timescales $t \lesssim t'_{CS}$ (the main plot of figure 7), or at ‘long’ timescales $t \gtrsim t'_{CS}$ (the inset of figure 7), it saturates at a constant value as J is increased. In each case, that occurs at different ensemble sizes: J_{CS} and J'_{CS} , respectively. In the latter case, one can also identify a threshold value J_{SS} , above which the aMSE exhibits a $1/\sqrt{J}$ -scaling before saturating at J'_{CS} . All these can be determined by comparing the dominant behaviours of the error within each regime, and read:

$$J_{CS} = \frac{1}{t} \sqrt{\frac{3}{\eta M \gamma_y}}, \quad J_{SS} = \frac{3^{2/3}}{\gamma t^2 \sqrt{\eta M q_B}}, \quad (69)$$

while J'_{CS} coincides consistently with the definition (67).

5. Conclusions

We have studied the problem of sensing a magnetic field in real time within the canonical atomic magnetometry setting—a polarised spin-ensemble is being continuously probed in the perpendicular direction to induce spin-squeezing of the atoms, so that a quantum-enhanced precision in estimating the field can be maintained. Within our model we have importantly incorporated both the stochastic fluctuations of the field (in the form of an OU process) as well as collective decoherence (affecting the ensemble as a whole) into the magnetometer conditional dynamics, depending on a particular measurement record collected continuously in time.

As a result, while considering the magnetometer evolution at short timescales within the linear-Gaussian regime, we have computed explicitly the optimal estimator—the KF—and studied the behaviour of its error both in time, t , as well as the effective size of the atomic ensemble, $J = N/2$ with N being the total number of atoms. Moreover, we have developed a CS method thanks to which we have established a fundamental limit on the precision that is induced solely by the decoherence. Crucially, as the so-obtained limit applies to any type of state-preparation and continuous-measurement scheme—as long as the conditional dynamics of the magnetometer in between subsequent measurements is not changed—it has allowed us to prove that the continuous measurement model of our interest may often be considered to be optimal in presence of any, even infinitesimal, noise.

In particular, the corresponding aMSE of the KF, which at short timescales always follows the quantum-enhanced behaviour in both time and the number of atoms, i.e. $1/t^3$ and $1/N^2$, respectively, is bound to saturate eventually the limit dictated by the noise. From the perspective of the time dependence, the noise constrains the aMSE to follow a ‘classical’-like $1/t$ scaling before the KF reaches its steady-state solution—which, in fact, coincides with the ultimate long-time precision induced by the decoherence for large atomic ensembles. If instead we focus on the dependence of the aMSE with respect to the number of atoms N , we observe that the impact of the collective noise is even more drastic, as the aMSE (after following the Heisenberg-like behaviour $1/N^2$) saturates eventually at a constant value—whether or not the KF operates within the steady-state regime. Furthermore, our analysis allows to straightforwardly estimate both the times and the numbers of atoms at which the transitions between such regimes occur.

Our work paves the way for finding new methods of incorporating effects of decoherence in real-time sensing protocols, while stemming from Bayesian inference techniques combined with tools previously developed within noisy quantum metrology. In particular, as the CS method we have invoked relies on properties of the effective quantum channel describing the dynamics, and not the design of the continuous-measurement scheme under study, e.g. any adaptive control operations that it incorporates, it should also be directly applicable to sensing protocols involving quantum feedback [74]. On the other hand, although within our work we have focussed on the estimation task in which only past measurement data may be used for inference (filtering), we believe that our results can be naturally extended to smoothing protocols [67–69] that include also retrodiction of data, and have been recently implemented experimentally [58, 59]. Moreover, although we have dealt here with the setting of atomic magnetometry, let us stress that the techniques we have presented can also be applied to other real-time sensing platforms,

e.g.: cavity-based experiments with cold atoms incorporating feedback [75, 76], requiring similar continuous-measurement theory [77]; or optomechanical devices [78–80] and levitated nanoparticles [81, 82] that naturally evolve respecting Gaussian dynamics, and hence directly require Kalman-filtering and alike techniques.

Acknowledgments

This research was supported by the Foundation for Polish Science within the ‘Quantum Optical Technologies’ project carried out within the International Research Agendas programme cofinanced by the European Union under the European Regional Development Fund, as well as the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the EU’s Horizon 2020 Programme (C’MON-QSENS! project) via the National Science Centre Poland.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Unconditional dynamics of $\langle \hat{J}_x(t) \rangle$ with field fluctuations

The unconditional evolution of $\langle \hat{J}_x(t) \rangle$ can be computed from the stochastic set of differential equations (9)–(11), where B_t follows the OU process described in (3). Although the full system of differential equations can in principle be solved numerically, we would like to find approximate analytical solutions valid in particular parameter regimes.

In particular, we focus on timescales short enough, such that we can assure $\omega_L(t)t \ll 1$, where $\omega_L(t) = \gamma B_t$ is the instantaneous Larmor frequency following the fluctuations of B_t . In order to identify and ensure such timescales, we enforce $\overline{\omega_L(t)} t \ll 1$, where $\overline{\omega_L(t)} = \gamma \overline{B_t}$ is now the time-average over the duration t with

$$\overline{B_t} = \frac{1}{t} \int_0^t d\tau B_\tau. \quad (\text{A.1})$$

Replacing B_t by $\overline{B_t}$ within the system of differential equations (9)–(11) and treating it as a constant, we solve them for $\langle \hat{J}_x(t) \rangle$ to get

$$\langle \hat{J}_x(t) \rangle = \frac{J}{2\Theta} e^{-(M+\gamma_x+2\gamma_y+\gamma_z+\Theta)t/4} \left(M - \gamma_x + \gamma_z + \Theta - e^{t\Theta/2} (M - \gamma_x + \gamma_z - \Theta) \right), \quad (\text{A.2})$$

where $\Theta = \sqrt{(M - \gamma_x + \gamma_z)^2 - (4\gamma \overline{B_t})^2}$. Then, by expanding the above expression to leading order in $\overline{B_t}$, we obtain

$$\langle \hat{J}_x(t) \rangle \approx J e^{-(M+\gamma_y+\gamma_z)t/2} \left(1 + 2\gamma^2 \overline{B_t}^2 \frac{2 - 2e^{(M-\gamma_x+\gamma_z)t/2} + t(M - \gamma_x + \gamma_z)}{(M - \gamma_x + \gamma_z)^2} \right), \quad (\text{A.3})$$

$$\approx J e^{-(M+\gamma_y+\gamma_z)t/2} \left(1 + \frac{\gamma^2 \overline{B_t}^2 t^2}{2} \right), \quad (\text{A.4})$$

where in (A.4) we have further assumed $t \lesssim (M - \gamma_x + \gamma_z)^{-1}$, so that

$e^{(M-\gamma_x+\gamma_z)t/2} \approx 1 + \frac{1}{2}(M - \gamma_x + \gamma_z)t + \frac{1}{8}(M - \gamma_x + \gamma_z)^2 t^2$ holds up to the second order in t .

Hence, it seems that we may approximate the dynamics of the mean value of the spin-component \hat{J}_x as

$$\langle \hat{J}_x(t) \rangle \approx J e^{-(M+\gamma_y+\gamma_z)t/2}, \quad (\text{A.5})$$

which constitutes the basis for the linear-Gaussian approximation introduced in equation (12), as long as: (i) $t \lesssim (M - \gamma_x + \gamma_z)^{-1}$ and (ii) $t \lesssim \frac{\sqrt{2}}{\gamma |\overline{B_t}|}$; of which the latter condition allows us to neglect the term in the parenthesis in (A.4). Moreover, as the noise-parameter γ_x does not enter the dynamics any more, after letting $\gamma_x \approx 0$ and reparametrising the measurement strength as in section 2.4, $M \rightarrow M - \gamma_z$, the condition (i) simplifies to $t \lesssim 1/M$ and is naturally satisfied by the more stringent requirement on the rescaled time $t_s = (M + \gamma_y)t$ to always obey $t_s \lesssim 1$, which we ensure throughout the main text.

However, we must be more careful when ensuring the validity of condition (ii). As B_t follows a stochastic process, the time-average $\overline{B_t}$ defined in (A.1) is a random variable in itself. Therefore, we must

further assure that the condition (ii) holds for almost all stochastic trajectories of B_t . We achieve this by applying the 68–95–99.7 rule, which allows us to state that if

$$t \lesssim \frac{\sqrt{2}}{\gamma \left| \mathbb{E}[\overline{B}_t] \pm 2\sqrt{\text{Var}[\overline{B}_t]} \right|}, \quad (\text{A.6})$$

then the approximation (A.5) of (A.4) is valid with 95% probability.

Evaluating the mean of \overline{B}_t defined in (A.1), we obtain

$$\mathbb{E}[\overline{B}_t] = \frac{1}{t} \int_0^t \mathbb{E}[B_\tau] d\tau = \frac{1}{t} \int_0^t B_0 e^{-\chi\tau} d\tau = \frac{B_0}{\chi t} (1 - e^{-\chi t}), \quad (\text{A.7})$$

where $\mathbb{E}[B_\tau] = B_0 e^{-\chi\tau}$ for the OU process (3) [55]. Similarly, we calculate the variance of the time-averaged value of the magnetic field, \overline{B}_t , as

$$\text{Var}[\overline{B}_t] = \mathbb{E}[(\overline{B}_t)^2] - \mathbb{E}[\overline{B}_t]^2 \quad (\text{A.8})$$

$$= \frac{q_B}{2\chi^3 t^2} (4e^{-\chi t} + 2\chi t - e^{-2\chi t} - 3) - \frac{B_0^2}{\chi^2 t^2} (1 - e^{-\chi t})^2, \quad (\text{A.9})$$

using the expression for the two-time correlation function of the OU process (3), i.e. [55]:

$$\mathbb{E}[B_s B_t] = \frac{q_B}{2\chi} \left(e^{-\chi|t-s|} - e^{-\chi(t+s)} \right), \quad (\text{A.10})$$

and evaluating

$$\mathbb{E}[(\overline{B}_t)^2] = \frac{1}{t^2} \int_0^t d\tau_1 \int_0^t d\tau_2 \mathbb{E}[B_{\tau_1} B_{\tau_2}] \quad (\text{A.11})$$

$$= \frac{1}{t^2} \int_0^t d\tau_1 \int_0^t d\tau_2 \frac{q_B}{2\chi} \left(e^{-\chi|\tau_1-\tau_2|} - e^{-\chi\tau_1} e^{-\chi\tau_2} \right) \quad (\text{A.12})$$

$$= \frac{q_B}{2\chi t^2} \left[\int_0^t d\tau_1 \left(\int_0^{\tau_1} d\tau_2 e^{-\chi(\tau_1-\tau_2)} + \int_{\tau_1}^t d\tau_2 e^{-\chi(\tau_1-\tau_2)} \right) - \left(\frac{1 - e^{-\chi t}}{\chi} \right)^2 \right] \quad (\text{A.13})$$

$$= \frac{q_B}{2\chi t^2} \left[\frac{1}{\chi} \int_0^t d\tau_1 (1 - e^{-\chi\tau_1} - e^{\chi(\tau_1-t)} + 1) - \left(\frac{1 - e^{-\chi t}}{\chi} \right)^2 \right] \quad (\text{A.14})$$

$$= \frac{q_B}{2\chi^3 t^2} (4e^{-\chi t} + 2\chi t - e^{-2\chi t} - 3). \quad (\text{A.15})$$

Although when constructing the optimal estimator (KF) of B_t at finite time $t > 0$ we do not want to assume any prior knowledge about the value of the initial field B_0 , for the purpose of this calculation we take the magnetic field to always be initialised in $B_0 = 0$, so that it is solely the presence of the OU process (3) that invalidates the condition (A.6) over time. In such a case, we have

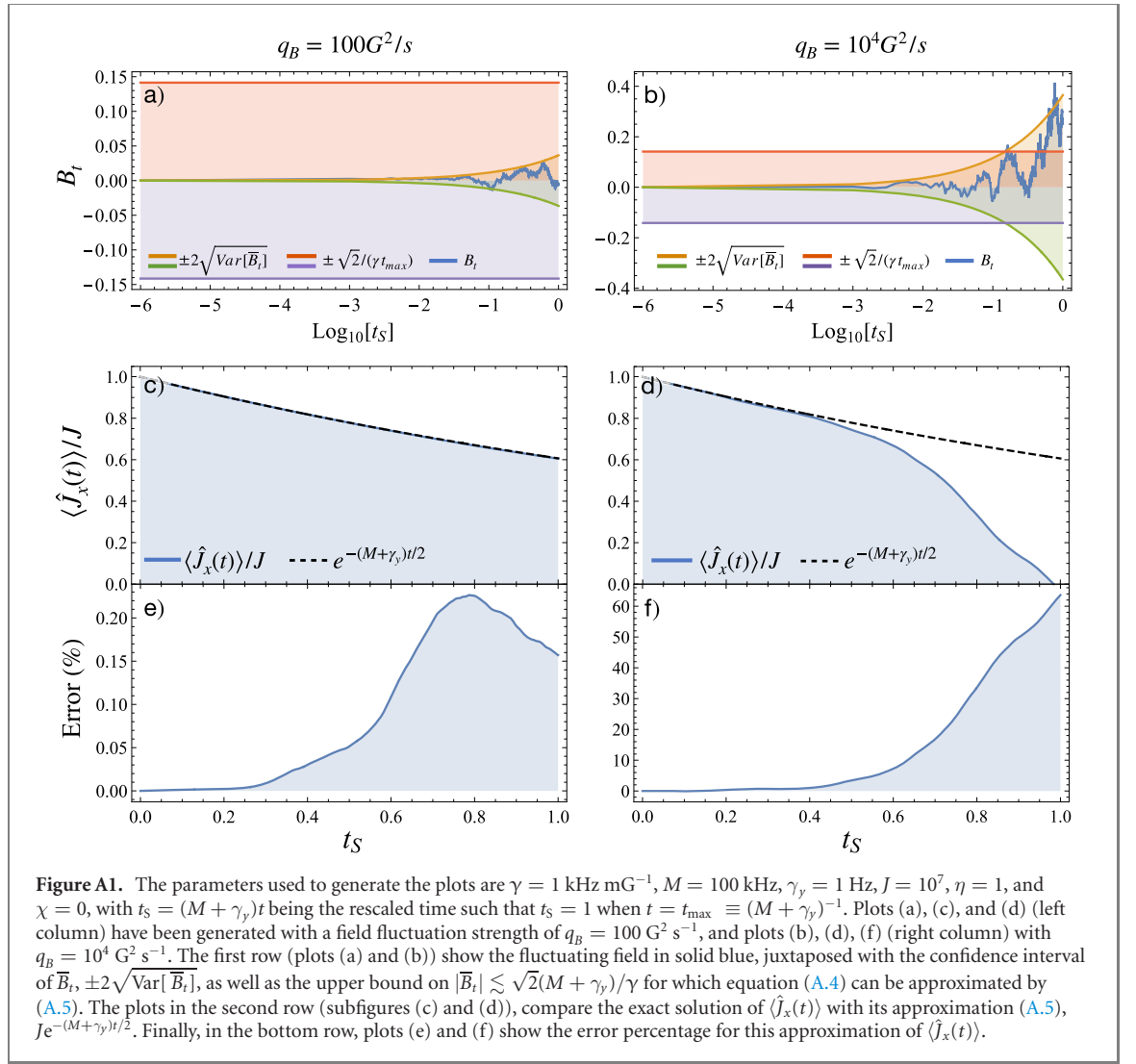
$$\left| \mathbb{E}[\overline{B}_t] \pm 2\sqrt{\text{Var}[\overline{B}_t]} \right| = \left| 2\sqrt{\mathbb{E}[(\overline{B}_t)^2]} \right| = \sqrt{\frac{2q_B}{\chi}} \sqrt{\frac{4e^{-\chi t} + 2\chi t - e^{-2\chi t} - 3}{\chi^2 t^2}} \quad (\text{A.16})$$

$$\approx \sqrt{\frac{2q_B}{\chi}} \sqrt{\frac{2\chi t}{3}} = \sqrt{\frac{4q_B t}{3}}, \quad (\text{A.17})$$

where the above approximation holds as long as $t \ll \frac{4}{3\chi}$. Hence, substituting (A.17) into (A.6), we obtain the desired condition setting an upper limit on valid timescales as a function of the fluctuations strength q_B ,

$$\text{i.e.: } t \lesssim \frac{1}{2\gamma} \sqrt{\frac{3}{q_B t}} \implies t \lesssim \left(\frac{3}{4\gamma^2 q_B} \right)^{1/3}.$$

In summary, we conclude that the linear-Gaussian approximation (A.5) is valid at timescales short enough such that all the three conditions $t \lesssim (M + \gamma_y)^{-1}$, $t \ll \frac{4}{3\chi}$ and $t \lesssim \left(\frac{3}{4\gamma^2 q_B} \right)^{1/3}$ hold. Throughout our work we assure these by focussing on the first one and considering the rescaled time $t_s = (M + \gamma_y)t$ to always fulfil $t_s \lesssim 1$ (as also done in figure A1). However, we must then also require the field fluctuations to be small enough such that the other two conditions are satisfied for any $t_s \lesssim 1$. In particular, this is



achieved by considering the decay parameter and the fluctuation strength of the OU process (3) to fulfil for any $t \lesssim (M + \gamma_y)^{-1}$:

$$\chi \ll \frac{4}{3t} \quad \Rightarrow \quad \chi \ll \frac{4(M + \gamma_y)}{3} \quad (\text{A.18})$$

$$q_B \lesssim \frac{3}{4\gamma^2 t^3} \quad \Rightarrow \quad q_B \lesssim \frac{3(M + \gamma_y)^3}{4\gamma^2}, \quad (\text{A.19})$$

which we more generally state for $r = M + \gamma_y + \gamma_z$ in equation (13) of the main text.

The importance of the upper constraint on the strength of field fluctuations q_B we demonstrate in figure A1, where we present two exemplary Wiener ($\chi = 0$) trajectories of B_t with $q_B = 10^2$ and $q_B = 10^4$ for $3(M + \gamma_y)^3/(4\gamma^2) \approx 10^3$ (all in $[\text{G}^2 \text{ s}^{-1}]$), such that in the latter case the condition (A.19) is clearly invalidated. Correspondingly, as directly seen from figure A1(b), the range of valid timescales (A.6) is surpassed around $t_S = 0.4$, at which the linear-Gaussian approximation (A.5) also ceases to hold, as shown in figure A1(d). In contrast, for $q_B = 10^2$ the approximation (A.5) is consistently maintained with accuracy within 0.25% for any $t_S \lesssim 1$ —see figure A1(e).

Appendix B. Conditional dynamics of the variance $\Delta^2 \hat{J}_z$

The differential equation for the conditional variance of \hat{J}_z introduced in (18) reads

$$d\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} = -4M\eta \langle \Delta^2 J_z(t) \rangle_{(c)}^2 dt + \gamma_y J^2 e^{-(M + \gamma_y)t} dt, \quad (\text{B.1})$$

Table B1. Series expansions of the Bessel functions for $1/\alpha$ and $1/\beta$ around $1/\alpha_0 = 0$ and $1/\beta_0 = 0$, stated to leading order.

$\mathcal{I}_1[2\beta] \approx \frac{e^{2\beta}}{2\sqrt{\pi\beta}}$	$\mathcal{K}_0[2\alpha] \approx \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}e^{-2\alpha}$	$\mathcal{K}_1[2\alpha] \approx \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}e^{-2\alpha}$
$\mathcal{K}_1[2\beta] \approx \frac{1}{2}\sqrt{\frac{\pi}{\beta}}e^{-2\beta}$	$\mathcal{I}_1[2\alpha] \approx \frac{e^{2\alpha}}{2\sqrt{\pi\alpha}}$	${}_0F_1[1, \alpha^2] \approx \frac{e^{2\alpha}}{2\sqrt{\pi\alpha}}$
${}_0F_1[1, \beta^2] \approx \frac{e^{2\beta}}{2\sqrt{\pi\beta}}$	$\mathcal{K}_0[2\beta] \approx \frac{1}{2}\sqrt{\frac{\pi}{\beta}}e^{-2\beta}$	${}_0F_1[2, \alpha^2] \approx \frac{e^{2\alpha}}{2\alpha\sqrt{\pi\alpha}}$

whose solution can be expressed in terms of modified Bessel functions of first and second kind, $\mathcal{I}_\beta[\cdot]$ and $\mathcal{K}_\beta[\cdot]$, and regularized confluent hypergeometric functions ${}_0F_1[\cdot]$, i.e.:

$$\begin{aligned} \langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} = V_e(t) = J e^{-(M+\gamma_y)t/2} & \left(\mathcal{I}_1[2\beta] \left(\sqrt{\eta\gamma_y M} \mathcal{K}_0[2\alpha] - \gamma_y \mathcal{K}_1[2\alpha] \right) \right. \\ & + \mathcal{K}_1[2\beta] \left(\gamma_y \mathcal{I}_1[2\alpha] + \sqrt{\eta\gamma_y M} {}_0F_1[1, \alpha^2] \right) \Big) / \\ & \left(2{}_0F_1[1, \beta^2] \left(\sqrt{\eta\gamma_y M} \mathcal{K}_1[2\alpha] - M\eta \mathcal{K}_0[2\alpha] \right) \right. \\ & \left. + \frac{2\eta M}{M + \gamma_y} \mathcal{K}_0[2\beta] \left((M + \gamma_y) {}_0F_1[1, \alpha^2] + 2\gamma_y J {}_0F_1[2, \alpha^2] \right) \right), \end{aligned} \quad (\text{B.2})$$

where $\alpha = 2J\sqrt{\eta\gamma_y M}/(M + \gamma_y)$ and $\beta = \alpha e^{-(M+\gamma_y)t/2}$. The behaviour of the solution (B.2) can be better understood when broken down into different regimes.

In order to do so, the first step is to expand the modified Bessel functions and the regularized confluent hypergeometric functions around infinity up to leading order—such an approximation is assured due to $\alpha \gg 1$ and $\beta \gg 1$. The relevant expansions are shown in table B1.

Substituting the leading-order expansions of table B1 into the solution (B.2) and approximating $e^{4\beta} \approx e^{-2\alpha(-2+t(M+\gamma_y))}$, the variance of $\hat{J}_z(t)$ simplifies to

$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} \approx \frac{1}{2} J e^{-(M+\gamma_y)t/2} \frac{\sqrt{M\gamma_y\eta} \cosh(2Jt\sqrt{M\gamma_y\eta}) + \gamma_y \sinh(2Jt\sqrt{M\gamma_y\eta})}{\sqrt{M\gamma_y\eta} \cosh(2Jt\sqrt{M\gamma_y\eta}) + M\eta \sinh(2Jt\sqrt{M\gamma_y\eta})}. \quad (\text{B.3})$$

Note that if $2Jt\sqrt{M\gamma_y\eta} \gg 1$ then $\cosh(2Jt\sqrt{M\gamma_y\eta}) \approx \frac{1}{2}e^{2Jt\sqrt{M\gamma_y\eta}}$ and $\sinh(2Jt\sqrt{M\gamma_y\eta}) \approx \frac{1}{2}e^{2Jt\sqrt{M\gamma_y\eta}}$. As a result, expression (B.3) simplifies further to the form stated in equation (19b), i.e.:

$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} \approx V_{>t^*}(t) = \frac{1}{2} J e^{-(M+\gamma_y)t/2} \sqrt{\frac{\gamma_y}{\eta M}}. \quad (\text{B.4})$$

Analogously, if $2Jt\sqrt{M\gamma_y\eta} \ll 1$, then $\cosh(2Jt\sqrt{M\gamma_y\eta}) \approx 1$ and $\sinh(2Jt\sqrt{M\gamma_y\eta}) \approx 2Jt\sqrt{M\gamma_y\eta}$, and the expression (B.3) simplifies to the form stated in equation (19a):

$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} \approx V_{<t^*}(t) = J e^{-(M+\gamma_y)t/2} \frac{(1 + 2Jt\gamma_y)}{2 + 4JtM\eta}. \quad (\text{B.5})$$

Moreover, since $2Jt\sqrt{M\gamma_y\eta} \ll 1$, we can then derive the condition

$$t \ll t^* = \frac{1}{2J\sqrt{M\gamma_y\eta}} \implies 2Jt\gamma_y \ll \sqrt{\frac{\gamma_y}{M\eta}}, \quad (\text{B.6})$$

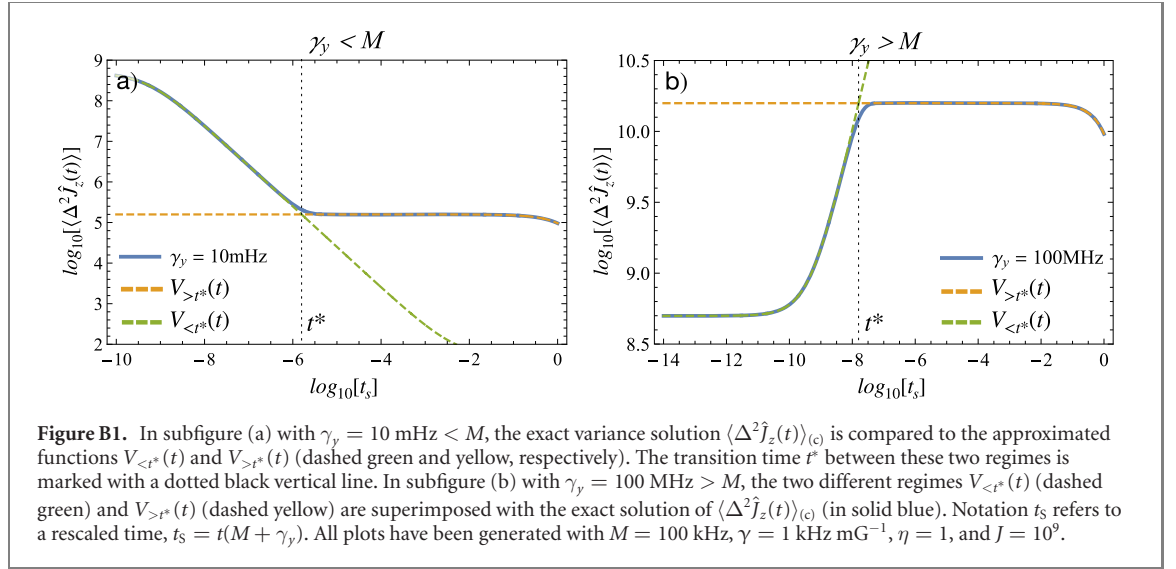
for which approximation (B.5) holds.

From equation (B.5) it also follows that if $\gamma_y < M$, then $(1 + 2Jt\gamma_y) \approx 1$ and, hence, for sufficiently short times scales $t \ll t^* < (M + \gamma_y)^{-1}$ we can always write

$$\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)} \approx \frac{J}{2 + 4JtM\eta} e^{-(M+\gamma_y)t/2} \approx \frac{J}{2 + 4JtM\eta}, \quad (\text{B.7})$$

which is the noiseless solution derived originally in [44]. In the other direction, by showing that $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ is a non-decreasing function at $t = 0_+$ if $\gamma_y \geq \eta M$, we can prove that (B.7) holds only if $\gamma_y < \eta M$ for $J \gg 1$. Namely, that we can consider the global decoherence γ_y to be insignificant at small times $t \ll t^*$ only when $\gamma_y < \eta M$. In order to do so, we differentiate the expression (B.5) w.r.t. t and then let $t \rightarrow 0$, i.e.:

$$\lim_{t \rightarrow 0} \frac{d}{dt} [V_{<t^*}(t)] = -\frac{1}{4} J [\gamma_y(1 - 4J) + M(1 + 4J\eta)]. \quad (\text{B.8})$$



Setting the above expression to zero, we find the value of γ_y for which the derivative changes signs at $t = 0$, i.e.:

$$\gamma_y = \frac{M + 4JM\eta}{4J - 1} = M\eta + \frac{M(\eta + 1)}{4J} + \frac{M(\eta + 1)}{16J^2} + O\left(\frac{1}{J^3}\right), \quad (\text{B.9})$$

which can be correctly approximated as $\gamma_y = \eta M$ when $J \gg 1$. Hence,

$$\lim_{t \rightarrow 0} \frac{d}{dt} [V_{< t^*}(t)] \geq 0 \quad \text{if } \gamma_y \geq \eta M, \quad \text{and} \quad \lim_{t \rightarrow 0} \frac{d}{dt} [V_{< t^*}(t)] < 0 \quad \text{if } \gamma_y < \eta M, \quad (\text{B.10})$$

what proves that $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ can be approximated as in (B.7) at short enough timescales *only* when $\gamma_y < \eta M$.

Finally, in figure B1 we compare the exact solution for the variance given in (B.2) with the short/long-time approximations $V_{< t^*}(t)$ and $V_{> t^*}(t)$ shown in (B.5) and (B.4), respectively.

Appendix C. Construction of the Kalman filter for correlated dynamics

Consider the following system of stochastic differential equations within the Itô calculus:

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{x}_t, t)dt + \mathbf{B}(\mathbf{x}_t, t)d\mathbf{w}_t, \quad (\text{C.1})$$

$$d\mathbf{y}_t = \mathbf{C}(\mathbf{x}_t, t)dt + d\mathbf{v}_t, \quad (\text{C.2})$$

where $d\mathbf{w}_t$ and $d\mathbf{v}_t$ are process and measurement Wiener noise-terms, respectively, that exhibit zero mean $\mathbb{E}[d\mathbf{w}_t] = \mathbb{E}[d\mathbf{v}_t] = 0$ and self-correlations:

$$\mathbb{E}[d\mathbf{w}_t d\mathbf{w}_s^T] = \mathbf{Q}_t \delta(t - s)dt, \quad (\text{C.3})$$

$$\mathbb{E}[d\mathbf{v}_t d\mathbf{v}_s^T] = \mathbf{R}_t \delta(t - s)dt, \quad (\text{C.4})$$

defined with help of (positive semi-definite) covariances $\mathbf{Q}_t, \mathbf{R}_t \geq 0$. Moreover, the cross-correlations between $d\mathbf{w}_t$ and $d\mathbf{v}_t$ are specified by the matrix \mathbf{S}_t :

$$\mathbb{E}[d\mathbf{w}_t d\mathbf{v}_t^T] = \mathbf{S}_t \delta(t - s)dt, \quad (\text{C.5})$$

which is not necessarily symmetric. Note that the matrices $\mathbf{Q}_t, \mathbf{R}_t$ and \mathbf{S}_t are not fully independent: since the covariance of the ‘overall’ noise $d\mathbf{T} = d\mathbf{w}_t \oplus d\mathbf{v}_t$, i.e.

$$\mathbb{E}[d\mathbf{T}_t d\mathbf{T}_t^T] = \begin{bmatrix} \mathbf{Q}_t & \mathbf{S}_t \\ \mathbf{S}_t^T & \mathbf{R}_t \end{bmatrix} \geq 0, \quad (\text{C.6})$$

must be positive semi-definite by definition (constituting an outer product of a vector), the form that matrices $\mathbf{Q}_t, \mathbf{R}_t$ and \mathbf{S}_t can take is then constrained.

Next, it is convenient to rewrite the system of differential equations (C.1) and (C.2) as [70]:

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{x}_t, t)dt + \mathbf{B}(\mathbf{x}_t, t)d\mathbf{w}_t + \mathbf{D}(\mathbf{x}_t, t) (d\mathbf{y}_t - \mathbf{C}(\mathbf{x}_t, t)dt - d\mathbf{v}_t), \quad (\text{C.7})$$

where $\mathbf{D}(\mathbf{x}_t, t)$ can be set arbitrarily, since $d\mathbf{y}_t - \mathbf{C}(\mathbf{x}_t, t)dt - d\mathbf{v}_t = 0$. By regrouping the terms in equation (C.7) we obtain

$$d\mathbf{x}_t = (\mathbf{A}(\mathbf{x}_t, t) - \mathbf{D}(\mathbf{x}_t, t)\mathbf{C}(\mathbf{x}_t, t))dt + \mathbf{D}(\mathbf{x}_t, t)d\mathbf{y}_t + d\mathbf{Z}_t, \quad (\text{C.8})$$

with $d\mathbf{Z}_t = \mathbf{B}(\mathbf{x}_t, t)d\mathbf{w}_t - \mathbf{D}(\mathbf{x}_t, t)d\mathbf{v}_t$ being now the effective process noise. This allows us to ensure that the correlations between the process and measurement noises, $d\mathbf{Z}_t$ and $d\mathbf{v}_t$, respectively, are vanishing—and after setting $\mathbf{D}(\mathbf{x}_t, t) = \mathbf{B}(\mathbf{x}_t, t)\mathbf{S}_t\mathbf{R}_t^{-1}$ without loss of generality we have

$$\mathbb{E}[d\mathbf{Z}_t d\mathbf{v}_t^T] = (\mathbf{B}(\mathbf{x}_t, t)\mathbf{S}_t - \mathbf{D}(\mathbf{x}_t, t)\mathbf{R}_t) dt = 0. \quad (\text{C.9})$$

As a result, we obtain a set of stochastic differential equations that are equivalent to (C.1) and (C.2):

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{x}_t, t)dt + \mathbf{B}(\mathbf{x}_t, t)\mathbf{S}_t\mathbf{R}_t^{-1} (d\mathbf{y}_t - \mathbf{C}(\mathbf{x}_t, t)dt) + \mathbf{B}(\mathbf{x}_t, t)d\mathbf{U}_t, \quad (\text{C.10})$$

$$d\mathbf{y}_t = \mathbf{C}(\mathbf{x}_t, t)dt + d\mathbf{v}_t, \quad (\text{C.11})$$

where the process noise $d\mathbf{U}_t = \mathbf{B}(\mathbf{x}_t, t)^{-1}d\mathbf{Z}_t = d\mathbf{w}_t - \mathbf{S}_t\mathbf{R}_t^{-1}d\mathbf{v}_t$ is now uncorrelated from the measurement noise, $\mathbb{E}[d\mathbf{U}_t d\mathbf{v}_t^T] = 0$, but the dynamics of the state \mathbf{x}_t explicitly depend on the observation \mathbf{y}_t in (C.11). Moreover, in contrast to (C.3) and (C.4), the self-correlations of process and measurement noises now read

$$\mathbb{E}[d\mathbf{U}_t d\mathbf{U}_t^T] = (\mathbf{Q}_t - \mathbf{S}_t\mathbf{R}_t^{-1}\mathbf{S}_t^T) dt, \quad (\text{C.12})$$

$$\mathbb{E}[d\mathbf{v}_t d\mathbf{v}_t^T] = \mathbf{R}_t dt. \quad (\text{C.13})$$

Now, in case of a linear model, the dynamical matrices in (C.1) and (C.2) simplify to:

$$\mathbf{A}(\mathbf{x}_t, t) = \mathbf{F}_t\mathbf{x}_t, \quad (\text{C.14})$$

$$\mathbf{B}(\mathbf{x}_t, t) = \mathbf{B}_t, \quad (\text{C.15})$$

$$\mathbf{C}(\mathbf{x}_t, t) = \mathbf{H}_t\mathbf{x}_t, \quad (\text{C.16})$$

so that the system of equations (C.10) and (C.11) reads

$$d\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_t dt + \mathbf{B}_t\mathbf{S}_t\mathbf{R}_t^{-1} (d\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_t dt) + \mathbf{B}_t d\mathbf{U}_t, \quad (\text{C.17})$$

$$d\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t dt + d\mathbf{v}_t. \quad (\text{C.18})$$

In such a special case, the optimal estimator $\tilde{\mathbf{x}}_t$ minimising the overall aMSE, which is defined as the trace of the covariance matrix $\Sigma_t = \mathbb{E}[(\mathbf{x}_t - \tilde{\mathbf{x}}_t)(\mathbf{x}_t - \tilde{\mathbf{x}}_t)^T]$, i.e. $\text{Tr}\{\Sigma_t\} = \mathbb{E}[\|\mathbf{x}_t - \tilde{\mathbf{x}}_t\|^2]$, is referred to as the KF and given by the solution of the so-called Kalman–Bucy equation [70]:

$$d\tilde{\mathbf{x}}_t = \mathbf{F}_t\tilde{\mathbf{x}}_t dt + \mathbf{\Gamma}_t (d\mathbf{y}_t - \mathbf{H}_t\tilde{\mathbf{x}}_t dt) \quad (\text{C.19})$$

where $\mathbf{\Gamma}_t$ is the Kalman gain

$$\mathbf{\Gamma}_t = (\Sigma_t\mathbf{H}_t^T + \mathbf{B}_t\mathbf{S}_t)\mathbf{R}_t^{-1}. \quad (\text{C.20})$$

Nonetheless, let us note that in order to evaluate the Kalman gain (C.20), one must determine beforehand the (optimal) covariance matrix Σ_t , which evolves according to the non-linear differential equation in the Riccati form [70]:

$$\frac{d\Sigma_t}{dt} = \mathbf{F}_t\Sigma_t + \Sigma_t\mathbf{F}_t^T - \mathbf{\Gamma}_t\mathbf{R}_t\mathbf{\Gamma}_t^T + \mathbf{B}_t\mathbf{Q}_t\mathbf{B}_t^T \quad (\text{C.21})$$

$$= (\mathbf{F}_t - \mathbf{B}_t\mathbf{S}_t\mathbf{R}_t^{-1}\mathbf{H}_t)\Sigma_t + \Sigma_t(\mathbf{F}_t - \mathbf{B}_t\mathbf{S}_t\mathbf{R}_t^{-1}\mathbf{H}_t)^T - \Sigma_t\mathbf{H}_t^T\mathbf{R}_t^{-1}\mathbf{H}_t\Sigma_t + \mathbf{B}_t(\mathbf{Q}_t - \mathbf{S}_t\mathbf{R}_t^{-1}\mathbf{S}_t^T)\mathbf{B}_t^T. \quad (\text{C.22})$$

From the practical perspective, however, the solution to equation (C.22) can be determined (often only numerically) and stored in advance, so that in real-life applications the construction of the KF, $\tilde{\mathbf{x}}_t$, as the

solution to equation (C.19) can still be performed fast and ‘on the fly’, while the observations y_t are constantly gathered.

Appendix D. Steady-state solution of the Kalman filter for $\chi \neq 0$

Let us consider the Riccati differential equation (C.22) (equivalent to (33)), which specifies the evolution of the covariance matrix Σ_t for the KF with the dynamical matrices F_t , B_t , and H_t , as well as the noise-correlation matrices Q_t , R_t , and S_t defined in equations (28) and (32) of the main text, respectively.

For simplicity, we rename the elements of the covariance matrix as

$$\Sigma_t = \begin{pmatrix} x(t) & y(t) \\ y(t) & z(t) \end{pmatrix} \geq 0, \quad (\text{D.1})$$

so that by resorting to the evolution of Σ_t in (C.22) and setting $\frac{d\Sigma}{dt} = 0$ —with (C.22) formally constituting then a continuous-time algebraic Riccati equation [70]—we obtain a system of equations describing the SS as

$$\begin{aligned} -8M\eta V_{>t^*}(t) x(t) - 4M\eta x^2(t) - 2\gamma J e^{-(M+\gamma_y)t/2} y(t) &= 0, \\ -\chi y(t) - 4M\eta V_{>t^*}(t) y(t) - 4M\eta x(t)y(t) - \gamma J e^{-(M+\gamma_y)t/2} z(t) &= 0, \\ q_B - 4M\eta y^2(t) - 2\chi z(t) &= 0, \end{aligned} \quad (\text{D.2})$$

where we have used the fact that the variance $\langle \Delta^2 \hat{J}_z(t) \rangle_{(c)}$ for $t \gg t^*$ (and, hence, in the SS) equals $V_{>t^*}(t)$ as defined in equation (19b).

Being interested only in the SS solution for the aMSE of the magnetic field, $\Delta^2 \tilde{B}_t^{\text{SS}} \equiv z(t)$, one may explicitly solve for $z(t)$ in (D.2) to obtain

$$\begin{aligned} \Delta^2 \tilde{B}_t^{\text{SS}} &= -\frac{\gamma_y \chi}{\gamma^2} - \frac{\chi^3}{4\gamma^2 J^2 M\eta} e^{(M+\gamma_y)t} - \frac{\chi}{\gamma^2 J \sqrt{M\eta}} \sqrt{q_B \gamma^2 + \gamma_y \chi^2} e^{(M+\gamma_y)t/2} \\ &\quad + \frac{1}{2\gamma^2 J M\eta} \sqrt{q_B \gamma^2 + \gamma_y \chi^2} \left(\sqrt{M\eta} + \frac{\chi^2 e^{(M+\gamma_y)t/2}}{2J \sqrt{q_B \gamma^2 + \gamma_y \chi^2}} \right) \\ &\quad \times \sqrt{\chi^2 e^{(M+\gamma_y)t} + 4J \left(\gamma_y J M\eta + e^{(M+\gamma_y)t/2} \sqrt{M\eta (q_B \gamma^2 + \gamma_y \chi^2)} \right)}, \end{aligned} \quad (\text{D.3})$$

which for $J \gg 1$ can be approximated by performing the Taylor expansion in $1/J$ to first order, as follows:

$$\Delta^2 \tilde{B}_t^{\text{SS}} \underset{J \gg 1}{\approx} -\frac{\gamma_y \chi}{\gamma^2} + \frac{\gamma_y}{\gamma^2} \sqrt{\frac{q_B \gamma^2 + \gamma_y \chi^2}{\gamma_y}}. \quad (\text{D.4})$$

In contrast, by letting $\chi \rightarrow 0$ in equation (D.3), we obtain the exact steady-state solution for the aMSE, when the magnetic field follows a Wiener (rather than the OU (3)) process:

$$\Delta^2 \tilde{B}_t^{\text{SS}}|_{\chi=0} = \left(\frac{q_B \gamma_y}{\gamma^2} + \frac{1}{\gamma J} \sqrt{\frac{q_B^3}{M\eta}} e^{(M+\gamma_y)t/2} \right)^{1/2}, \quad (\text{D.5})$$

which is the expression stated in equation (65) of the main text.

Appendix E. Derivation of the CS limit for a fluctuating field

After discretising the dynamics of the magnetic field B_t and the measurement outcomes y_t into $k = t/\delta t$ steps, their particular trajectories follow discrete-time stochastic processes being described by (time-ordered) sets:

$$B_k = \{B_0, B_1, \dots, B_k\} \quad \text{and} \quad y_k = \{y_0, y_1, \dots, y_k\}, \quad (\text{E.1})$$

respectively, while the marginal conditional BCRB (40) after the k th step then reads

$$\Delta^2 \tilde{B}_k \geq J_B^{-1} = (J_P + J_M)^{-1} = \frac{1}{F[p(B_k)] + \int dB_k p(B_k) F[p(y_k|B_k)]}. \quad (\text{E.2})$$

In the first subsection of this appendix, we compute the prior contribution to the BI, i.e. J_P above in (E.2). In the second subsection, we show how the quantum channel describing the dynamics of the atomic ensemble in the absence of continuous measurement can be decomposed into a probabilistic mixture of unitary channels. Finally, in the last subsection, we calculate the FI of $\mathbb{P}_{B_k}(\omega_k)$ defined in (55), which allows us to upper-bound the contribution of the measurement records to the BI, i.e. J_M above in (E.2).

E.1. Prior contribution to the Bayesian information

The marginal probability density function of the magnetic field at time t —or equivalently after the time-step $k = t/\delta t$ —can be written as:

$$p(B_k) = \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) = \int d\mathbf{B}_{k-1} \prod_{j=1}^k p(B_j|B_{j-1}) p(B_0), \quad (\text{E.3})$$

where the initial field B_0 is drawn from a Gaussian distribution which effectively specifies our *a priori* knowledge about the field at $t = 0$, i.e.:

$$B_0 \sim p(B_0) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{B_0^2}{2\sigma_0^2}\right), \quad (\text{E.4})$$

while each transition probability $p(B_j|B_{j-1})$ for all $j = 1, 2, \dots, k$ is given by the OU process (3) as a Gaussian distribution [55]:

$$p(B_j|B_{j-1}) = \sqrt{\frac{1}{2\pi V_P}} \exp\left(-\frac{(B_j - B_{j-1}e^{-\chi\delta t})^2}{2V_P}\right) \quad (\text{E.5})$$

with variance

$$V_P = \frac{q_B}{2\chi}(1 - e^{-2\chi\delta t}). \quad (\text{E.6})$$

Hence, after explicitly evaluating the multiple integrals in (E.3), we arrive at the marginal Gaussian distribution for B_k :

$$p(B_k) = \sqrt{\frac{1}{2\pi V_P^{(k)}}} \exp\left(-\frac{B_k^2}{2V_P^{(k)}}\right), \quad (\text{E.7})$$

whose mean is zero and its variance

$$V_P^{(k)} = \sigma_0^2 e^{-2k\chi\delta t} + \frac{q_B}{2\chi}(1 - e^{-2k\chi\delta t}), \quad (\text{E.8})$$

which in the case of a magnetic field following a Wiener process—a special case of the OU process (3) with $\chi = 0$ —simplifies to $\lim_{\chi \rightarrow 0} V_P^{(k)} = k\delta t q_B + \sigma_0^2$.

In general, the FI evaluated with respect to a Gaussian distribution, e.g. the marginal distribution (E.7), corresponds to the inverse of its variance. Hence, the prior contribution to the BI defined in equation (43) of the main text simply reads

$$J_P = F[p(B_k)] = \frac{1}{V_P^{(k)}} = \frac{1}{\sigma_0^2 e^{-2k\chi\delta t} + \frac{q_B}{2\chi}(1 - e^{-2k\chi\delta t})}, \quad (\text{E.9})$$

and in the continuous-time limit of $\delta t \rightarrow 0$ with $k = t/\delta t$, it becomes

$$J_P = \frac{1}{\sigma_0^2 e^{-2\chi t} + \frac{q_B}{2\chi}(1 - e^{-2\chi t})}. \quad (\text{E.10})$$

In the special case of a Wiener process ($\chi \rightarrow 0$), equation (E.10) reduces to $\lim_{\chi \rightarrow 0} J_P = (q_B t + \sigma_0^2)^{-1}$. On the other hand, for any $\chi, q_B, t \geq 0$, expression (E.10) vanishes if we do not possess any prior knowledge about the initial field B_0 , i.e. when we consider $\sigma_0 \rightarrow \infty$ in equation (E.4), for which $\lim_{\sigma_0 \rightarrow \infty} J_P = 0$.

E.2. Noisy dynamics as a convex mixture of unitary channels

Consider a unitary evolution, which is generated by $\xi \hat{H}$ with \hat{H} being a time-independent Hamiltonian and $\xi \in \mathbb{R}$, being applied to a state ρ_0 for a time interval τ , i.e.:

$$\mathcal{U}_{\xi, \delta t}[\rho_0] = e^{-i\xi \hat{H}\tau} \rho_0 e^{i\xi \hat{H}\tau} \quad (\text{E.11})$$

where the frequency-like parameter ξ is randomly distributed according to a Gaussian probability density:

$$\xi \sim p_{\mu_\tau, \sigma_\tau}(\xi) = \frac{1}{\sqrt{2\pi\sigma_\tau^2}} e^{-\frac{(\xi - \mu_\tau)^2}{2\sigma_\tau^2}}, \quad (\text{E.12})$$

whose mean μ_τ and standard deviation σ_τ are some smooth time-dependent functions.

Theorem 1. *The quantum map Λ_τ obtained by averaging over ξ after a time τ , i.e.:*

$$\rho_\tau = \Lambda_\tau[\rho_0] = \mathbb{E}_{p(\xi)}[\mathcal{U}_{\xi\tau}[\rho_0]] = \int d\xi p_{\mu_\tau, \sigma_\tau}(\xi) e^{-i\xi\hat{H}\tau} \rho_0 e^{i\xi\hat{H}\tau}, \quad (\text{E.13})$$

corresponds to the solution of the master equation:

$$\frac{d\rho_\tau}{d\tau} = -i\omega(\tau)[\hat{H}, \rho_\tau] + \Gamma(\tau) \left(\hat{H}\rho_\tau\hat{H} - \frac{1}{2}\{\hat{H}^2, \rho_\tau\} \right) \quad (\text{E.14})$$

$$= -i\omega(\tau)[\hat{H}, \rho_\tau] - \frac{1}{2}\Gamma(\tau) [\hat{H}, [\hat{H}, \rho_\tau]], \quad (\text{E.15})$$

where the effective time-dependent frequency and decay parameters read, respectively:

$$\omega(\tau) = \mu_\tau + \tau\dot{\mu}_\tau \quad \text{and} \quad \Gamma(\tau) = 2\sigma_\tau^2 \tau \left(1 + \frac{\dot{\sigma}_\tau}{\sigma_\tau} \tau \right). \quad (\text{E.16})$$

Proof. By explicitly differentiating ρ_τ defined in (E.13) with respect to τ , we obtain:

$$\begin{aligned} \frac{d\rho_\tau}{d\tau} &= \frac{d}{d\tau} \Lambda_\tau[\rho_0] = \frac{d}{d\tau} \mathbb{E}_{p(\xi)}[\mathcal{U}_{\xi\tau}[\rho_0]] = \frac{d}{d\tau} \left(\int d\xi p_{\mu_\tau, \sigma_\tau}(\xi) \mathcal{U}_{\xi\tau}[\rho_0] \right) \\ &= \int d\xi \left(\frac{d}{d\tau} p_{\mu_\tau, \sigma_\tau}(\xi) \right) \mathcal{U}_{\xi\tau}[\rho_0] + \int d\xi p_{\mu_\tau, \sigma_\tau}(\xi) \frac{d}{d\tau} \mathcal{U}_{\xi\tau}[\rho_0] \\ &= \frac{1}{\sigma_\tau^3} \mathbb{E}_{p(\xi)} \left[((\xi - \mu_\tau)\sigma_\tau\dot{\mu}_\tau + ((\xi - \mu_\tau)^2 - \sigma_\tau^2)\dot{\sigma}_\tau) \mathcal{U}_{\xi\tau}[\rho_0] \right] - i[\hat{H}, \mathbb{E}_{p(\xi)}[\xi \mathcal{U}_{\xi\tau}[\rho_0]]] \\ &= \frac{\dot{\mu}_\tau}{\sigma_\tau^2} \mathbb{E}_{p(\xi)}[\xi \mathcal{U}_{\xi\tau}[\rho_0]] + \frac{\dot{\sigma}_\tau}{\sigma_\tau^3} \mathbb{E}_{p(\xi)}[(\xi - \mu_\tau)^2 \mathcal{U}_{\xi\tau}[\rho_0]] - \left(\frac{\mu_\tau\dot{\mu}_\tau}{\sigma_\tau^2} + \frac{\dot{\sigma}_\tau}{\sigma_\tau} \right) \rho_\tau \\ &\quad - i[\hat{H}, \mathbb{E}_{p(\xi)}[\xi \mathcal{U}_{\xi\tau}[\rho_0]]], \end{aligned} \quad (\text{E.17})$$

where by using relations for the moments of a Gaussian distribution we can further simplify the following expressions:

$$\mathbb{E}_{p(\xi)}[\xi \mathcal{U}_{\xi\tau}[\rho_0]] = -i\sigma_\tau^2 \tau [\hat{H}, \rho_\tau] + \mu_\tau \rho_\tau, \quad (\text{E.18})$$

$$\mathbb{E}_{p(\xi)}[(\xi - \mu_\tau)^2 \mathcal{U}_{\xi\tau}[\rho_0]] = -\sigma_\tau^4 \tau^2 [\hat{H}, [\hat{H}, \rho_\tau]] + \sigma_\tau^2 \rho_\tau. \quad (\text{E.19})$$

As a result, we can write the dynamics (E.17) as

$$\frac{d\rho_\tau}{d\tau} = -i(\mu_\tau + \dot{\mu}_\tau \tau) [\hat{H}, \rho_\tau] - \sigma_\tau^2 \tau \left(1 + \frac{\dot{\sigma}_\tau}{\sigma_\tau} \tau \right) [\hat{H}, [\hat{H}, \rho_\tau]], \quad (\text{E.20})$$

which is the desired form stated above in equation (E.15). \square

Corollary. *For the case of N spin-1/2 particles evolving for a time τ according to the dynamics (E.15) with $\hat{H} = \hat{J}_y$, $\omega(\tau) = \gamma B$ and $\Gamma(\tau) = \gamma_y$, i.e.:*

$$\frac{d\rho_\tau}{d\tau} = -i\gamma B [\hat{J}_y, \rho_\tau] - \frac{1}{2}\gamma_y [\hat{J}_y, [\hat{J}_y, \rho_\tau]] \quad (\text{E.21})$$

with B being constant over the time τ , the effective quantum map describing the evolution Λ_τ in (E.13) can be interpreted as mixture of unitary channels with the Gaussian mixing probability (E.12) of mean $\mu_\tau = \gamma B$ and standard deviation $\sigma_\tau = \sqrt{\frac{\gamma_y}{\tau}}$.

The above statement can be straightforwardly verified by explicitly computing the effective time-dependent frequency and decay parameters in (E.16) for the choice of $\mu_\tau = \gamma B$ and $\sigma_\tau = \sqrt{\frac{\gamma_y}{\tau}}$, which consistently simplify to:

$$\omega(\tau) = \gamma B + \tau 0 = \gamma B, \quad (\text{E.22})$$

$$\Gamma(\tau) = 2\frac{\gamma_y}{\tau}\tau \left(1 + \frac{1}{2}\left(\frac{\gamma_y}{\tau}\right)^{-1/2} \left(-\frac{\gamma_y}{\tau^2}\right) \left(\frac{\gamma_y}{\tau}\right)^{-1/2} \tau\right) = 2\gamma_y \left(1 - \frac{1}{2}\right) = \gamma_y. \quad (\text{E.23})$$

E.3. Fisher information of $\mathbb{P}_{B_t}(\omega_{<t})$

As discussed in the main text, upper-bounding the measurement-record contribution to the BI, i.e. J_M defined in equation (44), corresponds effectively—see inequality (59)—to computing the FI:

$$F[\mathbb{P}_{B_k}(\omega_k)] = \int d\omega_k \mathbb{P}_{B_k}(\omega_k) \left[-\partial_{B_k}^2 \log(\mathbb{P}_{B_k}(\omega_k)) \right] \quad (\text{E.24})$$

for the probability distribution defined in equation (55):

$$\mathbb{P}_{B_k}(\omega_k) = \frac{1}{p(B_k)} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\omega_k|\mathbf{B}_k). \quad (\text{E.25})$$

Note that $\mathbb{P}_{B_k}(\omega_k)$ contains a Gaussian marginal distribution $p(B_k)$ specified in (E.7), as well as products of Gaussian distributions:

$$p(\mathbf{B}_k) = \prod_{j=1}^k p(B_j|B_{j-1})p(B_0) \quad \text{and} \quad q(\omega_k|\mathbf{B}_k) = \prod_{j=0}^k q(\omega_j|B_j), \quad (\text{E.26})$$

where $p(B_j|B_{j-1})$ is the Gaussian transition probability of the OU process defined in (E.5) with variance V_P given by (E.6), $p(B_0)$ is the prior Gaussian distribution with zero mean and variance σ_0^2 , while each $q(\omega_j|B_j)$ is the mixing probability introduced within the CS method in equation (50), which is also a Gaussian with mean γB_j and variance

$$V_Q = \frac{\gamma_y}{\delta t}. \quad (\text{E.27})$$

It follows from the definitions in (E.26) that the integral in (E.25) can be rewritten as a set of nested integrals, i.e.:

$$\begin{aligned} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\omega_k|\mathbf{B}_k) &= \int d\mathbf{B}_{k-1} \prod_{j=1}^k p(B_j|B_{j-1}) q(\omega_j|B_j) p(B_0) q(\omega_0|B_0) \\ &= q(\omega_k|B_k) \int d\mathbf{B}_{k-1} p(\mathbf{B}_k|B_{k-1}) q(\omega_{k-1}|B_{k-1}) \dots \int dB_1 p(B_2|B_1) q(\omega_1|B_1) \\ &\quad \times \int dB_0 p(B_1|B_0) q(\omega_0|B_0) p(B_0), \end{aligned} \quad (\text{E.28})$$

which we would like to simplify. To do so, we first need to prove the following lemma:

Lemma 2. *Let us consider a recurrence relation between $\mathcal{P}_j(B_j)$ and $\mathcal{P}_{j-1}(B_{j-1})$ for $j = 0, 1, 2, \dots$ as a generalised convolution of Gaussian distributions:*

$$\mathcal{P}_j(B_j) = \int dB_{j-1} \frac{1}{\sqrt{2\pi V_P}} e^{-\frac{(B_j - B_{j-1})^2}{2V_P}} \frac{1}{\sqrt{2\pi V_Q}} e^{-\frac{(\omega_{j-1} - \gamma B_{j-1})^2}{2V_Q}} \mathcal{P}_{j-1}(B_{j-1}), \quad (\text{E.29})$$

where $V_P, V_Q \geq 0$ and $\mathcal{P}_0(B_0) = C_0 e^{-\frac{(B_0 - \mu_0)^2}{2V_0}}$ with some fixed $C_0, V_0 \geq 0$ and $\mu_0 \in \mathbb{R}$.

Then, for all $j \geq 1$:

$$\mathcal{P}_j(B_j) = C_j e^{-\frac{(B_j - \mu_j)^2}{2V_j}} \quad (\text{E.30})$$

where the parameters C_j , μ_j and V_j are given as the solution to the following (coupled) recurrence relations:

$$C_j = C_{j-1} \left(2\pi \left(\gamma^2 V_P + V_Q + \frac{V_P V_Q}{V_{j-1}} \right) \right)^{-1/2} e^{-\frac{(\omega_{j-1} - \gamma \mu_{j-1})^2}{2(V_Q + \gamma^2 V_{j-1})}} \quad (\text{E.31})$$

$$\mu_j = \frac{V_Q \mu_{j-1} + V_{j-1} \gamma \omega_{j-1}}{V_Q + \gamma^2 V_{j-1}} \quad (\text{E.32})$$

$$V_j = V_P + \frac{V_Q V_{j-1}}{V_Q + \gamma^2 V_{j-1}}. \quad (\text{E.33})$$

Proof. As for any recurrence problem, it is sufficient to prove that the solution (E.30) holds for $j = 0$, and that the recurrence relation (E.29) is fulfilled for any $j \geq 1$. The first part is trivially satisfied by definition, while to prove the latter we substitute $P_{j-1}(B_{j-1})$, defined according to (E.30), into (E.29) and explicitly perform the integration, i.e.:

$$\begin{aligned} \mathcal{P}_j(B_j) &= \int dB_{j-1} \frac{1}{\sqrt{2\pi V_P}} e^{-\frac{(B_j - B_{j-1})^2}{2V_P}} \frac{1}{\sqrt{2\pi V_Q}} e^{-\frac{(\omega_{j-1} - \gamma B_{j-1})^2}{2V_Q}} C_{j-1} e^{-\frac{(B_{j-1} - \mu_{j-1})^2}{2V_{j-1}}} \\ &= \frac{C_{j-1}}{\sqrt{2\pi \left(V_Q + \gamma^2 V_P + \frac{V_P V_Q}{V_{j-1}} \right)}} \exp \left\{ -\frac{B_j^2 - 2\alpha B_j + \beta}{2 \left(V_P + \frac{V_Q V_{j-1}}{V_Q + \gamma^2 V_{j-1}} \right)} \right\}, \end{aligned} \quad (\text{E.34})$$

where α and β are constants independent of B_j and B_{j-1} , and equal to

$$\alpha = \frac{V_Q \mu_{j-1} + V_{j-1} \gamma \omega_{j-1}}{V_Q + \gamma^2 V_{j-1}}, \quad (\text{E.35})$$

$$\beta = \frac{V_Q \mu_{j-1}^2 + V_{j-1} \omega_{j-1}^2 + V_P (\omega_{j-1} - \gamma \mu_{j-1})^2}{V_Q + \gamma^2 V_{j-1}}. \quad (\text{E.36})$$

Hence, by ‘completing the square’ we rewrite (E.34) as

$$\mathcal{P}_j(B_j) = \frac{C_{j-1} e^{-\frac{-\alpha^2 + \beta}{2 \left(V_P + \frac{V_Q V_{j-1}}{V_Q + \gamma^2 V_{j-1}} \right)}}}{\sqrt{2\pi \left(V_Q + \gamma^2 V_P + \frac{V_P V_Q}{V_{j-1}} \right)}} \exp \left\{ -\frac{(B_j - \alpha)^2}{2 \left(V_P + \frac{V_Q V_{j-1}}{V_Q + \gamma^2 V_{j-1}} \right)} \right\}, \quad (\text{E.37})$$

and, after substituting for α and β from (E.35) and (E.36), we arrive at the expression (E.30) for $P_j(B_j)$ with C_j , μ_j and V_j specified by the recurrence relations (E.31)–(E.33). \square

Now, using the above lemma we may rewrite equation (E.28) as

$$\int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\omega_k | \mathbf{B}_k) = q(\omega_k | B_k) \mathcal{P}_k(B_k), \quad (\text{E.38})$$

with $\mathcal{P}_k(B_k)$ being now defined according to equation (E.30) with variances V_P and V_Q in (E.31)–(E.33) specified in our case by equations (E.6) and (E.27), respectively. Consequently, the FI of $\mathbb{P}_{B_k}(\omega_k)$ in equation (E.24) reads

$$\begin{aligned} F[\mathbb{P}_{B_k}(\omega_k)] &= \int d\omega_k \mathbb{P}_{B_k}(\omega_k) \left[-\partial_{B_k}^2 \log \left(\frac{1}{p(B_k)} \int d\mathbf{B}_{k-1} p(\mathbf{B}_k) q(\omega_k | \mathbf{B}_k) \right) \right] \\ &= \int d\omega_k \mathbb{P}_{B_k}(\omega_k) \left[-\partial_{B_k}^2 \log \left(\frac{q(\omega_k | B_k)}{p(B_k)} \mathcal{P}_k(B_k) \right) \right] \end{aligned} \quad (\text{E.39})$$

$$= \frac{\gamma^2}{V_Q} - \frac{1}{V_P^{(k)}} + \frac{1}{V_k}, \quad (\text{E.40})$$

where the expression (E.40) follows from the fact that we are dealing with a product (and quotient) of Gaussian distributions within $\log(\dots)$ in (E.39), so that the FI becomes just the sum (and difference) of the inverses of their respective variances. In particular, V_Q/γ^2 is the variance of $q(\omega_k | B_k)$ when treating B_k as the random variable, $V_P^{(k)}$ is the variance of $p(B_k)$ specified in (E.8); while V_k is the variance of $\mathcal{P}_k(B_k)$ given by the recurrence relation (E.33) that must still be solved.

Although the recursive relation (E.33) (with $V_0 = \sigma_0^2$) admits a general solution, for simplicity we present only its form for the relevant setting, in which we do not possess any prior knowledge about the initial field B_0 , i.e. when $\sigma_0 \rightarrow \infty$ in (E.4). Then,

$$V_k|_{\sigma_0 \rightarrow \infty} = \frac{1}{2V_P} + \frac{1}{2\gamma} \sqrt{V_P(4V_Q + V_P\gamma^2)} \left(1 + \frac{2}{-1 + \left(\frac{2V_Q + \gamma(V_P\gamma + \sqrt{V_P(4V_Q + V_P\gamma^2)})}{2V_Q + \gamma(V_P\gamma - \sqrt{V_P(4V_Q + V_P\gamma^2)})} \right)^k} \right), \quad (\text{E.41})$$

which—after substituting for $V_P = \frac{q_B}{2\chi}(1 - e^{-2\chi\delta t})$ and $V_Q = \gamma_y/\delta t$ according to expressions (E.6) and (E.27), respectively, and taking the continuous-time limit $\delta t \rightarrow 0$ with $k = t/\delta t$ —takes the form:

$$\lim_{\delta t \rightarrow 0} \left\{ V_{k=t/\delta t} |_{\sigma_0 \rightarrow \infty} \right\} = \sqrt{\frac{\gamma_y q_B}{\gamma^2}} \coth \left(t \sqrt{\frac{q_B \gamma^2}{\gamma_y}} \right). \quad (\text{E.42})$$

Finally, by noting that the term $\gamma^2/V_Q = \gamma^2\delta t/\gamma_y$ in (E.40) vanishes when letting $\delta t \rightarrow 0$, and so does the term $1/V_P^{(k)}$ when $\sigma_0 \rightarrow \infty$ (see equation (E.8)), we arrive at the FI of $\mathbb{P}_{B_t}(\omega_{<t})$ defined within the continuous-time $\delta t \rightarrow 0$ limit as

$$F[\mathbb{P}_{B_t}(\omega_{<t})]_{\sigma_0 \rightarrow \infty} = \frac{1}{\lim_{\delta t \rightarrow 0} \left\{ V_{k=t/\delta t} |_{\sigma_0 \rightarrow \infty} \right\}} = \sqrt{\frac{\gamma^2}{\gamma_y q_B}} \tanh \left(t \sqrt{\frac{q_B \gamma^2}{\gamma_y}} \right), \quad (\text{E.43})$$

which is the expression stated in equation (57) of the main text.

Appendix F. Effective dynamics averaged over the field fluctuations

We use the results obtained above in appendix E.2, in order to answer the question of what would the effective ensemble dynamics be, if one was *not* to use inference techniques such as the KF in order to track the magnetic field in real time, but rather ignore and, hence, average over the field fluctuations.

For that purpose, let us ignore the impact of continuous measurement on the atomic ensemble, and focus on the ensemble dynamics dictated only by the unitary evolution:

$$\frac{d\rho_t}{dt} = -i\gamma B_t [\hat{J}_y, \rho_t], \quad (\text{F.1})$$

where B_t follows a Wiener process $dB_t = \sqrt{q_B} dW_t$, such that $\mathbb{E}[dW_t^2] = dt$. Given that the ensemble is initially prepared in a pure state $|\psi_0\rangle$, its state at time t then reads

$$|\psi_t\rangle = e^{-i\gamma \hat{J}_y \int_0^t B_\tau d\tau} |\psi_0\rangle. \quad (\text{F.2})$$

The integral of a Wiener process, $\int_0^t W(\tau) d\tau$ with $W(t) := \int_0^t dW_\tau$, constitutes a random variable distributed normally with zero mean and variance $t^3/3$ [55]. Hence, upon defining

$$Z_t := \frac{\gamma}{t} \int_0^t B_\tau d\tau \quad (\text{F.3})$$

that satisfies then $Z_t \sim \mathcal{N}(0, \gamma^2 q_B t/3)$, we may rewrite (F.2) as

$$|\psi_t\rangle = e^{-iZ_t \hat{J}_y} |\psi_0\rangle, \quad (\text{F.4})$$

and the quantum state describing the atomic ensemble after averaging over the field fluctuations reads

$$\rho_t = \mathbb{E}_{Z_t} [|\psi_t\rangle\langle\psi_t|] = \mathbb{E}_{Z_t} \left[e^{-iZ_t \hat{J}_y} |\psi_0\rangle\langle\psi_0| e^{iZ_t \hat{J}_y} \right], \quad (\text{F.5})$$

being averaged over all potential stochastic trajectories of the variable Z_t .

Now, inspecting theorem 1 and noticing that equation (F.5) is just a special case of equation (E.13), we may directly conclude from (E.15) that the average dynamics is described by the following master equation

$$\frac{d\rho_t}{dt} = -\frac{1}{2} \Gamma(t) [\hat{J}_y, [\hat{J}_y, \rho_t]] = \Gamma(t) \mathcal{D}[\hat{J}_y] \rho_t, \quad (\text{F.6})$$

with the resulting decoherence rate given by

$$\Gamma(t) = \gamma^2 q_B t^2, \quad (\text{E.7})$$

which is obtained by substituting into equation (E.16) $\sigma_t^2 = \gamma^2 q_B t/3$, i.e. the variance of the random variable Z_t .

The above short derivation demonstrates that averaging over field fluctuations—instead of following a single trajectory, as done by resorting to, e.g. the KF—effectively leads to a collective noise in the direction of magnetic field (here, in the eigenbasis of \hat{J}_y), whose decoherence rate (E.7), however, increases quadratically with time, $\Gamma(t) \sim t^2$. This contrasts the collective noise model considered throughout the manuscript, whose decoherence rate is constant, $\gamma_y = 1/T_2^*$ in equation (5), being determined by the phenomenological (ensemble) spin-decoherence time T_2^* .

ORCID iDs

Júlia Amorós-Binefa  <https://orcid.org/0000-0002-0035-1236>

Jan Kołodzyński  <https://orcid.org/0000-0001-8211-0016>

References

- [1] Budker D and Romalis M 2007 Optical magnetometry *Nat. Phys.* **3** 227
- [2] Clarke J and Braginski A I 2004 *The SQUID Handbook* vol 1 (New York: Wiley)
- [3] Kitching J 2018 Chip-scale atomic devices *Appl. Phys. Rev.* **5** 031302
- [4] Jensen K, Skarsfeldt M A, Stærkind H, Arnbak J, Balabas M V, Olesen S-P, Bentzen B H and Polzik E S 2018 Magnetocardiography on an isolated animal heart with a room-temperature optically pumped magnetometer *Sci. Rep.* **8** 16218
- [5] Boto E *et al* 2018 Moving magnetoencephalography towards real-world applications with a wearable system *Nature* **555** 657
- [6] Limes M *et al* 2020 Portable magnetometry for detection of biomagnetism in ambient environments *Phys. Rev. Appl.* **14** 011002
- [7] Zhang R *et al* 2020 Recording brain activities in unshielded Earth's field with optically pumped atomic magnetometers *Sci. Adv.* **6** eaba8792
- [8] Pospelov M *et al* 2013 Detecting domain walls of axionlike models using terrestrial experiments *Phys. Rev. Lett.* **110** 021803
- [9] Pustelny S *et al* 2013 The global network of optical magnetometers for exotic physics (GNOME): a novel scheme to search for physics beyond the standard model *Ann. Phys., Lpz.* **525** 659
- [10] Degen C L, Reinhard F and Cappellaro P 2017 Quantum sensing *Rev. Mod. Phys.* **89** 035002
- [11] Budker D and Kozlov M G 2020 Sensing: equation one (arXiv:2011.11043[physics, physics:quant-ph])
- [12] Wang X R, Zheng Y S and Yin S 2005 Spin relaxation and decoherence of two-level systems *Phys. Rev. B* **72** 121303
- [13] Kay S M 1993 *Fundamentals of Statistical Signal Processing (Estimation Theory)* (Englewood Cliffs, NJ: Prentice-Hall)
- [14] Tóth G and Apellaniz I 2014 Quantum metrology from a quantum information science perspective *J. Phys. A: Math. Theor.* **47** 424006
- [15] Demkowicz-Dobrzański R, Jarzyna M and Kołodzyński J 2015 *Progress in Optics* ed E Wolf vol 60 (Amsterdam: Elsevier) pp 345–435 arXiv:1405.7703
- [16] Budker D, Gawlik W, Kimball D F, Rochester S M, Yashchuk V V and Weis A 2002 Resonant nonlinear magneto-optical effects in atoms *Rev. Mod. Phys.* **74** 1153
- [17] Pezzè L *et al* 2018 Quantum metrology with nonclassical states of atomic ensembles *Rev. Mod. Phys.* **90** 035005
- [18] Giovannetti V, Lloyd S and Maccone L 2004 Quantum-enhanced measurements: beating the standard quantum limit *Science* **306** 1330
- [19] Mitchell M W 2017 Number-unconstrained quantum sensing *Quantum Sci. Technol.* **2** 044005
- [20] Maccone L and Giovannetti V 2011 Quantum metrology: beauty and the noisy beast *Nat. Phys.* **7** 376
- [21] Escher B M, de Matos Filho R L and Davidovich L 2011 General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology *Nat. Phys.* **7** 406
- [22] Demkowicz-Dobrzański R, Kołodzyński J and Guţă M 2012 The elusive Heisenberg limit in quantum-enhanced metrology *Nat. Commun.* **3** 1063
- [23] Escher B M, Davidovich L, Zagury N and de Matos Filho R L 2012 Quantum metrological limits via a variational approach *Phys. Rev. Lett.* **109** 190404
- [24] Jeske J, Cole J H and Huelga S F 2014 Quantum metrology subject to spatially correlated Markovian noise: restoring the Heisenberg limit *New J. Phys.* **16** 073039
- [25] Wasilewski W, Jensen K, Krauter H, Renema J J, Balabas M V and Polzik E S 2010 Quantum noise limited and entanglement-assisted magnetometry *Phys. Rev. Lett.* **104** 133601
- [26] Koschorreck M *et al* 2010 Sub-projection-noise sensitivity in broadband atomic magnetometry *Phys. Rev. Lett.* **104** 093602
- [27] Sewell R J, Koschorreck M, Napolitano M, Dubost B, Behbood N and Mitchell M W 2012 Magnetic sensitivity beyond the projection noise limit by spin squeezing *Phys. Rev. Lett.* **109** 253605
- [28] Ma J, Wang X, Sun C P and Nori F 2011 Quantum spin squeezing *Phys. Rep.* **509** 89
- [29] Bar-Shalom Y, Li X R and Kirubarajan T 2004 *Estimation with Applications to Tracking and Navigation (Theory Algorithms and Software)* (New York: Wiley)
- [30] Jacobs K and Steck D A 2006 A straightforward introduction to continuous quantum measurement *Contemp. Phys.* **47** 279
- [31] Thomsen L K, Mancini S and Wiseman H M 2002 Spin squeezing via quantum feedback *Phys. Rev. A* **65** 061801
- [32] Thomsen L K, Mancini S and Wiseman H M 2002 Continuous quantum nondemolition feedback and unconditional atomic spin squeezing *J. Phys. B: At. Mol. Opt. Phys.* **35** 4937

- [33] Takahashi Y, Honda K, Tanaka N, Toyoda K, Ishikawa K and Yabuzaki T 1999 Quantum nondemolition measurement of spin via the paramagnetic faraday rotation *Phys. Rev. A* **60** 4974
- [34] Kuzmich A, Mandel L, Janis J, Young Y E, Eijnisman R and Bigelow N P 1999 Quantum nondemolition measurements of collective atomic spin *Phys. Rev. A* **60** 2346
- [35] Smith G A, Chaudhury S, Silberfarb A, Deutsch I H and Jessen P S 2004 Continuous weak measurement and nonlinear dynamics in a cold spin ensemble *Phys. Rev. Lett.* **93** 163602
- [36] Kuzmich A, Mandel L and Bigelow N P 2000 Generation of spin squeezing via continuous quantum nondemolition measurement *Phys. Rev. Lett.* **85** 1594
- [37] Shah V, Vasilakis G and Romalis M V 2010 High bandwidth atomic magnetometry with continuous quantum nondemolition measurements *Phys. Rev. Lett.* **104** 013601
- [38] Martin Ciurana F *et al* 2017 Entanglement-enhanced radio-frequency field detection and waveform sensing *Phys. Rev. Lett.* **119** 043603
- [39] Jiménez-Martínez R *et al* 2018 Signal tracking beyond the time resolution of an atomic sensor by Kalman filtering *Phys. Rev. Lett.* **120** 040503
- [40] van Trees H L and Bell K L 2007 *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking* (New York: Wiley)
- [41] Fritsche C *et al* 2014 17th Int. Conf. Information Fusion (FUSION) (IEEE) pp 1–8
- [42] Tsang M, Wiseman H M and Caves C M 2011 Fundamental quantum limit to waveform estimation *Phys. Rev. Lett.* **106** 090401
- [43] Deutsch I H and Jessen P S 2010 Quantum control and measurement of atomic spins in polarization spectroscopy *Opt. Commun.* **283** 681
- [44] Geremia J M, Stockton J K, Doherty A C and Mabuchi H 2003 Quantum Kalman filtering and the Heisenberg limit in atomic magnetometry *Phys. Rev. Lett.* **91** 250801
- [45] Mølmer K and Madsen L B 2004 Estimation of a classical parameter with Gaussian probes: magnetometry with collective atomic spins *Phys. Rev. A* **70** 052102
- [46] Stockton J K *et al* 2004 Robust quantum parameter estimation: coherent magnetometry with feedback *Phys. Rev. A* **69** 032109
- [47] Petersen V and Mølmer K 2006 Estimation of fluctuating magnetic fields by an atomic magnetometer *Phys. Rev. A* **74** 043802
- [48] Auzinsh M, Budker D, Kimball D F, Rochester S M, Stalnaker J E, Sushkov A O and Yashchuk V V 2004 Can a quantum nondemolition measurement improve the sensitivity of an atomic magnetometer? *Phys. Rev. Lett.* **93** 173002
- [49] Albarelli F, Rossi M A C, Paris M G A and Genoni M G 2017 Ultimate limits for quantum magnetometry via time-continuous measurements *New J. Phys.* **19** 123011
- [50] Rossi M A C, Albarelli F, Tamascelli D and Genoni M G 2020 Noisy quantum metrology enhanced by continuous nondemolition measurement *Phys. Rev. Lett.* **125** 200505
- [51] Kalman R E 1960 A new approach to linear filtering and prediction problems *J. Basic Eng.* **82** 35
- [52] Kalman R E and Bucy R S 1961 New results in linear filtering and prediction theory *J. Basic Eng.* **83** 95
- [53] Kołodyński J and Demkowicz-Dobrzański R 2013 Efficient tools for quantum metrology with uncorrelated noise *New J. Phys.* **15** 073043
- [54] Matsumoto K 2010 On metric of quantum channel spaces (arXiv:1006.0300[quant-ph])
- [55] Gardiner C W *et al* 1985 *Handbook of Stochastic Methods* vol 3 (Berlin: Springer)
- [56] Vasilakis G, Shah V and Romalis M V 2011 Stroboscopic backaction evasion in a dense alkali-metal vapor *Phys. Rev. Lett.* **106** 143601
- [57] Vasilakis G, Shen H, Jensen K, Balabas M, Salart D, Chen B and Polzik E S 2015 Generation of a squeezed state of an oscillator by stroboscopic back-action-evading measurement *Nat. Phys.* **11** 389
- [58] Bao H *et al* 2020 Spin squeezing of 1011 atoms by prediction and retrodiction measurements *Nature* **581** 159
- [59] Bao H, Jin S, Duan J, Jia S, Mølmer K, Shen H and Xiao Y 2020 Retrodiction beyond the Heisenberg uncertainty relation *Nat. Commun.* **11** 5658
- [60] Wiseman H M and Milburn G J 1993 Quantum theory of field-quadrature measurements *Phys. Rev. A* **47** 642
- [61] van Handel R, Stockton J K and Mabuchi H 2005 Modelling and feedback control design for quantum state preparation *J. Opt. B: Quantum Semiclassical Opt.* **7** S179
- [62] Geremia J M, Stockton J K and Mabuchi H 2006 Tensor polarizability and dispersive quantum measurement of multilevel atoms *Phys. Rev. A* **73** 042112
- [63] de Echaniz S R *et al* 2008 Hamiltonian design in atom–light interactions with rubidium ensembles: a quantum-information toolbox *Phys. Rev. A* **77** 032316
- [64] Breuer H-P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)
- [65] Madsen L B and Mølmer K 2004 Spin squeezing and precision probing with light and samples of atoms in the Gaussian description *Phys. Rev. A* **70** 052324
- [66] Wineland D J, Bollinger J J, Itano W M and Heinzen D J 1994 Squeezed atomic states and projection noise in spectroscopy *Phys. Rev. A* **50** 67
- [67] Tsang M 2010 Optimal waveform estimation for classical and quantum systems via time-symmetric smoothing: II. Applications to atomic magnetometry and hardy's paradox *Phys. Rev. A* **81** 013824
- [68] Zhang J and Mølmer K 2017 Prediction and retrodiction with continuously monitored Gaussian states *Phys. Rev. A* **96** 062131
- [69] Huang Z and Sarovar M 2018 Smoothing of Gaussian quantum dynamics for force detection *Phys. Rev. A* **97** 042106
- [70] Crassidis J and Junkins J 2011 *Optimal Estimation of Dynamic Systems* (Chapman & Hall/CRC Applied Mathematics & Nonlinear Science) (Boca Raton, FL: CRC Press)
- [71] Särkkä S 2013 *Bayesian Filtering and Smoothing*, *Bayesian Filtering and Smoothing* (Cambridge: Cambridge University Press)
- [72] Bobrovsky B Z, Mayer-Wolf E and Zakai M 1987 Some classes of global Cramér–Rao bounds *Ann. Stat.* **15** 1421
- [73] Jarzyna M and Kołodyński J 2020 Geometric approach to quantum statistical inference *IEEE J. Sel. Areas Inf. Theory* **1** 367
- [74] Zhang J, Liu Y-x, Wu R-B, Jacobs K and Nori F 2017 Quantum feedback: theory, experiments, and applications *Phys. Rep.* **679** 1
- [75] Hosten O, Engelsen N J, Krishnakumar R and Kasevich M A 2016 Measurement noise 100 times lower than the quantum-projection limit using entangled atoms *Nature* **529** 505
- [76] Cox K C *et al* 2016 Deterministic squeezed states with collective measurements and feedback *Phys. Rev. Lett.* **116** 093602
- [77] Shankar A, Greve G P, Wu B, Thompson J K and Holland M 2019 Continuous real-time tracking of a quantum phase below the standard quantum limit *Phys. Rev. Lett.* **122** 233602
- [78] Wiczorek W, Hofer S G, Hoelscher-Obermaier J, Riedinger R, Hammerer K and Aspelmeyer M 2015 Optimal state estimation for cavity optomechanical systems *Phys. Rev. Lett.* **114** 223601

- [79] Rossi M, Mason D, Chen J and Schliesser A 2019 Observing and verifying the quantum trajectory of a mechanical resonator *Phys. Rev. Lett.* **123** 163601
- [80] Iwasawa K, Makino K, Yonezawa H, Tsang M, Davidovic A, Huntington E and Furusawa A 2013 Quantum-limited mirror-motion estimation *Phys. Rev. Lett.* **111** 163602
- [81] Setter A *et al* 2018 Real-time Kalman filter: cooling of an optically levitated nanoparticle *Phys. Rev. A* **97** 033822
- [82] Magrini L *et al* 2021 Optimal quantum control of mechanical motion at room temperature: ground-state cooling *Nature* **595** 373–7