Sequential Bayesian experiment design for adaptive Ramsey sequence measurements

Robert D. McMichael,[1] Sergey Dushenko,[1, 2] and Sean M. Blakley[1]

[1)] *Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*[a)]

[2)] *Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA*

(Dated: 13 September 2021)

The Ramsey sequence is a canonical example of a quantum phase measurement for a spin qubit. In Ramsey measurements, the measurement efficiency can be optimized through careful selection of settings for the phase accumulation time setting, τ. This paper implements a sequential Bayesian experiment design protocol in low-fidelity Ramsey measurements, and its performance is compared to a previously reported adaptive heuristic protocol, a quantum phase estimation algorithm, and random setting choices. A workflow allowing measurements and design calculations to run concurrently largely eliminates computation time from measurement overhead. When precession frequency is the lone parameter to estimate, the Bayesian design is faster by factors of roughly 2 and 4 and 5 relative to the adaptive heuristic, random τ choices and the quantum phase estimation algorithm respectively. When four parameters are to be determined, Bayesian experiment design and random τ choices can converge to roughy equivalent sensitivity, but the Bayesian method converges 4 times faster.

[a)]Electronic mail: rmcmichael@nist.gov

I. INTRODUCTION

The development of diamond NV centers as measurement tools is one of the most significant achievements of applied quantum physics in recent years. From initial concept[1] and first demonstrations,[2–4] NV-based magnetometry measurements[5] have been developed for condensed-matter physics,[6–10] engineering,[11–13] biology[14–19], nanoscale nuclear magnetic resonance,[20–25] and commercial instrumentation is now becoming available.

The qubit that draws attention to the NV center is the $S = 1$ spin of the electronic ground state, which forms around a nitrogen atom, neighboring vacancy, and trapped electron in a diamond crystal. The coherence time of the spin state can extend into the millisecond range.[26,27] At ambient temperatures, the spin state can be initialized and read out using laser light and detection of emitted photons. With incident green light, an NV center will begin to cycle between its electronic ground state and an excited state, absorbing a photon and relaxing. The key factor is that the relaxation process in NV centers is spin-dependent. Centers excited from the $m_z = 0$ state will relax by emitting a photon and return to the same $m_z = 0$ state, but centers excited from the $m_z = \pm 1$ states can also relax without emitting a photon, and may switch one-way from $m_z = \pm 1$ to $m_z = 0$ in the process. Because readout resets the state to $m_z = 0$, only photons from the first few absorb-relax cycles carry information about the initial spin state.

A persistent problem in NV center measurements is that it is often difficult to efficiently collect these few meaningful, spin-dependent photons. Improvements in collection efficiency have been demonstrated using solid immersion lenses[28], optical resonators,[29,30] and fabricated diamond nanostructures.[31–33] Other characteristics of NV centers have also been successfully exploited. Additional information can be gleaned from the timing of emitted photons[34]. More exotic approaches include spin-dependent ionization of the NV center into different long-lasting charge states where differences in emission spectra can be measured through many excitation-emission cycles.[35–41] Another approach has been to exploit features of a metastable state in the "dark" relaxation path of $m_z = \pm 1$ states.[42,43]

While these methods improve readout fidelity, additional gains can be achieved through efficient measurement design, which allows experiments to measure faster and/or more precisely. Measurement design in this context refers to the choice of experimental settings. Previously, sequential Bayesian experiment design (SBED) produced order-of-magnitude

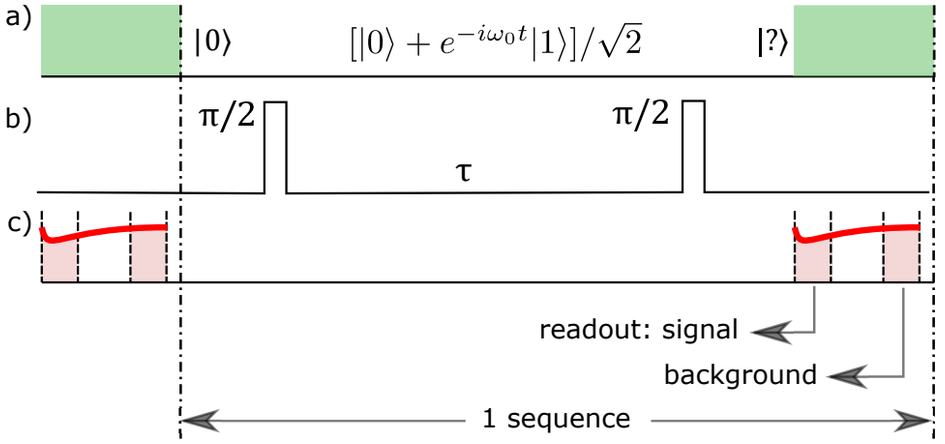


FIG. 1. Timing diagram of Ramsey pulse sequences. Green areas in trace a) denote green laser excitation. In trace b), a microwave $\pi/2$ pulse initiates precession. After a time τ selected by the experiment design, another $\pi/2$ pulse projects the state for readout. In trace c), red curves indicate photoluminescence emission and pink zones show time intervals where emitted photons are counted. Signal photons are counted early in the laser pulse, and background photons are counted after a steady state is reached. The horizontal axis corresponds to elapsed time.

decreases in both the number of measurements and measurement time as compared to swept-frequency measurements in optically detected magnetic resonance of NV centers.[44] In this paper, we apply sequential Bayesian experiment design methods to Ramsey sequence measurements, and we compare results with other adaptive and non-adaptive methods. Section II reviews Ramsey measurements, and section III describes the measurement simulation and introduces the measurement methods. Section IV presents the results of measurement simulations comparing the different protocols, showing substantial improvement of the Bayesian experiment design over the other methods.

II. BACKGROUND

A. Ramsey experiment

The Ramsey sequence is an example of quantum interferometry, and it is the canonical approach to determine the energy difference $\hbar\omega_0$ between two states. When the energy difference depends on a signal, the Ramsey sequence is the basis of quantum sensing.[4,45]

In a Ramsey experiment, a system with $|0\rangle$ and $|1\rangle$ states differing in energy by $\Delta E = \hbar\omega$ is prepared in the $|0\rangle$ and put into a superposition state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ with a $\pi/2$ operator. The system is allowed to propagate for a time τ, yielding a state $|\psi(\tau)\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\omega\tau}|1\rangle)$. After another $\pi/2$ operation, the final state, $|\psi_{\mathrm{f}}(\tau)\rangle = \frac{1}{2}(1 + e^{-i\omega\tau})|0\rangle + \frac{1}{2}(1 - e^{-i\omega\tau})|1\rangle$, can be read out. In the ideal case, measurement of this final state yields a result m ($m = \{0, 1\}$ for $\{|0\rangle, |1\rangle\}$) with conditional probability $P(m|\tau, \omega)$ given τ and ω:

$$P(m|\tau, \omega) = \frac{1}{2}[1 + (-1)^m \cos\omega\tau]. \tag{1}$$

Under conditions of high-fidelity, single-shot readout, quantum phase estimation schemes allow a phase $\phi = \omega\tau_0$ to be determined with precision $\Delta\phi \propto 2^{-M}$ in a total time $\Delta t \propto 2^M$.[46,47] Since $\Delta\phi\Delta t$ is a constant, quantum phase estimation is said to achieve Heisenberg scaling of the precision. Such scaling can be achieved using entangled states, and approximated by adaptive measurement designs[48–52] and non-adaptive designs.[53–56]

However, in ambient-temperature measurements using NV centers, readout fidelity is typically far from ideal, and the measurements have low efficiency, falling in the "averaged readout" category.[45]

The NV center Ramsey sequence is illustrated in figure 1. First, the spin of the NV's electronic ground is initialized to $|0\rangle$ with a few microseconds of green light illumination followed by time to allow optically excited states to relax. Next, the spins are put into a superposition of $|0\rangle$ and $|1\rangle$ states with a calibrated $\pi/2$ microwave pulse. The spin state then evolves for a time τ according to the interactions included in the Hamiltonian. A second microwave pulse projects the phase-shifted state back onto $|0\rangle$ and $|1\rangle$ states for readout.

During readout, maximum spin-state contrast is achieved by counting signal photons during the first 100 ns to 200 ns of green light illumination before the NV center's spin state is reinitialized to $|0\rangle$. After a steady state is reached, during the last 100 ns to 200 ns of the laser pulse, a second count of background photons is collected to monitor and compensate for experimental drift.

III. METHODS

A. Experiment simulation

The measurement process consists of a series of *epochs*, each epoch comprising m_s repeated Ramsey sequences with the same τ setting value. Signal photon counts collected near the beginning of the laser pulse are summed for each epoch, yielding n_s signal photons. Both τ and m_s may be chosen flexibly, but we attempt to keep m_s roughly consistent in order to focus on the effects of τ choices in the different protocols.

Background photon counts are collected near the trailing end of the laser pulse in order to monitor and compensate for changing laser intensity and other slow experimental drifts.

To simulate the count data, we model the shot noise of the n_s and n_b photon counts as samples from Poisson distributions

$$n_\mathrm{s} \sim \mathrm{Pois}(m_\mathrm{s}\lambda_\mathrm{s}(t)) \tag{2}$$

$$n_\mathrm{b} \sim \mathrm{Pois}(m_\mathrm{s}\lambda_\mathrm{b}(t)) \tag{3}$$

where $\lambda_\mathrm{s}(t)$ and $\lambda_\mathrm{b}(t)$ are counts-per-Ramsay-sequence photon rates and the time dependence denotes slow drifts.

To model the rates, we assume that signal and background channels drift with the same slow time dependence, and we define a drift-free ratio, $R(\theta, \tau)$ which depends on $\lambda_\mathrm{b}(t)$, time-independent parameters $\theta \equiv \{a, c, \omega_0, T_2^*\}$, and setting τ.

$$\lambda_\mathrm{s}(t) = R(\theta, \tau)\lambda_\mathrm{b}(t) \tag{4}$$

$$R(\theta, \tau) = a + c\cos(\omega_0\tau)e^{-(\tau/T_2^*)^2}. \tag{5}$$

The parameters include background a, contrast c, precession frequency ω_0, and decoherence time T_2^*.

As a guide to measurement simulations, we use the Ramsey experimental data from Ulm University[57] and the descriptions provided in ref. 58. The experimental overhead includes approximately 3 μs of laser illumination, 1 μs of relaxation time and 70 ns for microwave pulses and photodetector for a total overhead of 4.07 μs per sequence. Allowed values of τ range from 0.1 μs to 20 μs of precession time with 50 ns resolution.

We simulate photon counts using (4) and (5) with "true" parameter values $a = 0.8$, $c = 0.13$, $\omega_0 = 9.4\ \mu\text{s}^{-1}$, and either $T_2^* = 10\ \mu$s or $T_2^* \to \infty$. We use an average background count rate $\lambda_\text{b} = 0.15$ photons per sequence.

B. Bayesian Inference

For all of the experiment designs described below, the simulated data is interpreted using Bayesian inference to refine the probability distribution of the model parameters.

In the i^th epoch, $n_{\text{s},i}$ and $n_{\text{b},i}$ photons are detected and recorded from the signal and background channels, respectively. Then $P_i(\theta|\mathbf{y}_i, \mathbf{d}_i)$ is inferred given measured values $\mathbf{y}_i \equiv \{\mathbf{n}_{\text{s},i}, \mathbf{n}_{\text{b},i}\}$ and setting (design) values $\mathbf{d}_i \equiv \{\boldsymbol{\tau}_i, \mathbf{m}_i\}$. We use bold font to denote recorded values from the current epoch and all preceding epochs, i.e. $\mathbf{n}_{\text{s},i} \equiv \{n_{\text{s},1}, n_{\text{s},2}, \ldots, n_{\text{s},i}\}$.

The distribution of time-independent parameters θ is obtained recursively incorporating data using Bayes' rule. In each epoch, the *posterior* $P(\theta|\mathbf{y}_i, \mathbf{d}_i)$ is proportional to the likelihood and the prior,

$$P(\theta|\mathbf{y}_i, \boldsymbol{\tau}_i) \propto P(n_{\text{s},i}|\theta, d_i, \tilde{n}_{\text{b},i}, \tilde{m}_{\text{b},i}) P(\theta|\mathbf{y}_{i-1}, \mathbf{d}_{i-1}) \tag{6}$$

with likelihood

$$P(n_{\text{s},i}|\theta, \tau_i, \tilde{n}_{\text{b},i}, \tilde{m}_{\text{b},i}) \propto R(\theta, \tau_i)^{n_\text{s}} \times \left[\frac{m_i + \tilde{m}_{\text{b},i}}{m_\text{i} R(\theta, \tau_i) + \tilde{m}_{\text{b},i}}\right]^{n_{\text{s},i}+\tilde{n}_{\text{b},i}}, \tag{7}$$

and the posterior of the preceding epoch, $P(\theta|\mathbf{y}_{i-1}, \mathbf{d}_{i-1})$ as the prior. The first epoch uses $P(\theta|\mathbf{y}_0, \mathbf{d}_0)$, the prior conditioned by no data. The derivation of (7) is given in Appendix A.

Background signal information is included in (7) using sums $\tilde{n}_{\text{b},i}$ and $\tilde{m}_{\text{b},i}$ over only the N_Σ most recent epochs.

$$\tilde{n}_{\text{b},i} = \sum_{j=\max(i-N_\Sigma+1,1)}^{i} n_{\text{b},j}; \tag{8}$$

$$\tilde{m}_{\text{b},i} = \sum_{j=\max(i-N_\Sigma+1,1)}^{i} m_{\text{s},j}. \tag{9}$$

Using finite N_Σ allows $P_i(\lambda_\text{b})$ to forget old data and respond to experimental drift. The simulations presented below use $N_\Sigma = 10$.

C. Experiment Design

The goal of the experiment designs is to choose τ values to make the measurements efficiently. In the simulations, we compare the effectiveness of four designs, including two adaptive protocols: sequential Bayesian experiment design (SBED), and an adaptive heuristic design[58], and two non-adaptive designs: random setting selection and version of the quantum phase estimation algorithm (QPEA)[34].

To motivate the SBED protocol, we first provide a brief overview of optimal Bayesian experiment design (OBED). More detailed descriptions of OBED may be found in the references.[44,59–63] Software and documentation for SBED are available through ref. 63.

In each epoch, optimal Bayesian experiment design calculates which of the candidate τ' settings would, on average, produce the greatest improvement in the parameter distribution relative to the cost of the measurement. To quantify improvement in the parameter distribution, the OBED method uses the information entropy,

$$H_\theta[P] \equiv -\int P(\theta)\log[P(\theta)]d\theta \tag{10}$$

as a metric. We add the subscript to H_θ to identify the entropy of a distribution over θ.

Suppose that a future measurement using setting design d' yields result y'. (Primes denote predicted values to differentiate from unprimed, recorded values.) Bayesian inference would yield a projected new posterior distribution $P'(\theta|y',d')$ from the prior $P(\theta)$. The beneficial information entropy difference between $P(\theta)$ and $P'(\theta|y',d')$ is gauged using Kullback-Leibler divergence,

$$D_{\mathrm{KL}}(y',d') = -\int P'(\theta|y',d')\log\left[\frac{P(\theta)}{P'(\theta|y',d')}\right]d\theta. \tag{11}$$

The utility $U(\tau')$ is then the average D_{KL} benefit,

$$U(d') = \int D_{\mathrm{KL}}(y',d')P(y'|d')dy', \tag{12}$$

The optimal design is the design d' that maximizes $U(d')$.

Application of Bayes rule allows the integrand in (11) to be expressed in terms of distributions over predicted y'. The result is the difference between two entropy-like terms:

$$U(d') = H_{y'|d}(d') - \int H_{y'|\theta,d'}(\theta,d')P(\theta)d\theta, \tag{13}$$

where the first term is the entropy of

$$P(y'|d') = \int P(y'|\theta,d')P(\theta)d\theta, \tag{14}$$

which is the distribution of measurement predictions broadened by the parameter distribution. The second term in (13) is essentially the entropy of measurement noise for fixed θ, averaged over θ.

The double integral prescribed in (13) and (14) is computationally demanding, so for SBED we desire a calculable proxy utility $U^*(d')$ with maxima near the maxima of the information-theoretic $U(d')$. Our proposed proxy utility is motivated by the following plausibility arguments.

We start by looking at an oversimplified case where both $P(y'|\theta, d')$ and $P(y'|d')$ are normal distributions. In particular, $P(y'|\theta, d') \sim \mathcal{N}(y_m(\theta, d'), v_\delta)$ is a normal distribution around model value $y_m(\theta, d')$ with a noise variance v_δ. In the Ramsay experiment, both the mean y_m and the variance v_δ of the Poisson-distributed signal count are equal to $m_\mathrm{s}\lambda_b R(\theta, d')$.

Also in this simple case, the distribution of model values

$$P(y_m|d') = \int P(y_m|\theta, d')P(\theta)d\theta \tag{15}$$

is assumed to be a normal distribution $\mathcal{N}(\langle y_m \rangle, v_\mathrm{m})$ around a mean model value $\langle y_m \rangle$ with variance v_m stemming from the distribution of parameter values.

Under these over-simplifying assumptions, the utility can be expressed as

$$U(d')|_\mathcal{N} = \log\left[\frac{v_\mathrm{m}(d') + v_\delta(d')}{v_\delta(d')}\right], \tag{16}$$

Where the effects of the parameter distribution on model values and the effects of measurement noise are separated into variances $v_\mathrm{m}(d')$ and $v_\delta(d')$ respectively.

In many cases of interest, the measurement noise can be modeled by a normal distribution. However, with nonlinear model functions, $P(\theta)$ and $P(y_\mathrm{m}|d')$ are generally not normal, and we observe in simulation that $P(\theta)$ is often multimodal. Nevertheless, we have found in previous work that SBED protocols based on (16) perform well compared to non-adaptive protocols.[44,63]

In modeling the Ramsay experiment model, however, we found that (16) often performed worse than random setting selection, especially during the first few epochs of a run. This failure motivated us to reintroduce information entropy to gauge the effects of parameter distribution on measurement outcomes, but we keep the form of (16) where measurement noise and parameter distribution effects are separated. For SBED, we propose a proxy utility

$$U^*(d') = \left[\frac{v^*_\mathrm{m}(d') + v_\delta(d')}{v_\delta(d')}\right], \tag{17}$$

with effective variance $v_{\mathrm{m}}^*(d')$ defined as

$$v_{\mathrm{m}}^*(d') = \frac{e^{2H_{y_{\mathrm{m}}}(d')}}{2\pi e}. \tag{18}$$

In this expression, the entropy $H_{y_{\mathrm{m}}}(d')$ of the model value distribution (15) is cast as a variance by inverting an expression for the entropy of a normal distribution. The absence of the log() function in U^* (17) relative to (16) does not affect the maximal-utility d' values because log() is monotonic.

Applied to Ramsay experiment modeling, $y' \equiv m_{\mathrm{s}}\lambda_{\mathrm{b}}R(\theta, d)$ is the signal channel count rate, $\theta \equiv \{a, c, \omega_0, T_2^*\}$ and $\theta \equiv \omega_0$, and $d' \equiv \tau'$. The number of repeats per epoch, m_s is set by the allocated measurement time.

Our computational method for selecting settings is presented in Algorithm 1. We draw $N_s = 100$ parameter samples and calculate corresponding model values y' for the signal emission rate (4) using (5), we then estimate the entropy[64] $H_{y'}(d')$ and compute the effective variance and the proxy utility. We estimate $U^*(d')$ corresponding to all allowed ≈ 400 discrete τ' values, all using the same θ samples. Using a consistent set of θ samples endows $U^*(d')$ with a smoothness, eliminating sampling noise that might obscure the maxima of $U^*(d')$. Finally, we select the d' setting with the greatest $U^*(d')$.

For comparison against SBED, we also consider an adaptive design for Ramsey experiments reported by Santagati et al.[58] This heuristic design, "Tau" is attractively simple, with the τ setting determined by $\tau = h/\sigma_{\omega_0}$, where σ_{ω_0} is the width of the frequency distribution and $h \approx 1$ is a tuning parameter. Empirically, we found that a value of $h \approx 0.5$ generated the best results in our simulations. A scaling argument (see Appendix B) suggests that the Tau protocol might show $1/T$ scaling, but the reported behavior appears to be closer to exponential decay.[58] In Tau runs, σ_{ω_0} eventually becomes small enough that $1/(2\sigma_{\omega_0}) > 20$ μs is outside the allowed setting range. In this case, τ is selected randomly from the top 10 % of τ values.

Third, we consider a non-adaptive protocol, "Random," where τ is selected randomly from the allowed settings. This protocol provides a baseline, non-adaptive case for comparison with the adaptive protocols.

Finally, we test a non-adaptive design, "QPEA," which incorporates many of the features of the quantum phase estimation algorithms,[46–56] but follows Dinani et al. in application of these methods to a low-fidelity readout.[34]

Algorithm 1 Pseudocode for SBED setting selection.

Input: Parameter distribution $P(\theta)$, candidate designs d_i

```
for j ← 1 ... N_s do
    θ_j ← sample from P(θ)
end for
for i ← 1 ... N_d do                          ▷ All candidate settings d_i
    for j ← 1 ... N_s do                      ▷ Reuse θ samples
        y_j ← model(θ_j, d_i)                 ▷ Evaluate model function
    end for
    H_i ← H({y_1 ... y_{N_s}})                ▷ Estimate entropy of P(y)
    v_i ← e^{2h_i}/(2πe)                      ▷ Effective variance
    U_i^* ← (v_i + v_δ)/v_δ
end for
i_best ← argmax(U_i^*)                        ▷ Choose max utility setting
```

Output: $d_{i_{\text{best}}}$ setting

In the QPEA design, the experimental model (5) as a function of parameters θ is extended to include a readout phase ψ in addition to the precession time τ.

$$R(\theta, \tau, \psi) = a + c\cos(\omega_0\tau + \psi)e^{-(\tau/T_2^*)^2} \tag{19}$$

The precession time takes on values $\tau = \tau_0 2^k$ for $k = 0, 1, ...K$, and measurements are repeated M_k times with ψ values at increments of π/M_k[34,53]. In our implementation, we create a randomly scrambled sequence of all (τ, ψ) setting combinations, and then repeat that sequence as needed, simulating $m_{\text{s}} = 500$ repeats for each setting combination. The number of measurements allocated to each τ value is given by

$$M_k = G + F * (K - k), \tag{20}$$

where G and $G + F$ measurements are allocated to the largest and smallest τ values, respectively.[34,52–56]. In its original context, each τ value contributes one binary bit to the result. The M_k formula above allocates additional measurements to more significant bits

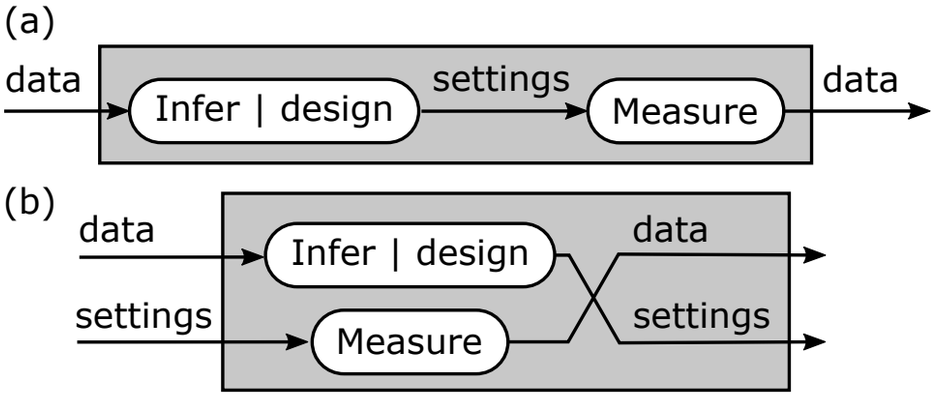


FIG. 2. Flow diagrams of computation and measurement using adaptive experiment design. Grey blocks indicate single measurement epochs. In a series workflow (a), data collection stops for data analysis and new setting decisions. In a concurrent workflow (b), measurements continue until new settings are ready.

where errors have more severe consequences.[53] In the simulations reported below, we used $G = 25$ and $F = 1$ with $K = 8$ and $\tau_0 = (20\,\mu\text{s})/2^8$.

To compare these four designs/protools, we value elapsed time as the resource to be allocated. The exception is QPEA where measurement repeats are allocated according to M_k. Accordingly, measurement efficiency is judged on the basis of elapsed time. Therefore, in Tau and Random runs, a consistent time interval is allocated to each epoch, and the Ramsey sequence is repeated with the selected τ until the time is spent. In the Bayes runs, the measurement time is determined by computation time, as described in the following subsection.

D. Concurrent design and measurement

The computational time t_{calc} demanded by SBED may be the method's greatest disadvantage. In a sequential measure-infer-design loop as illustrated in fig. 2(a), the measurement waits for instructions during the design calculations. This idle time adds to the experimental overhead and degrades the overall efficiency of the protocol. In the calculations for this paper, SBED computation required a few milliseconds per epoch. In the worse case of one sequence per epoch, a few milliseconds would be an enormous overhead compared to a Ramsey τ of a few microseconds.

To avoid adding computation time to overhead, we propose a concurrent workflow where measurements and design calculations run at the same time (See fig. 2(b)). In epoch i, the measurement process accumulates and averages data y_i using a setting d_i designed in the previous iteration $i-1$. Meanwhile, Bayesian regression incorporates the previous epoch's measurement data y_{i-1} and chooses settings d_{i+1} for the next iteration. After the calculations are completed, the measurement system reports data and receives the next setting design. By running measurements and calculations concurrently, the measurement can continue to collect data virtually nonstop. The fact that d_{i+1} is based off accumulated data $\{y_0, \ldots, y_{i-1}\}$, i.e. not including the current measurements y_i has negligible effects in measurements that iterate for many epochs.

We also suggest that for adaptive measurement schemes, the computation time is more appropriately compared to the time t_{SNR} required to reach a signal to noise ratio (SNR) ≈ 1. We propose that the infer/design calculations are fast enough if they can generate designs in a time comparable to the time required for measurement data to significantly change the parameter distribution. The parameter distribution only changes significantly following measurements with SNR $\gtrsim 1$ or equivalent. Using shorter measurement times with fewer repetitions would produce incremental changes in the parameter distribution between computations. At the other extreme, many more repetitions between calculations might hurt efficiency by missing opportunities to switch to higher-utility settings.

In our simulations, t_{SNR} can be estimated. Noise standard deviation equal to the contrast of $c = 0.13$ on a background of $a = 0.8$ would require ≈ 40 signal photons, or $m_{\mathrm{s}} \approx 300$ repeats, or about 4.2 ms at the median setting $\tau = 10$ μs. By coincidence, in the Bayes runs, the mean measurement time was 4.4 ms, which is comparable to t_{SNR}.

IV. RESULTS AND DISCUSSION

Here we test sequential Bayesian experiment design (SBED) against the adaptive heuristic method of Santagati et al.,[58] (Tau) Random selection of settings (Random), and a quantum phase estimation algorightm (QPEA).

Figure 3 displays statistics from 100-run batches of simulated Ramsey measurements where the precession frequency ω_0 is the only unknown parameter, and $T_2^* \to \infty$ makes decoherence unimportant. The measurement time in each epoch is roughly equivalent. In

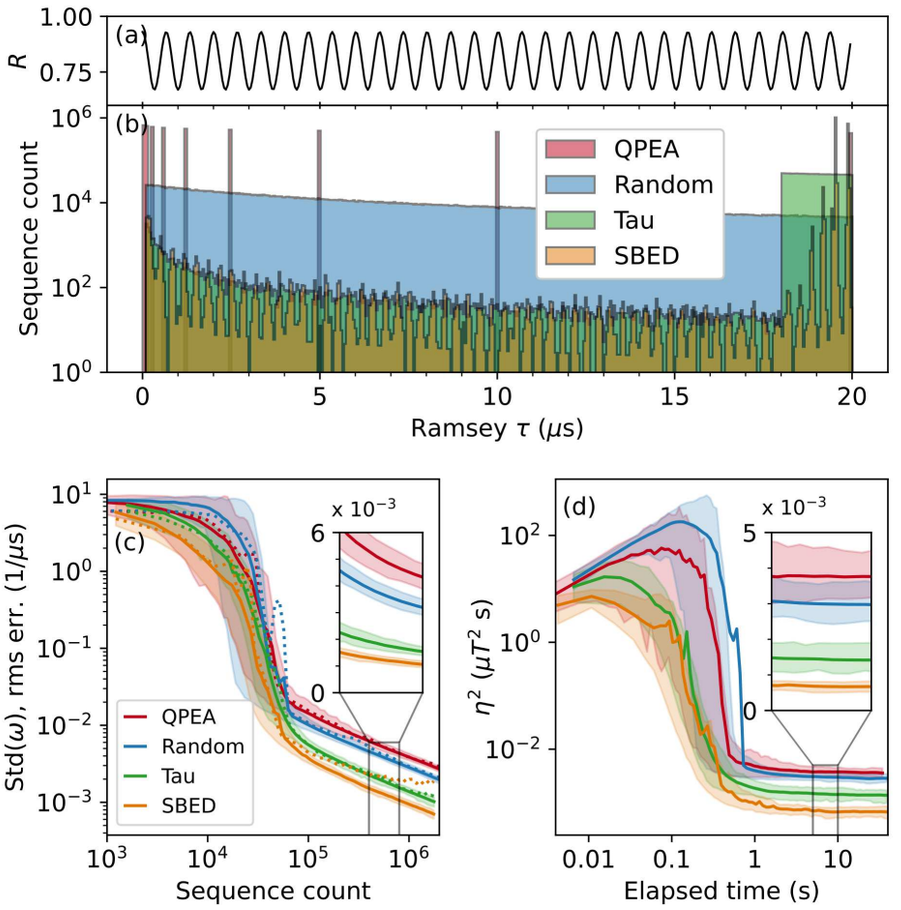


FIG. 3. Measurement simulations where only ω_0 is treated as an unknown. (a) True R curve plotted as a function of τ with $a = 0.8$, $c = 0.13$, $\omega_0 = 9.4$ μs^{-1} and $T_2^* \to \infty$ (b) Histograms comparing the number of times each τ setting was used chosen by four protocols. QPEA and Random are non-adaptive and Tau and SBED are adaptive protocols. (c) Evolution of mean σ_{ω_0}. The Bayesian protocol outperforms the heuristic protocol, and the random protocol produces roughly double the uncertainty. Shaded areas are 90 % credibility regions. Dotted lines are standard deviation of error. (d) Evolution of absolute sensitivity vs. elapsed "wall clock" time. The SBED protocol is roughly twice as fast as the Tau protocol, four times as fast as Random and five times as fast as QPEA. Statistics are calculated over 100 individual runs.

the QPEA simulations, each epoch repeats (τ, ψ) settings 500 times. The average time per epoch was 3.9 ms. Tau and Random runs are allocated 4 ms of measurement time per epoch, i.e. $m_s = 4$ ms$/(\tau + 4.07$ μs$)$ repeats of the Ramsey sequence. SBED measurement simulations mimic a concurrent workflow, using the execution times of the inference and

experiment design code to determine the measurement times. In these simulations, the mean epoch time is 4.4 ms.

The histograms in fig. 3(b) provide a comparison of how frequently the various protocols use different τ settings. The QPEA protocol places a very strong emphasis on small τ; in fact the median τ is 0.625 μs. The Random protocol also emphasizes low τ by virtue of the fact that measurement time is allocated uniformly, and more repeats are possible when τ is small.

The adaptive Tau and SBED protocols produce similar overall distributions of setting choices, shown in fig. 3(b). These protocols both exhibit an initial emphasis on small τ, later moving to large τ. A striking feature of the Bayes histogram of selected τ values is the comb-like structure indicating that the Bayesian method adaptively concentrates τ values where the model true curve (a) has maximum slope. By comparison, the Tau and Random protocols show no such phase selectivity, although such selectivity could be programmed into an improved heuristic.

The standard deviation of the ω distributions are plotted in fig. 3(c) as a function of the number of Ramsey sequences. Above $\approx 10^5$ sequences, all four protocols converge to a $(\Sigma m_\mathrm{s})^{-1/2}$ scaling. The inset replots the boxed region on linear scales over a doubling of the sequence count.

Figure 3(d) provides a comparison of the absolute sensitivity η of the Ramsey ω_0 measurement used as a magnetometer. The uncertainties in field and frequency, σ_B and σ_{ω_0} respectively, are related by $\sigma_B = \sigma_{\omega_0}/\gamma$, where $\gamma = 2\pi\, 28$ GHz/T is the gyromagnetic ratio. For uncertainties that scale as $t_\mathrm{lab}^{-1/2}$, $\eta^2 \equiv \sigma_B^2 t_\mathrm{lab}$ is a constant. All three protocols produce nearly constant η^2 after about 1 s of laboratory time. The inset replots the boxed region on linear scales. To achieve equivalent uncertainties, the Bayes protocol is about twice as fast as Tau, four times as fast as Random, and five times as fast as QPEA.

At several places, in figure 3(c) and 3(d), and in figure 4 below, the plotted mean value lines fall outside the shaded 90% credibility region. In a typical run, the standard deviation will vary slowly and then rapidly fall orders of magnitude toward the asymptotic value. In a batch of 100 runs, there will be instances where most of the runs will have converged, while a few remain with high standard deviation. In these instances, the mean is heavily influenced by a few, very large outlier values that fall outside the credibility interval.

Fig. 4 presents results from 100-run batches of measurement simulations where a, c, ω_0,

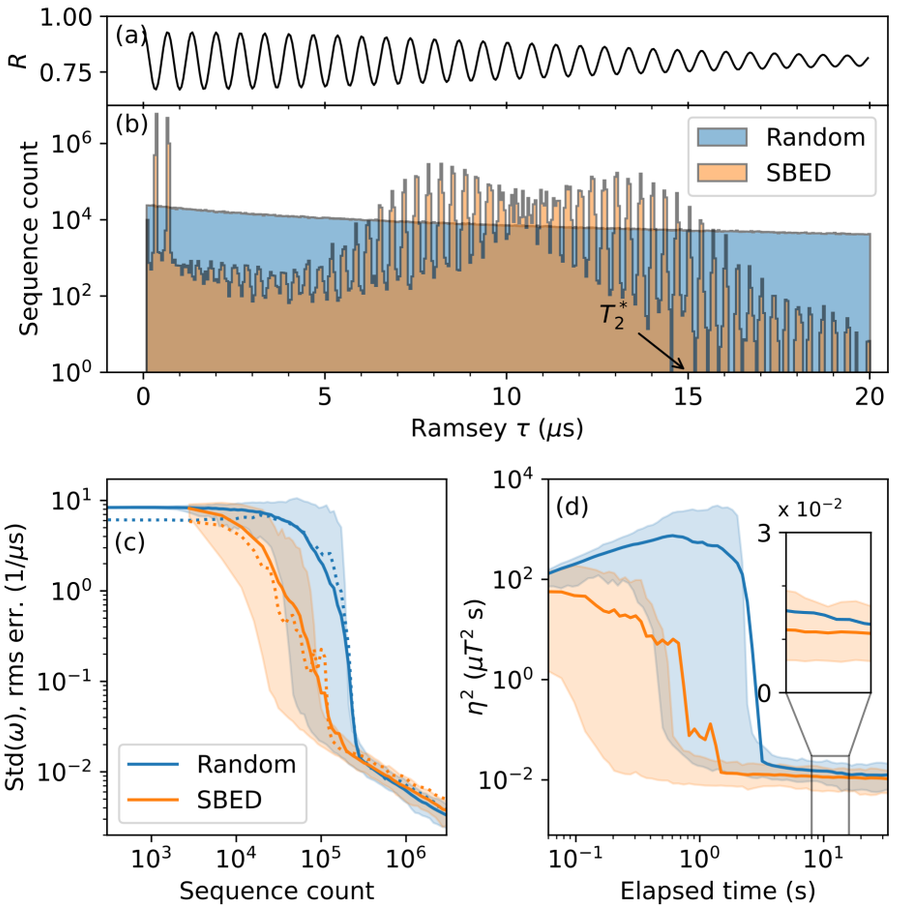


FIG. 4. Simulations of measurements to estimate ω_0, a, c and T_2^*. (a) True R curve plotted as a function of τ with $a = 0.8$, $c = 0.13$, $\omega_0 = 9.4$ μs^{-1} and $T_2^* = 15$ μs. (b) Histograms of τ setting values determined by Random selection and by SBED protocols. (c) Uncertainty in frequency, σ_{ω_0} vs. Ramsey sequence count. (d) Absolute uncertainty η^2 for SBED and random selections of τ. The Bayesian method uses concurrent measurement and calculation. In the random design data, the time required for data analysis has been discounted. Statistics are calculated over 100 individual runs.

and T_2^* are all treated as unknowns. We compare only SBED and Random protocols, as no heuristic is available and QPEA is designed for frequency/phase determination only. As above, 4 ms were allotted for measurements in the Random protocol, but SBED code execution was slower with 4 variables, and an average of 13 ms were allocated per epoch in Bayes protocol runs.

Compared to the histogram in fig. 3(b), the SBED histogram structure in fig. 4(b) is

much richer. For a qualitative interpretation of this structure, we note that the modeled signal given by (5) is most sensitive to contrast c at small τ extrema, is most sensitive to T_2^* for extrema near $\tau \approx T_2^*(= 15\ \mu\text{s})$ and is most sensitive to ω_0 for large slope of the model function and $\tau \approx T_2^*/\sqrt{2}$. The maxima of the envelope are skewed to slightly smaller τ values by the fact that low-τ designs produce more photons per epoch, and deliver smaller uncertainty.

Presumably, an adaptive heuristic could be developed and tuned for this four-variable measurement. It is likely that the development would require significant labor, however. Implementation of the Bayes protocol was comparatively simple, requiring only minor adjustments to include a, c and T_2^* as unknowns in the model function.

Figs. 4(c)-(d) show that the Random protocol is slower to converge than the Bayes protocol, but the contrast in asymptotic performance is weaker than in the single-unknown case above. We attribute the apparent efficiency of the random approach to the fact high-utility τ values selected by SBED are widely distributed among the available settings, not tightly grouped at large τ as they appear in fig. 3(b). Our selection of T_2^* =15 μs in a 20 μs τ is close to optimal for the random protocol, artificially permitting τ selections to cover the same τ range that SBED protocol chooses adaptively. In other calculations using T_2^* =10 μs, (not shown) the SBED protocol self-limits to τ values less than 14 μs, where the signal is strong. The resulting OBED measurement is roughly 2 times faster than random in this case.

V. SUMMARY

The simulation results show that sequential Bayesian experiment design is an efficient protocol for single-NV Ramsey measurements, outperforming a tuned heuristic protocol, Tau, random setting selection, and the established QPEA method in efficient use of laboratory, "wall-clock" time. A key factor for the Bayes protocol is the introduction of a concurrent workflow, which allows measurements to continue until design calculations are complete, and effectively eliminates computation time from the overhead, at least when computation time is less than the time needed to attain SNR ≈ 1. Longer measurements might reduce efficiency through missed opportunities to select higher-utility settings.

The efficiency gains produced by sequential Bayesian experiment design in this work are significant, but modest compared to the order-of-magnitude gains previously reported

for measurements of peaks on a wide, flat background.[44] We explain this difference noting that measurements in the broad background of a spectrum offer little compared to the information-rich settings near the peaks, so large gains are possible by focusing measurements on the peak regions. In contrast, the Ramsey measurements tend to be informative over a larger range of setting values.

While this paper has focused on efficiency during data acquisition, sequential Bayesian experiment design also offers efficiency advantages in the time periods before and after data acquisition. For development of a protocol before measurement, the Bayes methods are easily adapted to new experiment models or to different measurement goals, especially when compared to the demands of designing efficient heuristics. Even in cases where Bayes calculations would be prohibitively slow, the Bayes protocols may serve as a reliable guide for heuristic protocol development. Also, if post-measurement data analysis is required, it might be reasonable to count analysis time as part of measurement overhead. In the Random protocol results above, data was analyzed at each epoch to show progress but the analysis time was discounted. But in a more typical use, a quick protocol would yield raw data, and any required data analysis might contribute to overhead. The Bayes protocol offers (almost) instant results because the data is continuously analyzed as part of the protocol.

VI. DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

S.D. acknowledges support under the cooperative research agreement between the University of Maryland and the National Institute of Standards and Technology Physical Measurement Laboratory (Grant No. 70NANB14H209) through the University of Maryland. The authors thank Adam Pintar and Hossein T. Dinani for many helpful discussions.

Appendix A: Derivation of likelihood

This appendix provides a derivation of the likelihood, $P(n_{\mathrm{s},i}|\theta, d_i, \tilde{n}_{\mathrm{b},i}, \tilde{m}_{\mathrm{b},i})$ of receiving $n_{\mathrm{s},i}$ signal counts given parameters θ, settings $d_i = \{\tau_i, m_i\}$ with precession time τ_i and $m_{\mathrm{s},}$ repeats. Background channel counts are incorporated via $\tilde{n}_{\mathrm{b},i}$ and $\tilde{m}_{\mathrm{b},i}$ given in (8) and (9) respectively.

$$P(\theta|\mathbf{y}_i, \mathbf{d}_i) \propto P(n_{\mathrm{s},i}|\theta, d_i, \tilde{n}_{\mathrm{b},i}, \tilde{m}_{\mathrm{b},i}) P(\theta|\mathbf{y}_{i-1}, \mathbf{d}_{i-1}). \tag{A1}$$

where $P(\theta)$ incorporates n and m data from all preceding epochs. The derivation expresses the likelihood of n_{b} counts in m_{b} repeats as a function of a signal count rate λ_{s}, which is the product of the ratio $R(\theta, \tau)$ and the background count λ_{b}.

The distribution $P(\lambda_{\mathrm{b}})$ is determined from n_{b} and m_{b}, then eliminated from the final expression by integration. To simplify notation, we suppress the explicit τ dependence in the following.

The likelihood of receiving n_{s} signal counts in m_{s} repeats is a Poisson distribution with mean $m_{\mathrm{s}} R(\theta) \lambda_{\mathrm{b}}$.

$$P(n_{\mathrm{s}}|\theta, m_{\mathrm{s}}, \lambda_{\mathrm{b}}) = \frac{(m_{\mathrm{s}} R(\theta) \lambda_{\mathrm{b}})^{n_{\mathrm{s}}} e^{-m_{\mathrm{s}} R(\theta) \lambda_{\mathrm{b}}}}{n_{\mathrm{s}}!}. \tag{A2}$$

The background count rate distribution is determined by

$\tilde{n}_{\mathrm{b}}$ and $\tilde{m}_{\mathrm{b}}$

$$P(\lambda_{\mathrm{b}}) \propto \frac{(\tilde{m}_{\mathrm{b}} \lambda_{\mathrm{b}})^{\tilde{n}_{\mathrm{b}}} e^{-\tilde{m}_{\mathrm{b}} \lambda_{\mathrm{b}}}}{\tilde{n}_{\mathrm{b}}!} \lambda_{\mathrm{b}}^{\nu}. \tag{A3}$$

The fraction in this expression is a Poisson distribution expressing the likelihood of counting $\tilde{n}_{\mathrm{b}}$ photons. With $\tilde{m}_{\mathrm{b}} = 0$ experiments and $\tilde{n}_{\mathrm{b}} = 0$ photons collected, the trailing $\lambda_{\mathrm{b}}^{\nu}$ can be viewed as a prior with exponent ν to be determined later. For the Jeffreys prior, $\nu = -1/2$.

The distribution of λ_{b} values is incorporated into the likelihood of n_{s} by integration.

$$P(n_{\mathrm{s}}|\theta, m_{\mathrm{s}}, \tilde{n}_{\mathrm{b}}, \tilde{m}_{\mathrm{b}}) = \int P(n_{\mathrm{s}}|\theta, m_{\mathrm{s}}, \lambda_{\mathrm{b}}) P(\lambda_{\mathrm{b}}|\tilde{n}_{\mathrm{b}}, \tilde{m}_{\mathrm{b}}) d\lambda_{\mathrm{b}} \tag{A4}$$

Substituting from (A2) and (A3), the integral is tractable, yielding

$$P(n_{\mathrm{s}}|\theta, m_{\mathrm{s}}, \tilde{n}_{\mathrm{b}}, \tilde{m}_{\mathrm{b}}) = C \left[\frac{R(\theta)^{n_{\mathrm{s}}}}{m_{\mathrm{s}} R(\theta) + \tilde{m}_{\mathrm{b}}} \right]^{n_{\mathrm{s}} + \tilde{n}_{\mathrm{b}} + 1 + \nu}. \tag{A5}$$

The prior exponent ν is selected so that the net effect of two likelihoods from two identical measurement results is the same as a single likelihood with combined data. That is, we require

$$P(n_\mathrm{s}|\theta, m_\mathrm{s}, \tilde{n}_\mathrm{b}, \tilde{m}_\mathrm{b})^2 \propto P(2n_\mathrm{s}|\theta, 2m_\mathrm{s}, 2\tilde{n}_\mathrm{b}, 2\tilde{m}_\mathrm{b}), \tag{A6}$$

which is satisfied by $\nu = -1$.

Testing also confirmed $\nu = -1$. Using $\nu = 0$, the simulations converged to incorrect mean parameter values. With $\nu = -1$ the simulations regularly converged to the true values within a standard deviation.

We also note that with the choice of $\nu = -1$ (A3) is a gamma distribution of the general form

$$\lambda_\mathrm{b} \sim \Gamma(\alpha, \beta) = \frac{\beta^\alpha \lambda_\mathrm{b}^{\alpha-1} e^{-\beta\lambda_\mathrm{b}}}{\Gamma(\alpha)}, \tag{A7}$$

with $\alpha \to 0$ and $\beta \to 0$ and where $\Gamma(x)$ is the gamma function. Gamma distributions are conjugate priors for Poisson likelihoods, and in the case $\alpha \to 0$ and $\beta \to 0$, the gamma distribution becomes "vague" as its variance $\to \infty$.

The factor C in (A5) contains exponential functions $m_\mathrm{s}!$, $\tilde{m}_b!$, and $(m_\mathrm{s} + \tilde{m}_\mathrm{b})!$, which are computationally challenging for large m values. Since the θ distribution is explicitly normalized in software, we are free to choose a convenient pseudo normalization for the likelihood provided that the θ dependence is preserved. To eliminate factorials we choose the following pseudo-normalization,

$$P(n_\mathrm{s}|\theta, m_\mathrm{s}, \tilde{n}_\mathrm{b}, \tilde{m}_\mathrm{b}) \propto R(\theta)^{n_\mathrm{s}} \left[\frac{m_\mathrm{s} + \tilde{m}_\mathrm{b}}{m_\mathrm{s}R(\theta) + m_\mathrm{b}}\right]^{n_\mathrm{s}+\tilde{n}_\mathrm{b}}, \tag{A8}$$

where C has been replaced by 1 and the θ-independent numerator $m_\mathrm{s} + \tilde{m}_\mathrm{b}$ has been inserted to ensure that the fraction in square brackets is of order unity.

Figure 5 illustrates the performance and behavior (7) or (A8) for different numbers of background measurements, m_b. The inset shows that the likelihood is a peaked function that approaches a Poisson distribution as $\tilde{m}_\mathrm{b}$ increases. The main plot shows the effect of the $\tilde{m}_\mathrm{b}$ on the performance of the SBED method. Only marginal improvements are gained by extending background measurements beyond $n_\mathrm{b} \gtrsim 10 \times n_\mathrm{s}$.

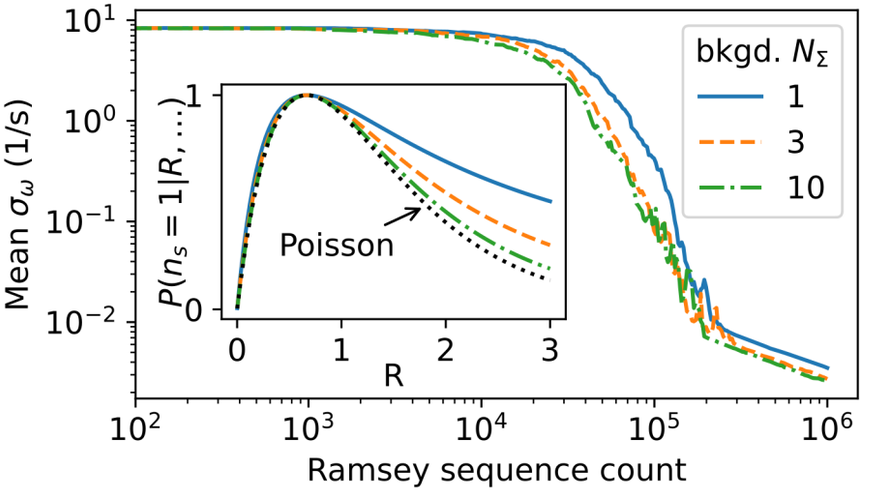


FIG. 5. Influence of background count averaging on the performance of the Bayesian protocol. The inset plots the likelihood of $n_\mathrm{s} = 1$ as a function of R for $m = 10$ repeats per epoch, background counts summed over N_Σ epochs, and $\lambda_\mathrm{b} = 0.15$. The line labeled "Poisson" is proportional to the Poisson distribution for $n_\mathrm{s} = 1$ assuming $\lambda_\mathrm{s} = R\lambda_\mathrm{b}$. In both the main figure and inset, the benefits of background averaging saturate for $N_\Sigma \gtrsim 10$.

Appendix B: Scaling of Tau protocol

Each epoch begins with a consistent phase uncertainty $\sigma_\phi = \tau\sigma_{\omega_0} = h$. If the number of repeats per epoch is constant, each epoch decreases uncertainty by a constant factor $\beta < 1$ on average. After the k^{th} epoch,

$$\sigma \propto \beta^k \tag{B1}$$

Each epoch requires measurement time proportional to τ, neglecting overhead, so total measurement time T scales as

$$T \propto \beta^{-(k+1)} \tag{B2}$$

The net behavior is therefore predicted to follow Heisenberg scaling,

$$\sigma \propto 1/T, \tag{B3}$$

under the assumption that the number of repeats per epoch is constant.

REFERENCES

[1]B. M. Chernobrod and G. P. Berman, Journal of applied physics **97**, 014903 (2005).

[2]G. Balasubramanian, I. Y. Chan, R. Kolesov, M. Al-Hmoud, J. Tisler, C. Shin, C. Kim, A. Wojcik, P. R. Hemmer, A. Krueger, T. Hanke, A. Leitenstorfer, R. Bratschitsch, F. Jelezko, and J. Wrachtrup, Nature **455**, 648 (2008).

[3]J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. G. Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin, Nature **455**, 644 (2008).

[4]J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin, Nature Physics **4**, 810 (2008).

[5]L. Rondin, J.-P. Tetienne, T. Hingant, J.-F. Roch, P. Maletinsky, and V. Jacques, Reports on Progress in Physics **77**, 056503 (2014).

[6]T. Hingant, J.-P. Tetienne, L. J. Martínez, K. Garcia, D. Ravelosona, J.-F. Roch, and V. Jacques, Physical Review Applied **4**, 014003 (2015).

[7]T. van der Sar, F. Casola, R. Walsworth, and A. Yacoby, Nature Communications **6**, 7886 (2015).

[8]J.-P. Tetienne, T. Hingant, J.-V. Kim, L. H. Diez, J.-P. Adam, K. Garcia, J.-F. Roch, S. Rohart, A. Thiaville, D. Ravelosona, and V. Jacques, Science **344**, 1366 (2014).

[9]J.-P. Tetienne, N. Dontschuk, D. A. Broadway, A. Stacey, D. A. Simpson, and L. C. L. Hollenberg, Science Advances **3**, e1602429 (2017).

[10]L. Rondin, J. P. Tetienne, S. Rohart, A. Thiaville, T. Hingant, P. Spinicelli, J. F. Roch, and V. Jacques, Nature Communications **4**, 2279 (2013).

[11]M. J. Turner, N. Langellier, R. Bainbridge, D. Walters, S. Meesala, T. M. Babinec, P. Kehayias, A. Yacoby, E. Hu, M. Lončar, R. L. Walsworth, and E. V. Levine, Physical Review Applied **14**, 014097 (2020).

[12]A. Horsley, P. Appel, J. Wolters, J. Achard, A. Tallaire, P. Maletinsky, and P. Treutlein, Physical Review Applied **10**, 044039 (2018).

[13]P. Appel, M. Ganzhorn, E. Neu, and P. Maletinsky, New Journal of Physics **17**, 112001 (2015).

[14]G. Kucsko, P. C. Maurer, N. Y. Yao, M. Kubo, H. J. Noh, P. K. Lo, H. Park, and M. D. Lukin, Nature **500**, 54 (2013).

[15]L. P. McGuinness, Y. Yan, A. Stacey, D. A. Simpson, L. T. Hall, D. Maclaurin, S. Prawer, P. Mulvaney, J. Wrachtrup, F. Caruso, R. E. Scholten, and L. C. L. Hollenberg, Nature Nanotechnology **6**, 358 (2011).

[16]J. M. McCoey, M. Matsuoka, R. W. Gille, L. T. Hall, J. A. Shaw, J. Tetienne, D. Kisailus, L. C. L. Hollenberg, and D. A. Simpson, Small Methods **4**, 1900754 (2020).

[17]F. C. Ziem, N. S. Götz, A. Zappe, S. Steinert, and J. Wrachtrup, Nano Letters **13**, 4093 (2013).

[18]A. Ermakova, G. Pramanik, J.-M. Cai, G. Algara-Siller, U. Kaiser, T. Weil, Y.-K. Tzeng, H. C. Chang, L. P. McGuinness, M. B. Plenio, B. Naydenov, and F. Jelezko, Nano Letters **13**, 3305 (2013).

[19]J. F. Barry, M. J. Turner, J. M. Schloss, D. R. Glenn, Y. Song, M. D. Lukin, H. Park, and R. L. Walsworth, Proceedings of the National Academy of Sciences **113**, 14133 (2016).

[20]S. J. DeVience, L. M. Pham, I. Lovchinsky, A. O. Sushkov, N. Bar-Gill, C. Belthangady, F. Casola, M. Corbett, H. Zhang, M. Lukin, H. Park, A. Yacoby, and R. L. Walsworth, Nature Nanotechnology **10**, 129 (2015).

[21]D. Rugar, H. J. Mamin, M. H. Sherwood, M. Kim, C. T. Rettner, K. Ohno, and D. D. Awschalom, Nature Nanotechnology **10**, 120 (2014).

[22]T. Häberle, D. Schmid-Lorch, F. Reinhard, and J. Wrachtrup, Nature Nanotechnology **10**, 125 (2015).

[23]H. J. Mamin, M. Kim, M. H. Sherwood, C. T. Rettner, K. Ohno, D. D. Awschalom, and D. Rugar, Science **339**, 557 (2013).

[24]V. P. Bhallamudi and P. C. Hammel, Nature nanotechnology **10**, 104 (2015).

[25]T. Staudacher, F. Shi, S. Pezzagna, J. Meijer, J. Du, C. A. Meriles, F. Reinhard, and J. Wrachtrup, Science **339**, 561 (2013).

[26]K. D. Jahnke, B. Naydenov, T. Teraji, S. Koizumi, T. Umeda, J. Isoya, and F. Jelezko, Applied Physics Letters **101**, 012405 (2012).

[27]G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, Nature Materials **8**, 383 (2009).

[28]J. P. Hadden, J. P. Harrison, A. C. Stanley-Clarke, L. Marseglia, Y.-L. D. Ho, B. R. Patton, J. L. O'Brien, and J. G. Rarity, Applied Physics Letters **97**, 241901 (2010).

[29]K. Jensen, N. Leefer, A. Jarmola, Y. Dumeige, V. Acosta, P. Kehayias, B. Patton, and D. Budker, Physical Review Letters **112**, 160802 (2014).

[30]J. Riedrich-Möller, S. Pezzagna, J. Meijer, C. Pauly, F. Mücklich, M. Markham, A. M. Edmonds, and C. Becher, Applied Physics Letters **106**, 221103 (2015).

[31]T. M. Babinec, B. J. M. Hausmann, M. Khan, Y. Zhang, J. R. Maze, P. R. Hemmer, and M. Lončar, Nature Nanotechnology **5**, 195 (2010).

[32]P. Maletinsky, S. Hong, M. S. Grinolds, B. Hausmann, M. D. Lukin, R. L. Walsworth, M. Loncar, and A. Yacoby, Nature Nanotechnology **7**, 320 (2012).

[33]M. Pelliccione, A. Jenkins, P. Ovartchaiyapong, C. Reetz, E. Emmanouilidou, N. Ni, and A. C. Bleszynski Jayich, Nature Nanotechnology **11**, 700 (2016).

[34]H. T. Dinani, D. W. Berry, R. Gonzalez, J. R. Maze, and C. Bonato, Physical Review B **99**, 125413 (2019).

[35]G. Waldherr, P. Neumann, S. F. Huelga, F. Jelezko, and J. Wrachtrup, Physical Review Letters **107**, 090401 (2011).

[36]N. Aslam, G. Waldherr, P. Neumann, F. Jelezko, and J. Wrachtrup, New Journal of Physics **15**, 013064 (2013).

[37]B. Shields, Q. Unterreithmeier, N. de Leon, H. Park, and M. Lukin, Physical Review Letters **114**, 136402 (2015).

[38]D. Bluvstein, Z. Zhang, and A. C. B. Jayich, Physical Review Letters **122**, 076101 (2019).

[39]J.-C. Jaskula, B. Shields, E. Bauch, M. Lukin, A. Trifonov, and R. Walsworth, Physical Review Applied **11**, 064003 (2019).

[40]L. Hacquebard and L. Childress, Physical Review A **97**, 063408 (2018).

[41]D. A. Hopper, R. R. Grote, A. L. Exarhos, and L. C. Bassett, Physical Review B **94**, 241201 (2016).

[42]V. M. Acosta, A. Jarmola, E. Bauch, and D. Budker, Physical Review B **82**, 201202 (2010).

[43]V. M. Acosta, E. Bauch, A. Jarmola, L. J. Zipp, M. P. Ledbetter, and D. Budker, Applied Physics Letters **97**, 174104 (2010).

[44]S. Dushenko, K. Ambal, and R. D. McMichael, Physical Review Applied **14**, 054036 (2020).

[45]C. Degen, F. Reinhard, and P. Cappellaro, Reviews of Modern Physics **89**, 035002 (2017).

[46]A. Y. Kitaev, arXiv:quant-ph/9511026 (1995), arXiv: quant-ph/9511026.

[47]A. Y. Kitaev, in *Electr. Coll. Comput. Complex.* (1996) pp. TR96–003.

[48]B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature **450**, 393 (2007).

[49]D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, Physical Review A **80**, 052114 (2009).

[50]R. S. Said, D. W. Berry, and J. Twamley, Physical Review B **83**, 125410 (2011).

[51]P. Cappellaro, Physical Review A **85**, 030301 (2012).

[52]C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, Nature Nanotechnology **11**, 247 (2016).

[53]B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, New Journal of Physics **11**, 073023 (2009).

[54]N. M. Nusran, M. U. Momeen, and M. V. G. Dutt, Nature Nanotechnology **7**, 109 (2012).

[55]G. Waldherr, J. Beck, P. Neumann, R. S. Said, M. Nitsche, M. L. Markham, D. J. Twitchen, J. Twamley, F. Jelezko, and J. Wrachtrup, Nature Nanotechnology **7**, 105 (2012).

[56]A. J. F. Hayes and D. W. Berry, Physical Review A **89**, 013838 (2014).

[57]A. Gentile, R. Santagati, S. Schmitt, and L. P. McGuinness, "Experimental Datasets," (2019).

[58]R. Santagati, A. Gentile, S. Knauer, S. Schmitt, S. Paesani, C. Granade, N. Wiebe, C. Osterkamp, L. McGuinness, J. Wang, M. Thompson, J. Rarity, F. Jelezko, and A. Laing, Physical Review X **9**, 021019 (2019).

[59]K. Chaloner and I. Verdinelli, Statistical Science **10**, 273 (1995).

[60]C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, New Journal of Physics **14**, 103013 (2012).

[61]C. Granade, C. Ferrie, I. Hincks, S. Casagrande, T. Alexander, J. Gross, M. Kononenko, and Y. Sanders, Quantum **1**, 5 (2017), arXiv: 1610.00336.

[62]X. Huan and Y. M. Marzouk, Journal of Computational Physics **232**, 288 (2013).

[63]R. D. McMichael, S. M. Blakley, and S. Dushenko, Journal of Research of the National Institute of Standards and Technology **126**, 126002 (2021).

[64]O. Vasicek, Journal of the Royal Statistical Society. Series B (Methodological) **38**, 54 (1976), publisher: [Royal Statistical Society, Wiley].

FIGURE 1

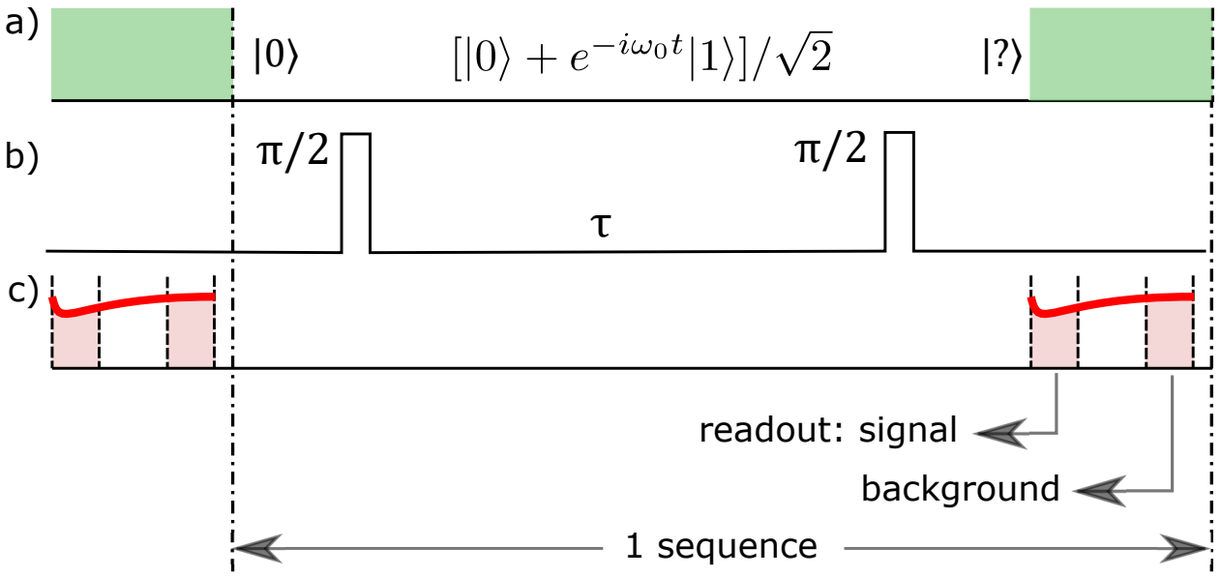


FIGURE 2

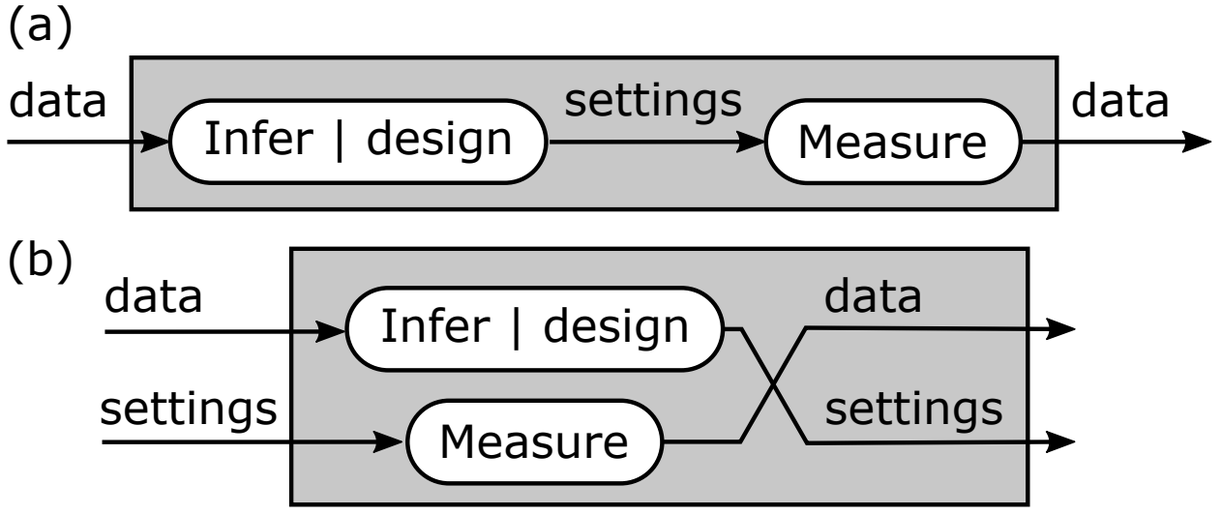


FIGURE 3

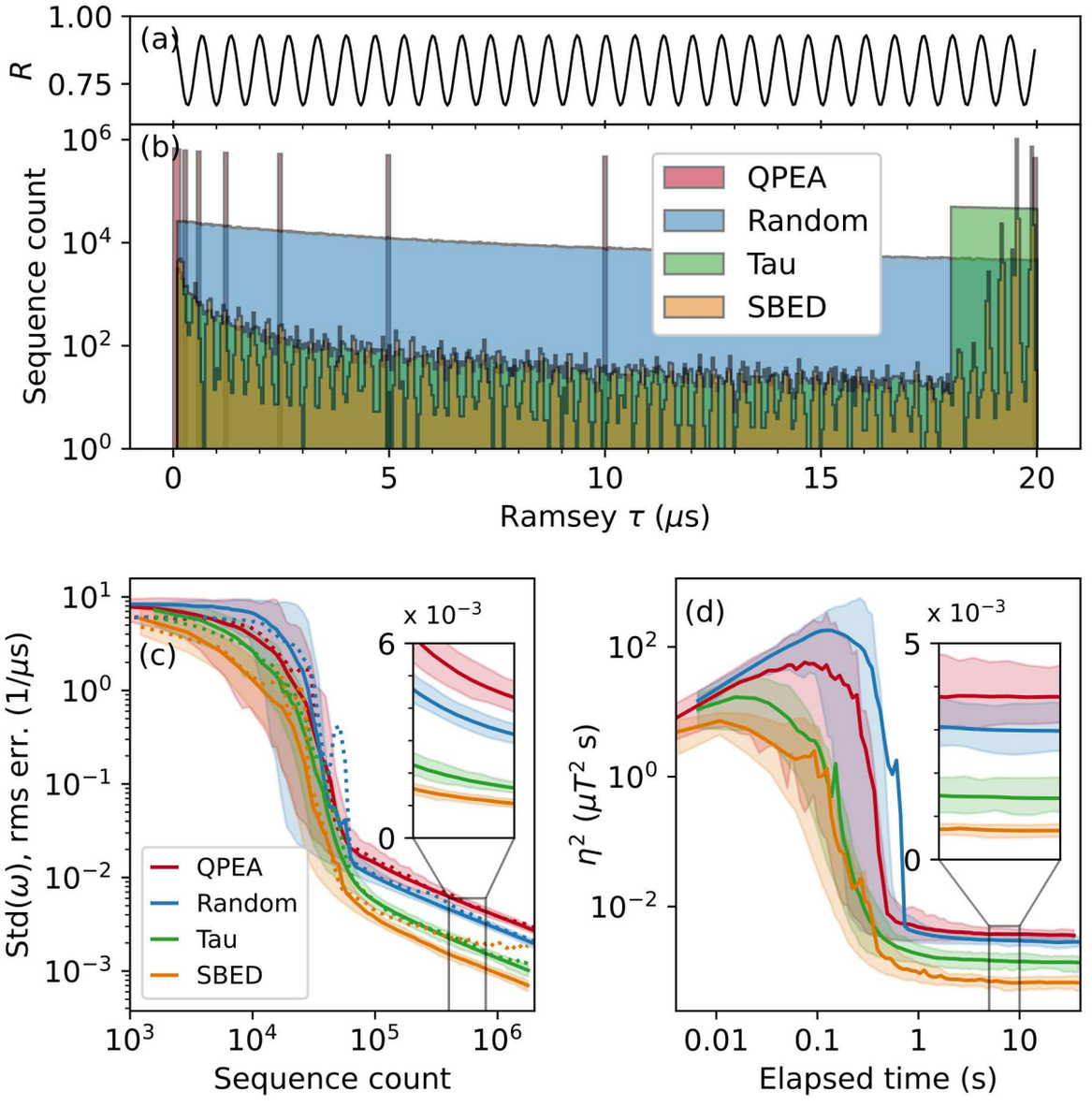


FIGURE 4

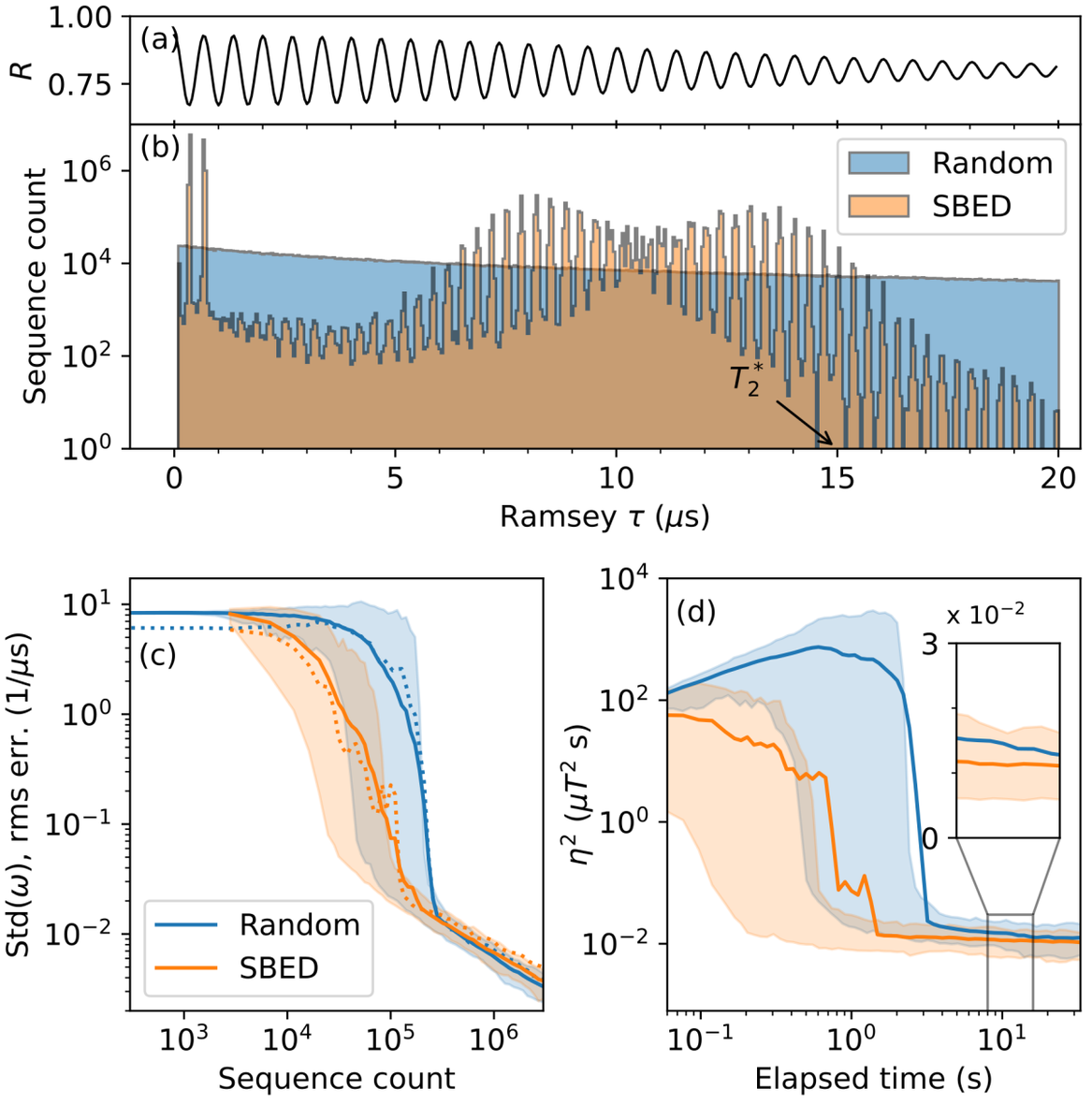
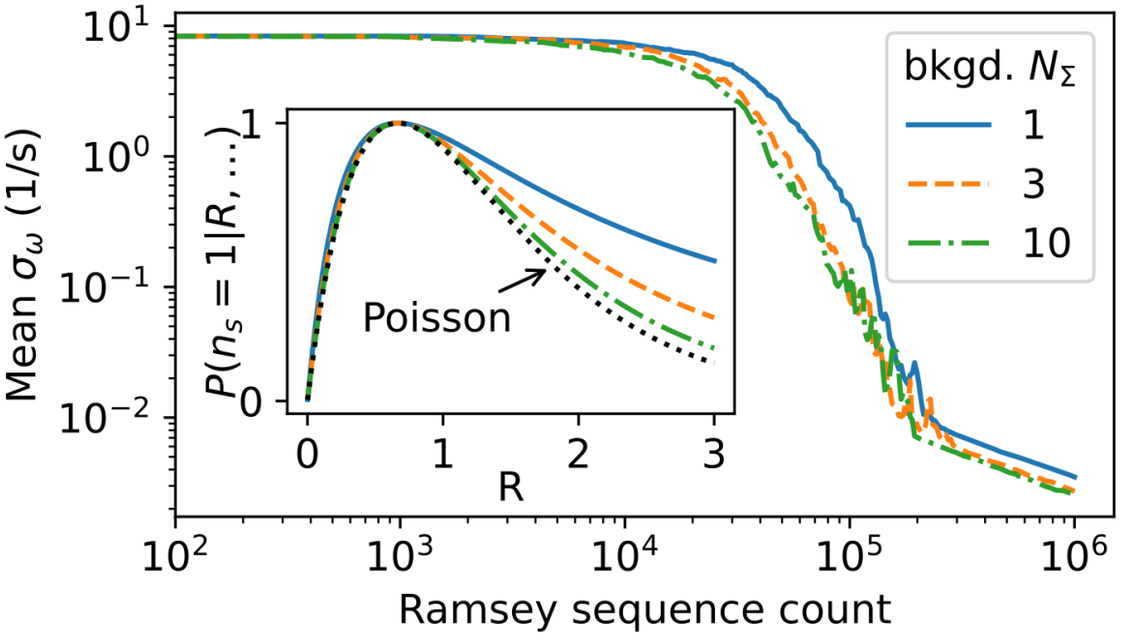
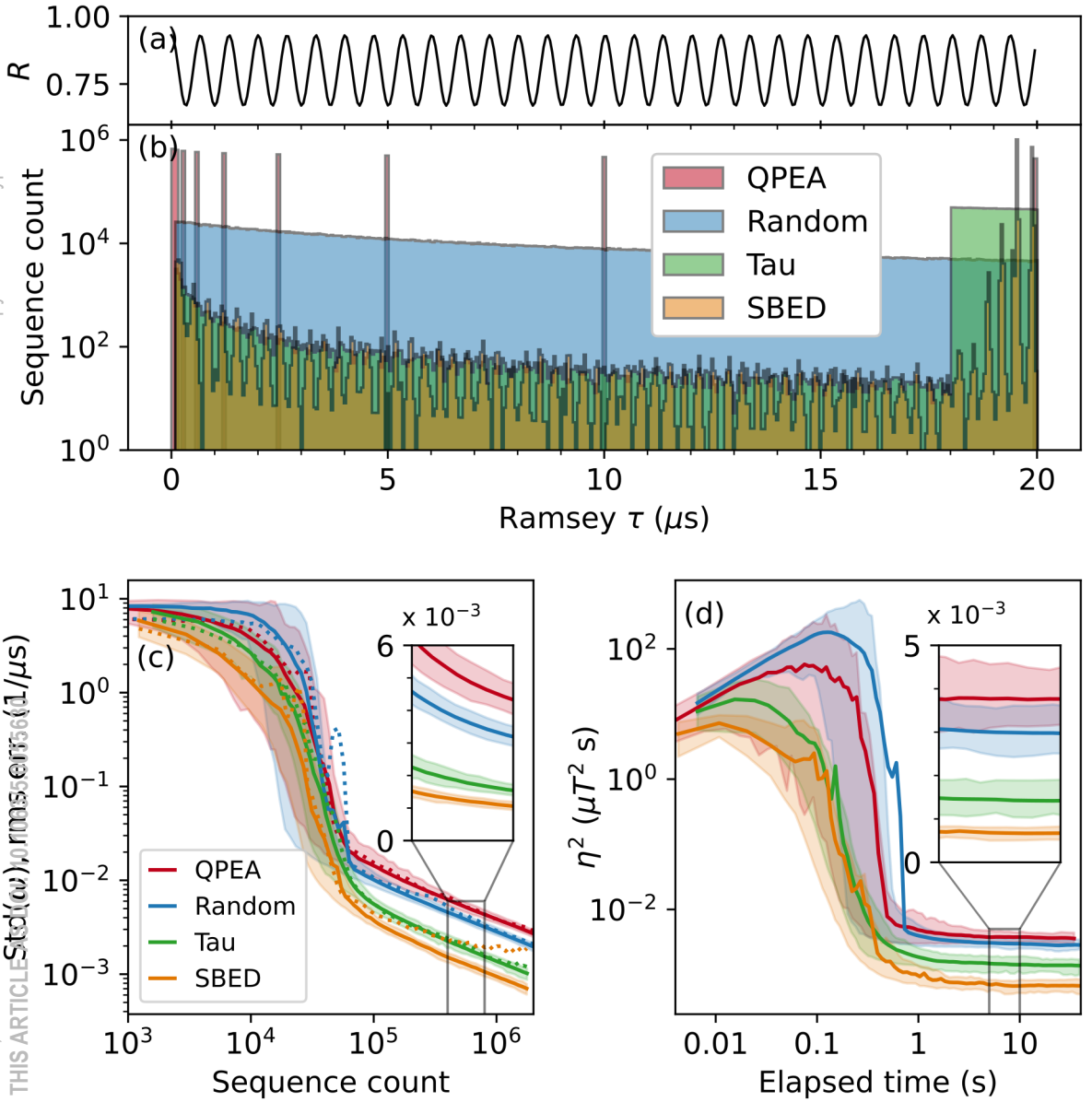


FIGURE 5

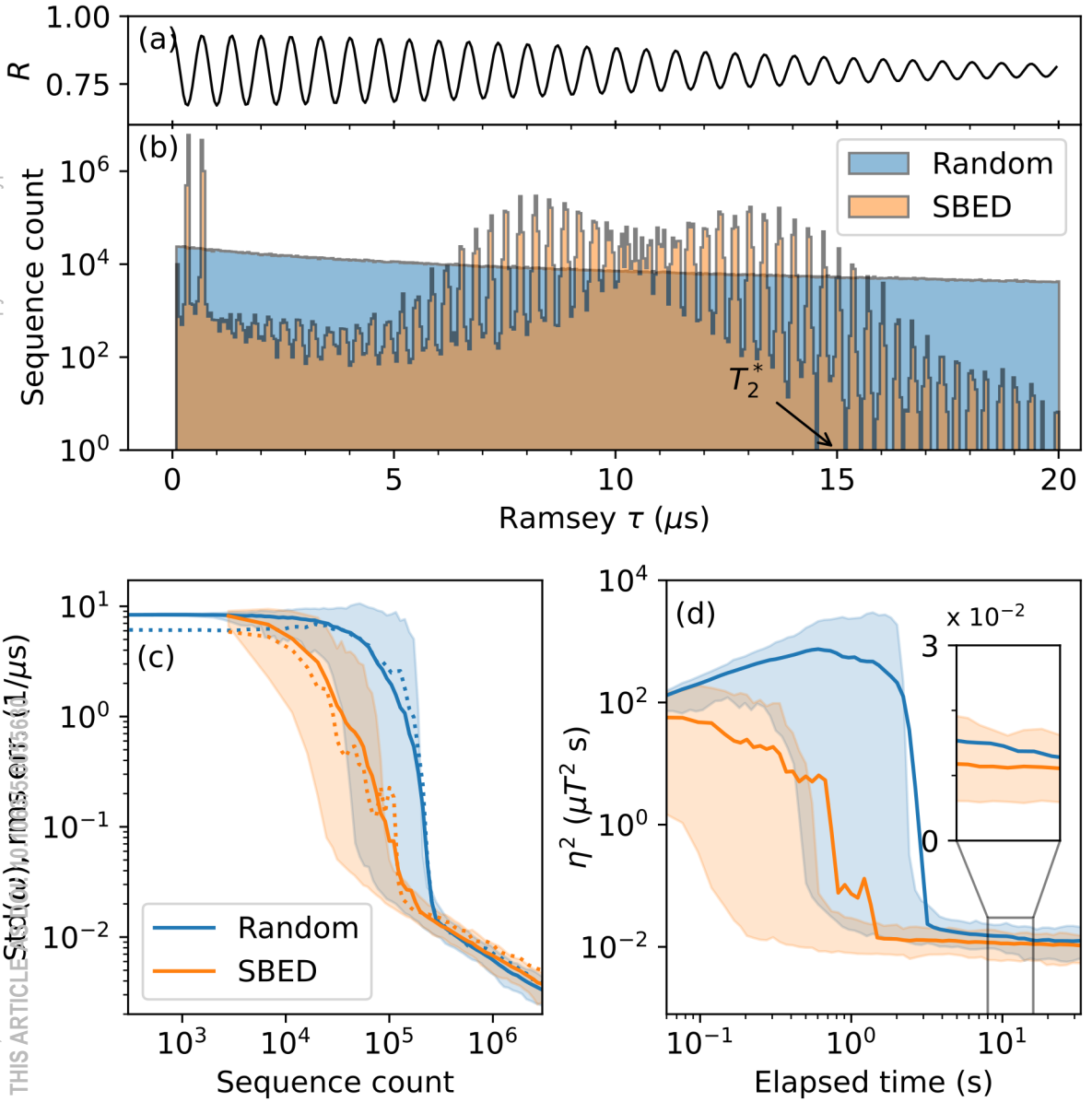


Journal of Applied Physics



Journal of Applied Physics

AIP Publishing



AIP Publishing

Journal of Applied Physics

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0055630