Quantum Science and Technology

**PAPER**

Real-time frequency estimation of a qubit without single-shot-readout

I Zohar\(^1\), B Haylock\(^2\), Y Romach\(^3\), M J Arshad\(^1\), N Halay\(^1\), N Drucker\(^1\), R Stöhr\(^4\), A Denisenko\(^3\), Y Cohen\(^3\), C Bonato\(^2\) and A Finkler\(^1\)

\(^1\) Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 7610001, Israel
\(^2\) Institute of Photonics and Quantum Sciences, SUPA, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom
\(^3\) Quantum Machines, Tel-Aviv 6744332, Israel
\(^4\) Third Institute of Physics, University of Stuttgart, Stuttgart 70569, Germany

\* Author to whom any correspondence should be addressed.

E-mail: inbar.aharon@weizmann.ac.il

Keywords: statistical inference, standard quantum limit, quantum sensing, single-shot readout, real time adaptive, magnetometry, frequency estimation

**Abstract**

Quantum sensors can potentially achieve the Heisenberg limit of sensitivity over a large dynamic range using quantum algorithms. The adaptive phase estimation algorithm (PEA) is one example that was proven to achieve such high sensitivities with single-shot readout (SSR) sensors. However, using the adaptive PEA on a non-SSR sensor is not trivial due to the low contrast nature of the measurement. The standard approach to account for the averaged nature of the measurement in this PEA algorithm is to use a method based on 'majority voting'. Although it is easy to implement, this method is more prone to mistakes due to noise in the measurement. To reduce these mistakes, a binomial distribution technique from a batch selection was recently shown theoretically to be superior, as all ranges of outcomes from an averaged measurement are considered. Here we apply, for the first time, real-time non-adaptive PEA on a non-SSR sensor with the binomial distribution approach. We compare the mean square error of the binomial distribution method to the majority-voting approach using the nitrogen-vacancy center in diamond at ambient conditions as a non-SSR sensor. Our results suggest that the binomial distribution approach achieves better accuracy with the same sensing times. To further shorten the sensing time, we propose an adaptive algorithm that controls the readout phase and, therefore, the measurement basis set. We show by numerical simulation that adding the adaptive protocol can further improve the accuracy in a future real-time experiment.

**1. Introduction**

Quantum sensing is a promising technology with many possible applications in fields such as renewable energy [1], condensed matter physics [2–5], biology [6–8], and chemistry [9, 10]. Different quantum systems are studied as quantum sensors [11], and depending on the systems’ interactions with the environment it can be used to sense different physical quantities such as magnetic fields [5], electric fields [8], temperature [12], strain [13], or pressure [14]. One of the advantages of these sensors is the possibility of achieving high sensitivity while overcoming the standard quantum limit (SQL) and reaching the Heisenberg limit (HL) [15].

Recent studies have pushed the sensitivity to the HL using entanglement [16], or quantum algorithms [17]. One algorithm, widely studied, is the phase estimation algorithm (PEA), suggested by Kitaev [18]. This algorithm aims to estimate a phase (\(\phi\)) that a quantum sensor is accumulating due to some interaction with frequency \(f\) with an unknown parameter in the environment. The sensor accumulates the phase at \(K+1\) different sensing times (\(\tau\)) that grow exponentially, \(\tau = 2^k \tau_0\), where \(k\) is an index going from 0 to \(K\). The shortest sensing time \(\tau_0\) limits the dynamic range (DR) of the sensor to
Figure 1. (a) Graphical illustration of the adaptive phase estimation algorithm comprising four steps: (1) A pulse sequence suitable for the estimation of the unknown parameter, given the nature of the interaction between the sensor and the parameter. This pulse sequence will be applied with exponentially growing sensing times. The state of the sensor is measured after every sensing time. (2) Calculating the probability function for the state of the sensor given the unknown parameter. (3) Using Bayes’ Theorem to update the probability function for the parameter. (4) Calculating the optimal variables for extracting maximal information from the next iteration. After $M_k$ iterations for each sensing time, the final distribution will be the estimate of the unknown parameter. (b), (c) Schematic illustration of the measurement outcome of a single-shot (b) or averaged (c) sensor.

The longest sensing time is bounded from above by the dephasing time, $T_2^*$, of the sensor $2^k T_0 < T_2^*$ [19].

The full algorithm is based on a quantum system with multiple quantum bits that carry the process of estimating the phase simultaneously using the quantum Fourier transform [17]. Such multi-qubit systems are still challenging and sometimes not available for every sensing environment. Moreover, a single qubit sensor such as a spin-1/2 offers the ultimate spatial resolution, and any additional gain from entangling it with additional spins is canceled by the increase in sensor size. In these cases, therefore, a system of a single qubit that is incorporated with an adaptive PEA, and making use of quantum–classical interfaces [20, 21], can be of benefit. Here we experimentally demonstrate, in real-time, a non-adaptive PEA scheme in non-ideal but very realistic sensing conditions, and show numerically the advantage of moving this method to an adaptive one.

1.1. Adaptive PEA

The general scheme for applying adaptive PEA (figure 1(a)) consists of a cyclic process of four steps. The first is a pulse sequence applied on the sensor depending on the target frequency, $f$, in question and its interaction with the sensor as expressed in the Hamiltonian, $H(f)$. This pulse sequence will use the same exponentially growing sensing time as in the quantum PEA only in sequential order, from the shortest to the longest, and not simultaneously, similar to the Kitaev’s iterative PEA [22]. After each pulse sequence with one sensing time, the sensor is measured, and the outcome ($u$) can be one of the two states of the sensor—zero or one. This outcome is used in the second step to update the probability function, $P_m(u|f)$, to measure the sensor state, $|0\rangle$ or $|1\rangle$, given that there is interaction with the unknown parameter with frequency $f$. The nature of the sensor’s interaction with the target parameter in the pulse sequence is encoded in the probability function.

The critical step of the algorithm is in step 3, where one applies a Bayesian update to estimate the unknown parameter [23–25]

$$P_{\text{posterior}}(f|u) \propto P_m(u|f)P_{\text{prior}}(f|u)$$

where $P(f|u)$ is the probability function of the measurement outcome given the target parameter, subscript posterior is the new probability after each Bayesian update and prior is the old one from the last update.
\( P(u|f) \) is the probability function of the target parameter given the outcome of the measurement is \( u \), the subscript \( m \) denotes a single outcome. Since the adaptive PEA applies the sensing scheme with different sensing times sequentially, each measurement holds less information about the phase than the quantum PE. The penalty in the full scheme is that each sensing time is measured multiple times by changing one of the sensing variables, as is illustrated in step (3). The number of iterations \( M_k = G + (K - k)F \) for each sensing time grows as the sensing time gets shorter, where \( G \) and \( F \) are optimized parameters, and \( k \) is the index of the sensing time [26]. The adaptive character of the scheme is established in step (2). In this step, the optimal variables for gaining maximal information are calculated based on the last probability function and then transferred to the pulse sequence of the next iteration.

Adaptive PEA has been studied extensively. Theoretical works suggested controlling the sensing phase or the sensing time [27] to enhance sensitivity. Others used numerical simulations [28, 29], and several did experimental studies with different sensors [23, 30] to prove the feasibility and benefits of this protocol. All of these studies were performed with a single-shot readout (SSR) sensor, where the state of the sensor can be measured after one measurement with high fidelity (figure 1(b)). Nevertheless, in some cases, non-SSR sensors are the only possible sensing approach, for instance, for imaging nanoscale biological samples with high special resolution and in ambient conditions. These sensors are characterized by the high ratio of classical noise added in the measurement, for example, low photon collection efficiency in optically read-out systems, compared to the quantum projection noise of the system [11]. This causes the histogram of the measurement outcomes to mix ‘0’ and ‘1’. Therefore, assigning the measurement outcome to one state of the sensor with high fidelity, i.e. in one shot, is impossible (figure 1(c)).

For a non-SSR sensor, the pulse-sequence and the measurement should be applied for many repetitions to assess the sensor state, still with a non-negligible error. This situation requires adjusting the probability function \( P(u|f) \) used in the Bayesian update to the averaged measurement result. The most common and simple solution is to use a threshold that is calculated based on the probability to measure a positive outcome from the sensor at each of the states, which can be a collection of photons for an optically measurable sensor (see appendix ‘Visibility’). In this method the measurement is repeated for \( R \) times and the number of positive outcomes \( r \) is assigned to a state of the sensor, \( u \), based on the calculated threshold; we call this method ‘majority voting’. This approach results in a binary outcome from a large batch of size \( R \) repetitions of the measurement. This method’s benefit is the possibility of using the probability function and Bayesian update as in the SSR sensor scheme [31]. However, it disregards most of the possible outcomes from the \( R \) repetitions by using only a binary span of results. Therefore, it is also more prone to noise, where a noisy measurement can be assigned to the wrong binary option [32].

Since a non-SSR measurement entails repeating the measurement \( R \) times to improve the readout fidelity, we consider the number of positive outcomes, \( r \), out of the full \( R \) batch. This probability distribution, then, is binomial,

\[
P(f|r) = \binom{R}{r} P_d(1|f)^r (1 - P_d(1|f))^{R-r}
\]

(3)

where \( P_d(1|f) \) is the probability of detecting a positive outcome given the sensor state was ‘1’ calculated for the full range of the unknown parameter \( f \). The subscript \( d \) denotes the detection of the positive outcome, and \( r \) is the number of positive outcomes of the measurement [32]. In this case, the information about the phase accumulated due to the external target parameter is encoded in \( P_d \). This method considers the full range of possible outcomes for the averaged measurement. Therefore, a noisy measurement will not result in a mistake but with an error within the range of the noise of the measurement—and therefore leading to a more sensitive estimation [32]. So far, experimental demonstration of the binomial distribution approach and the enhancement in accuracy it offers has not been demonstrated.

1.2. DC magnetometry with non-SSR sensor

In this work, we use the nitrogen-vacancy (NV) center in diamond, a widely used non-SSR sensor at ambient conditions [33, 34]. The NV center is a spin-1 system with a zero field energy splitting between \( m_s = 0 \) and \( m_s = \pm 1 \) spin states of 2.87 GHz (implicitly \( h \) is taken to be equal to \( 1 \)). It is sensitive to DC magnetic fields due to the Zeeman effect \( H(B) = \gamma_e B \cdot S \), where \( \gamma_e \) is the electron gyromagnetic ratio and \( S \) is the spin operator. When the magnetic field is aligned with the \( z \) axis of the spin, the Hamiltonian of the system is simplified to

\[
H_{NV}(B) = DS_z^2 + \gamma_e B_0 S_z.
\]

(4)
This interaction results in an energy splitting between the two ($m_s = \pm 1$) degenerate spin levels of $\Delta \omega = 2\gamma_e B$. Under these conditions and in most instances, each of the two single quantum transitions of the NV center can be practically considered as a two-level spin system (figure 2(a)).

The first step in the PEA scheme is applying the pulse-sequence sensitive to the target field (figure 2(c)). For a DC magnetic field, we use Ramsey interferometry ([35], figure 2(b)). The evolution of the spin in such a pulse-sequence can be simplified when considered in the rotating frame. After initialization to $|0\rangle$, to prepare the sensor, a $\pi/2$ microwave pulse resonant with the eigenenergy of the sensor $\Delta \omega$ is applied, placing the sensor in a superposition state of the two eigenstates $|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\gamma_e\Delta B t}|1\rangle)$.

When the sensor interacts with a small external magnetic field $\Delta B$, it will accumulate a relative phase between the two eigenstates in the rotating frame, which is proportional to the external magnetic field $|\psi(t)\rangle = 2^{-1/2}(|0\rangle + e^{-i\gamma_e\Delta B t}|1\rangle) = 2^{-1/2}(|0\rangle + e^{-i2\pi f\Delta B t - \phi}|1\rangle)$, where $\Delta B$ can also be considered as a small frequency detuning from the resonance frequency of the sensor (as illustrated in figure 2(a)), and therefore, throughout in the manuscript we use the two terms interchangeably. When applying another $\pi/2$ microwave pulse, we project the spin to the eigenstates of $\sigma_z$. If this pulse is rotated by an angle $\phi$ from the preparation pulse, we will project the sensor to a rotated spin basis $\sigma_x e^{-i\phi}$, where $\phi$ is the phase we change in the fourth step of the scheme in figure 2(c). This pulse sequence can estimate external magnetic fields that are within the DR of the measurement (equation (1)).

The second step of the PEA (figure 2(c)) is to calculate the probability function according to the prior-measured state of the sensor $u = |0\rangle/|1\rangle$. This probability function is based on the Ramsey interferometry model for sensing a small external magnetic field $\Delta B$ (figure 2(b)),

$$P_m(u|\Delta B) = \frac{1}{2} \left[ 1 + (-1)^r e^{-(t/T_2^*)} \cos(2\pi f\Delta B t - \phi) \right], \quad (5)$$

where $T_2^*$ is the dephasing time of the sensor. This probability function can be used for an SSR sensor, like the NV center at cryogenic conditions [36], or with the majority voting approach for the non-SSR sensor, like the NV center at ambient conditions [31].

However, for the binomial approach we want to use equation (3), which accounts for all $r$ possible outcome of the repeated measurement. This probability function depends on the probability of detecting a
positive outcome given the external target parameter $P_d(1|f)$. As shown in the theoretical derivation from [32], this probability for sensing an external DC magnetic field is

$$P_d(1|f) = \alpha \left[ 1 - V e^{-(t/T_\alpha)^2} \cos(2\pi f_{\Delta} t - \phi) \right]$$

where $\alpha$ is the sensor’s threshold, and $V$ is the visibility of the sensor (see appendix ‘Visibility’).

### 2. Real-time Bayesian update comparison

We report for the first time on the advantage of the binomial approach over the majority voting in a real-time experiment. The experiment was done at ambient conditions setup using an quantum machines operator-X (QM OPX) to conduct real-time calculations (see appendix ‘Experimental setup’). We collected data from a single NV center with a dephasing time of $T_\alpha^2 = 3.5 \mu s$ (see appendix ‘Sample’). We used five sensing times ($K = 4$) with $t_0 = 100$ ns in the Ramsey measurement pulse sequence (First step in figure 2(c)). The probability function estimating the external magnetic field was constructed with a resolution (binning) of 25 kHz. For each external magnetic field, we applied the scheme twice, once with the majority voting probability function (equation (5)) in the second step and once with the binomial distribution probability function (equation (3)). After the Bayesian update (third step in figure 2(c)), we change the phase of the second $\frac{\pi}{2}$ pulse linearly between zero and $\pi$ in a non-adaptive manner following a predetermined measurement sequence [26] (fourth step in figure 2(c)).

Figure 3(a) presents the iteration number of a measurement of a random magnetic field using the approach described above (Bayesian, non-adaptive). The probability function starts as a uniform distribution. The first iterations apply the shortest sensing time, which guides the probability function to a rough estimation of the frequency. As the iterations advance, the sensing time gets longer, and the estimated frequency gets focused and narrower to a more precise estimate. The frequency at the peak of the probability function in the last iteration is the final estimation for this measurement.

We applied the two approaches for a non-SSR sensor in a non-adaptive scheme on 500 randomly chosen magnetic fields $f_{\Delta R}$ in the range $[-2, 2]$ MHz. For more information about the choice of the range, see appendix ‘Choice of range for random detunings’. We applied the external magnetic field as an off-resonance microwave tone relative to the $|0\rangle \rightarrow | -1\rangle$ transition of the NV, $\omega_{-1}$, at an applied magnetic field of 551 Gauss (close to the NV’s excited-state level anti-crossing), corresponding to $\omega_{-1} = 2\pi \times 1.3322$ GHz. All 500 magnetic fields were measured with seven different repetition numbers $R = (100, 250, 500, 750, 1000, 2500, 5000)$. For each detuning we measured the two approaches in a random order and with all repetition numbers also randomized in the order. After each detuning we refocused the frequency and position of the confocal setup.

To compare the two sensing methods, we calculated the mean square error (MSE):

$$\text{MSE} = \sqrt{V_B} = \sqrt{\left\langle \left( \tilde{f}_B - f_B \right)^2 \right\rangle}$$

based on the estimated frequency $\tilde{f}_B$ calculated from the $P(f|r)$ after every iteration. Our results show a reduction of the MSE with the same measurement time when using the binomial distribution approach. The best MSE achieved was $\approx 0.6$ MHz for $R = 2500$ with a total sensing time of $T = 0.75$ s when using the binomial distribution method. The majority voting method reached this value with almost twice the sensing time ($R = 5000$, $T = 1.25$ s). The lowest possible MSE is limited by the decoherence time of the sensor where $\text{MSE} \geq \sqrt{\frac{T}{T_\alpha}}$.

We note that the MSE for larger $R$ does not improve by much, and we attribute this to the slight improvement of the contrast (see appendix ‘Contrast’). The superiority of the binomial distribution approach is evident also for shorter sensing times, starting from $R = 250$ with $T = 0.34$ s (see figure A5 in appendix ‘Contrast’). To see the improvement in MSE, we plot it as a function of the iteration number for the $R = 2500$ case in figure 3(b). It improves as the iterations progress due to significant improvement in the estimation precision, smaller MSE, compared to the addition of the total sensing time needed for this improvement. We see a good agreement between the experiment and a simulation based on the experimental parameters used in the experiment. The small discrepancy between experiment and simulation can be explained by a little uncertainty in the detection probability for the 0 and 1 state, i.e. $P_d(1|m_0)$ and $P_d(1|m_1)$.

The MSEs calculated for each number of repetitions, $R$, are plotted as a function of the contrast, averaged over all 500 frequencies with the same number of repetitions in figure A5 in appendix ‘Contrast’.
3. Adaptive Bayesian update with binomial distribution

As shown with an SSR sensor, using an adaptive scheme where the measurement variables are optimized based on the updated probability function further improves the sensitivity of the method [23]. In DC magnetometry we look for the optimal readout phase \( \phi \). While one can optimize quantities such as the information gain, this is typically quite complex and adds significant computational overhead. Simpler adaptive rules can be obtained through the Cramér–Rao lower bound (CRLB), which represents the minimum reachable variance for any (unbiased) estimator of \( \phi \). As the CRLB of \( \phi \) is inversely proportional to the Fisher information \( I \), one can target the maximization of \( I \) to improve the estimate of \( \phi \). To find this phase we calculate the Fisher information of the probability function as it is written in equation (7a) and maximize it with respect to the phase \( \phi \),

\[
\mathcal{I}(f_{\Delta B}) = E \left[ \left( \frac{\partial}{\partial f_{\Delta B}} \log(P(r|f_{\Delta B})) \right)^2 \right],
\]

(7a)

\[
\frac{\partial}{\partial \phi} \mathcal{I}(f_{\Delta B}) = 0.
\]

(7b)

where \( E \) is the expectation value.
Figure 4. (a) Simulated experiment with the calculated phases for the different methods: non-adaptive (blue) and our adaptive protocol (orange). After reaching the end of the settings determined by the PEA scheme, measurement times $\tau = T_\star^2$, and phases are chosen at random (non-adaptive) or via adaptive algorithm. (b) Comparison between three different phase calculations for a single experiment with detuning of $f_{\Delta B} = -1.3$ MHz: non-adaptive (blue), our adaptive (orange) and the adaptive-optimized (green), the latter showing an improvement in the number of iterations ($=\text{time}$) needed to attain the correct phase (see main text). The vertical lines (different values of $k$) represent the move to the next $\tau$ in the algorithm.

By solving the optimum problem for the phase and taking the solution that results with the maximum, we find the optimal phase,

$$\phi_{\text{opt}} = 2\pi E[f_{\Delta B}] t - \cos^{-1}\left(\frac{-B}{A}\right),$$

where $A = \frac{r^2}{R^2} + (1 - 2R) \alpha$ and $B = (1 - 2R) \alpha V$ (see appendix ‘Adaptive phase calculation’). Using the number of positive results and the expectation value at each iteration with this optimal phase calculation will result in the next readout phase.

To evaluate the benefit of the adaptive scheme compared to the non-adaptive one, we simulated the experiments based on the dephasing time ($T_\star^2$) and threshold ($\alpha$) (see appendix ‘Sample’) of the NV used in the real-time experiment. Simulations were performed by numerically reproducing the experiment, randomly generating a simulated photon number $r$ from a binomial distribution as in equation (6), using experimental parameter ($R = 10^5$, $G = 3$, $F = 2$), and are presented in figure 4(a). We observe two different regimes. The first one where we increase the $\tau$ exponentially, which is the ‘high DR’ regime. Once we reach $T_\star^2$ we measure from that iteration number at $\tau = T_\star^2$, at the SQL, hence the clear change in slope at iteration number (approximately) 30. In both cases (non-adaptive and adaptive), the probability function was calculated based on the binomial distribution approach. In the simulation of the adaptive scheme, the phase of the Ramsey readout pulse in the next iteration was calculated based on the probability function (step 4 in
whereas in the non-adaptive scheme the phase was linearly ramped between zero and \( \pi \) (see figure 4(b)). The phases calculated for the adaptive simulated experiment show convergence of the phase in a small number of iterations, smaller than \( M_k \) iterations determined theoretically for each sensing time (figure 4(b), orange circles). This convergence raises the possibility of reducing the number of iterations for each sensing time by moving to the next sensing phase once the phase remains steady for three iterations with an error of \( \frac{\pi}{2} \) (figure 4(b), green squares). We denote this method as adaptive-optimized. It has the potential to reduce the total measurement time significantly, which will also improve the sensitivity.

The MSE calculated from simulated results of the two methods, non-adaptive and adaptive, is plotted in figure 4(a) as a function of increasing iterations, where each iteration consists of a new value for the phase. Both methods in the simulation used the binomial distribution approach for the probability function to calculate the final estimation, as this approach proved to be more sensitive also in the real-time experiment.

4. Conclusions

We performed a real-time Bayesian update with an NV center, a non-SSR sensor at ambient conditions. We compared the MSE of the sensor between two calculation methods—majority voting and binomial distribution, and showed that the latter approach has better sensitivity than the former.

We showed by simulation that an adaptive scheme can further improve the MSE, and suggested using it to also reduce the total number of iterations and therefore the total sensing time, and offer extra improvement on the sensitivity. Our simulations suggest that these schemes can achieve a sensitivity four times better than the non-adaptive approach.

This work demonstrates how one can use non-SSR sensors as practical tools in adaptive PEA, and serves as a proof of concept for a specific non-SSR sensor, the NV center in diamond. Nevertheless, it could also be implemented in other sensing systems, as the approach is general. This method can also be used in other sensing schemes, such as ac magnetometry using dynamical decoupling [37] for solid state spin sensors.

Data availability statement

All data shown in the main text and supplementary materials are available in the Zenodo repository [38].

Acknowledgments

We thank B Nadler for invigorating discussions. C B and A F are jointly supported by the ‘Making Connections’ Weizmann-UK program. C B is supported by the Engineering and Physical Sciences Research Council (EP/S000550/1 and EP/V053779/1). A F acknowledges financial support from the Israel Science Foundation (ISF Grant No. 963/19). A F is the incumbent of the Elaine Blond Career Development Chair in Perpetuity and acknowledges the historic generosity of the Harold Perlman Family, research grants from the Abramson Family Center for Young Scientists and the Willner Family Leadership Institute for the Weizmann Institute of Science.

Appendices

Sample

The NV layer was created by a 10 keV nitrogen ion \((^{15}\text{N}^+\) ) implantation with a flux of 80 ions per \( \mu \text{m}^2 \) in an electronic grade (e6) CVD diamond, subsequently annealed in vacuum at a temperature of 950 °C for two hours. A nanopillar structure was then etched in the diamond for enhanced photon collection efficiency [39]. All measurements were performed on a single NV center. The dephasing time of the NV center is \( T_2^* = 3.5 \mu \text{s} \), measured with a standard Ramsey (FID) sequence on resonance. The Rabi contrast of the NV center was about 30% with count rate of 80 kcounts per second.

Experimental setup

The NV center was measured on a custom-built (confocal microscope with a 520 nm laser diode for excitation, dichroic mirror for separating excitation and fluorescence, a band-pass filter for fluorescence counting and two avalanche photodiodes in a Hanbury-Brown and Twiss configuration). We used an OPMX to orchestrate all pulse sequence generation, photon readout, real-time Bayesian estimation and adaptive phase calculation. A local oscillator from a Windfreak SynthNV-Pro was mixed (Marki MLIQ-0218L) with two low-frequency (150 MHz) 90° phase-shifted sine signals from the OPX to produce a single-sideband modulated RF, amplified by an EliteRF (M.02006G424550) broadband amplifier.
As opposed to prior works with NVs [30, 31, 40], here the measurements and Bayesian update are performed by an field-programmable gate array (FPGA)-based computer in real-time (QM OPX). Together with on-the-fly pulse sequence generation, each Bayesian Update (in the non-adaptive case) takes only 0.4 ms to complete (for a probability distribution function of length 400, or 1 μs per frequency bin), with a small overhead of <1 μs for the optimal phase calculation (in the adaptive case, equation (8)). The OPX QUA code is available on github.

Measurements for figure 3(b) were taken on a similar setup, previously described by Arshad et al [41], using a laser-written NV center with $T_2^* = 5.5 \mu s$. A single photon count rate of ≈50 kcps was equivalent to $P_d(1|m_1) \approx 0.011$ and $P_d(1|m_0) \approx 0.016$.

Adaptive phase calculation

Taking equation (3) and the Ramsey model (equation (6)), we derive in this appendix the optimal phase in the adaptive case. First, define the model as $L(f_\beta, \theta) = \alpha [1 + V\cos(2\pi f_\beta \tau - \theta)]$ and calculate its derivative

$$L'(f_\beta, \theta) = \frac{\partial}{\partial f_\beta} L(f_\beta, \theta) = -2\alpha V\pi \tau \sin(2\pi f_\beta \tau - \theta).$$

Next, we use the binomial probability distribution to write down the mean and variance,

$$\mu_r = E[r|f_\beta] = R \cdot L(f_\beta, \theta)$$

$$\sigma_r^2 = E[(r - \mu_r)^2|f_\beta] = R \cdot L(f_\beta, \theta) [1 - L(f_\beta, \theta)].$$

We can approximate $L(f_\beta, \theta) \equiv L \approx \frac{r}{R} + \Delta$ if $\Delta \ll 1$, and then get an expression for the variance in leading orders of $\frac{r}{R}$:

$$\sigma_r^2 = R \cdot L(f_\beta, \theta) [1 - L(f_\beta, \theta)] = RL - RL^2 = R \left[ \left( \frac{r}{R} + \Delta \right) - \left( \frac{r}{R} \right)^2 \right]$$

$$= R \left[ \left( \frac{r}{R} + \Delta \right) - \frac{r^2}{R^2} - 2\Delta \frac{r}{R} + \Delta^2 \right] \approx R \left[ \left( \frac{r}{R} + \Delta \right) - \frac{r^2}{R^2} - 2\Delta \frac{r}{R} \right]$$

$$= R \left[ L - \frac{r^2}{R^2} - 2 \left( L - \frac{r}{R} \right) \frac{r}{R} \right] = R \left[ L + \frac{r^2}{R^2} - 2 \frac{r}{R} L \right] = R \left[ L \left( 1 - 2 \frac{r}{R} \right) + \frac{r^2}{R^2} \right].$$

Now we define the logarithm of the model (likelihood) function:

$$K(r, f_\beta) \equiv \log P(r|f_\beta) = \log \left( \frac{R}{r} \right) + r \log L(f_\beta, \theta) + (R - r) \log(1 - L(f_\beta, \theta))$$

and so

$$\frac{\partial}{\partial f_\beta} K(r, f_\beta) = \frac{rL'}{L} - \frac{(R-r)L'}{1-L} = \left[ \frac{r}{L} - \frac{R-r}{1-L} \right] L' = \left[ \frac{r-RL'}{L(1-L)} \right] L' = \frac{R}{\sigma_r^2} (r - \mu_r) L'(f_\beta, \theta).$$

As we wrote in section 3, the Fisher information can now be explicitly calculated,

$$I(f_\beta) = E \left[ \left( \frac{\partial}{\partial f_\beta} K(r, f_\beta) \right)^2 \right] = E \left[ (r - \mu_r)^2 \frac{R^2}{\sigma_r^2} (L'(f_\beta, \theta))^2 \right] = \frac{R^2}{\sigma_r^2} (L'(f_\beta, \theta))^2$$

$$\approx \frac{R^2}{\sigma_r^2} \left( \frac{R L'(f_\beta, \theta)}{2} \right)^2 = 4R^2\alpha^2 V^2 \pi^2 \tau^2 \frac{\sin^2(2\pi f_\beta \tau - \theta)}{\frac{R^2}{\sigma_r^2} (1 - 2\frac{r}{R}) \alpha [1 + V\cos(2\pi f_\beta \tau - \theta)]}.$$

The last term can be written in a more compact form by denoting

$$A = \frac{r^2}{R^2} + \left( 1 - 2 \frac{r}{R} \right) \alpha$$

$$B = \left( 1 - 2 \frac{r}{R} \right) \alpha V$$

$$C = 4R^2\alpha^2 V^2 \pi^2 \tau^2,$$

such that,

$$I(f_\beta) = \frac{C \sin^2(2\pi f_\beta \tau - \theta)}{A + B\cos(2\pi f_\beta \tau - \theta)}.$$
We maximize the Fisher information (or minimize the Cramér–Rao bound):
\[
\frac{\partial}{\partial \theta} I(f_B) = 0
\]
with two solutions. The first one is a minimum with \( \theta = 2\pi f_B \tau \) and the second solution is \( A \cos(2\pi f_B \tau - \theta) + B = 0 \). Using \( \hat{f}_B = E[f_B] \), gives:
\[
\theta_{\text{opt}} = 2\pi \hat{f}_B \tau - \cos^{-1}\left(\frac{-B}{A}\right).
\]

**Contrast**

As defined previously [23], the contrast, \( C \), for \( R \) repetitions scales as
\[
C = \left[ 1 + \frac{2(\alpha_0 + \alpha_1)}{(\alpha_0 - \alpha_1)^2 R} \right]^{-1/2},
\]
where \( \alpha_0 \) is the number of photons per shot when the NV is in the \( m_s = 0 \) state and \( \alpha_1 \) is the number of photons per shot when the NV is in the \( m_s = 1 \) state. In figure A5 we plot the MSE as a function of the contrast, \( C \). The data presented in this figure is from a different dataset than that plotted in figure 3(b), hence the difference in the MSE. Nevertheless, the point regarding the contrast is still valid.

**Visibility**

In equation (6) we introduced two parameters: the threshold, \( \alpha \) and the visibility, \( V \). Following [32] and for self-consistency, we define them here as
\[
\alpha = \frac{1}{2} \left[ P_d(1|m_0) + P_d(1|m_1) \right]
\]
\[
V = \frac{P_d(1|m_0) - P_d(1|m_1)}{P_d(1|m_0) + P_d(1|m_1)} e^{i(\pi/2)j},
\]
where \( P_d(1|m_i) \), for \( i = 0, 1 \), is the probability of a detector click for the spin in the state \( |i\rangle \). Typical values for our setup were \( P_d(1|m_1) = 0.0251 \) and \( P_d(1|m_0) = 0.03419 \).

**Choice of range for random detunings**

As we explained in the main text, the DR of the sensor is bounded in the range \([-\frac{1}{\tau_0}, \frac{1}{\tau_0}] \) which is \([-5, 5]\) MHz in our case. While we saw no discernible change in the conventional Ramsey curve for detunings larger than 2 MHz, the level of noise increased dramatically for the Bayesian update. We therefore limited the range to \([-2, 2]\) MHz when randomly selecting the 500 detuning used in our dataset. The increase in noise for larger detunings is currently being investigated, but we do not think it affects the results we presented in the main text.

**ORCID iDs**

I Zohar  
https://orcid.org/0000-0002-6278-9807

B Haylock  
https://orcid.org/0000-0002-3925-4484
M J Arshad https://orcid.org/0000-0001-5676-5246
C Bonato https://orcid.org/0000-0003-1550-8483
A Finkler https://orcid.org/0000-0002-8034-4000

References

[1] Crawford S E, Shugayev R A, Paudel H P, Lu P, Syamlal M, Ohodnicki P R, Chorpening B, Gentry R and Duan Y 2021 Adv. Quantum Technol. 4 2100049
[2] van der Sar T, Casola F, Walsworth R and Yacoby A 2015 Nat. Commun. 6 7886
[3] Gross I et al 2017 Nature 549 252–6
[4] Dovzhenko Y, Casola F, Schlotter S, Zhou T X, Büttner F, Walsworth R L, Beach G S D and Yacoby A 2018 Nat. Commun. 9 2712
[5] Jenkins A, Pelliccione M, Yu G, Ma X, Li X, Wang K L and Jayich A C B 2019 Phys. Rev. Mater. 3 083801
[6] Shi F et al 2015 Science 347 1135–8
[7] Lovchinsky I et al 2016 Science 351 836–41
[8] Barry J F, Turner M J, Schloss J M, Glenn D R, Song Y, Lukin M D, Park H and Walsworth R L 2016 Proc. Natl. Acad. Sci. 113 14133–8
[9] Schäfer–Nolte E, Schlipf L, Ternes M, Reinhard F, Kern K and Wrachtrup J 2014 Phys. Rev. Lett. 113 217204
[10] Finkler A and Dasari D 2021 Phys. Rev. Appl. 15 034066
[11] Degen C L, Reinhard F and Cappellaro P 2017 Rev. Mod. Phys. 89 035002
[12] Neumann P et al 2013 Nano Lett. 13 2738–42
[13] Trushheim M E and Englund D 2016 New J. Phys. 18 123023
[14] Ho K O, Wong K C, Leung M Y, Pang Y Y, Leung W K, Yip K Y, Zhang W, Xie J, Goh S K and Yang S 2021 J. Appl. Phys. 129 241101
[15] Higgins B L, Berry D W, Bartlett S D, Wiseman H M and Pryde G J 2007 Nature 450 393–6
[16] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 Phys. Rev. A 54 R4649–52
[17] Vorobyov V, Zaiser S, Abt N, Meinel J, Dasari D, Neumann P and Wrachtrup J 2021 npj Quantum Inf. 7 124
[18] Kitaev A Y 1995 arXiv:quant-ph/9511026
[19] Said R S, Berry D W and Tawam J 2011 Phys. Rev. B 83 125410
[20] Okamoto R, Iefuji M, Oyama S, Yamagata K, Inami H, Fujiwara A and Takeuchi S 2012 Phys. Rev. Lett. 109 130404
[21] Danilin S, Lebedev A V, Vepsäläinen A, Lesovik G B, Blatter G and Paraoanu G S 2018 npj Quantum Inf. 4 29
[22] Kitaev A, Shen A and Vyalii M 2002 Classical and Quantum Computation (Providence, RI: American Mathematical Society)
[23] Bonato C, Blok M S, Dinani H T, Berry D W, Markham M L, Twitchen D J and Hanson R 2015 Nat. Nanotechnol. 11 247–52
[24] Valeri M, Polino E, Poderini D, Gianani I, Corrielli G, Crespi A, Osellame R, Spagnolo N and Sciarrino F 2020 npj Quantum Inf. 6 92
[25] Gebhart V, Santagati R, Gentile A A, Gauger E, Craig D, Ares N, Banchi L, Marquardt F, Pezzé L and Bonato C 2023 Nat. Rev. Phys. 5 141–56
[26] Higgins B L, Berry D W, Bartlett S D, Mitchell M W, Wiseman H M and Pryde G J 2009 New J. Phys. 11 073023
[27] Cappellaro P 2012 Phys. Rev. A 85 030301
[28] Wiebe N and Granade C 2016 Phys. Rev. Lett. 117 010503
[29] Scerri E, Gauger E M and Bonato C 2020 New J. Phys. 22 035002
[30] Santagati R et al 2019 Phys. Rev. Lett. 123 020109
[31] Joas T, Schmitt S, Santagati R, Gentile A A, Bonato C, Laing A, McGuinness L P and Jelezko F 2021 npj Quantum Inf. 7 56
[32] Dinani H T, Berry D W, Gonzalez R, Maze J R and Bonato C 2019 Phys. Rev. B 99 125413
[33] Mehta J R et al 2008 Nature 455 648–51
[34] Balasubramanian G et al 2008 Nature 455 648–51
[35] Childress L, Gurudev Dutt M V, Taylor J M, Zibrov A S, Jelezko F, Wrachtrup J, Hemmer P R and Lukin M D 2006 Science 314 281–5
[36] Robledo L, Childress L, Bernien H, Hensen B, Alkemade P F A and Hanson R 2011 Nature 477 574–8
[37] Staudacher T, Shi F, Pezzagna S, Meijer J, Du J, Meriles C A, Reinhard F and Wrachtrup J 2013 Science 339 561–3
[38] Zohar I 2023 Replication Data for: Real-time frequency estimation of a qubit without single-shot-readout Zenodo (https://doi.org/10.5281/zenodo.7735976)
[39] Momenzadeh S A, Stöhr R J, de Oliveira F F, Brunner A, Denisenko A, Yang S, Reinhard F and Wrachtrup J 2015 Nano Lett. 15 165–9
[40] McMichael R D, Dushenko S and Blakley S M 2021 J. Appl. Phys. 130 144401
[41] Arshad M J, Bekker C, Haylock B, Skrzypczak K, White D, Griffiths B, Gore J, Morley G W, Salter P, Smith J, Zohar I, Finkler A, Altman Y, Gauger E M and Bonato C 2022 arXiv:2210.06103