Qubit control noise spectroscopy with optimal suppression of dephasing

Vivian Maloney[1,*], Yasuo Oda,[2] Gregory Quiroz,[1,2] B. David Clader[1,†] and Leigh M. Norris[1]

[1]*The Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland 20723, USA*
[2]*William H. Miller III Department of Physics & Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA*

We extend quantum noise spectroscopy (QNS) of amplitude control noise to settings where dephasing noise or detuning errors make significant contributions to qubit dynamics. Previous approaches to characterize amplitude noise are limited by their vulnerability to low-frequency dephasing noise and static detuning errors, which can overwhelm the target control noise signal and introduce bias into estimates of the amplitude noise spectrum. To overcome this problem, we leverage optimal control to identify a family of amplitude control waveforms that optimally suppress low-frequency dephasing noise and detuning errors, while maintaining the spectral concentration in the amplitude filter essential for spectral estimation. The waveforms found via numerical optimization have surprisingly simple analytic forms, consisting of oscillating sine waves obeying particular amplitude and frequency constraints. In numerically simulated QNS experiments, these waveforms demonstrate superior robustness, enabling accurate estimation of the amplitude noise spectrum in regimes where existing approaches are biased by low-frequency dephasing noise and detuning errors.

I. INTRODUCTION

Robust, accurate control is a requirement for quantum technologies ranging from quantum sensors to quantum computers. In addition to being limited by environmental noise sources or by crosstalk between neighboring quantum systems, current gate fidelities are constrained by noise in the drives and classical electronics used for control [1]. Common examples of control noise include jitter in the amplitude and phase of microwave fields used to implement gates in platforms such as superconducting qubits, trapped ions, quantum dots, and diamond nitrogen vacancy centers [2–5]. In numerous platforms, control noise is addressed using composite pulses (CPs), which originated in nuclear magnetic resonance to correct unknown sources of static amplitude and phase noise [6]. Approaches such as dynamical decoupling (DD) or dynamically corrected quantum gates (DCGs) are capable of refocusing both control and environmental noise sources, as long as the temporal correlations of the noise decay on timescales that are slow compared to the control [7–9]. While CPs, DD, and DCGs are powerful in that they can increase gate fidelities without detailed knowledge of the control noise, further gains can be realized by taking into account specific features of the noise affecting a quantum device and leveraging optimal control to design robust, customized gates [3,10–12].

Obtaining the knowledge of the noise necessary for customized error mitigation is the domain of quantum noise spectroscopy (QNS) [13]. Over the last decade, QNS has become a framework to characterize temporally correlated noise in quantum devices by reconstructing the associated noise power spectral densities or noise spectra. The noise spectra provide a means to perturbatively model the noisy dynamics of a quantum device, estimate gate fidelities, and define objective functions for optimal control [3,14]. QNS protocols were originally devised for classical, Gaussian, dephasing noise [15,16], but have been extended to multiaxis noise [17], non-Gaussian dephasing noise [18,19], and noise on multiple qubits [20–22]. Most QNS protocols follow a procedure in which a quantum system is (1) prepared in a state that is sensitive to a target noise source, (2) allowed to evolve under the influence of noise and control, and, finally, (3) measured to estimate the expected value of an observable that captures the dynamical effect of the noise. The control applied in step (2) modifies the spectral response of the quantum system to noise. Typically, the control is selected so the system is sensitive to noise at frequencies within a certain range or band, yielding an estimate of the noise spectrum in a localized region of the frequency domain.

The first QNS protocol for characterizing quantum control noise, introduced and experimentally validated on trapped ions in Ref. [23], utilized control based on Slepian sequences from classical signal processing to estimate the spectrum of multiplicative amplitude noise acting on a single qubit. Waveforms based on Slepian sequences were later extended to enable simultaneous estimation of amplitude and dephasing noise spectra in Ref. [24]. Slepian sequences are notable for their optimal spectral concentration or degree of localization in a target frequency band [25]. In both classical and quantum settings, the spectral concentration of Slepians translates into estimates of the noise spectrum with minimal *leakage bias* or contamination from noise at frequencies outside the target band [25–27]. In addition to leveraging Slepians and other tools from classical signal processing, Refs. [23,24] applied a quantum tomographic procedure to isolate the dynamical contribution of the amplitude noise from dephasing, the secondary noise source in the trapped ion device. Because it relies on a perturbative expansion, the tomographic procedure

---

*Vivian.Maloney@jhuapl.edu
†Present address: Goldman Sachs & Co.

ensures the estimate of the amplitude noise is unbiased by the presence of dephasing provided that both noise sources are sufficiently weak and, furthermore, that the dynamical contribution of the amplitude noise is on the same order or stronger than that of dephasing. While this condition holds in the trapped ion device owing to the dominance of amplitude noise, higher order terms in the perturbative expansion can become significant in platforms where dephasing is strong relative to amplitude noise and/or predominantly low frequency, characteristic of superconducting qubits and other solid-state devices. In addition to dephasing, higher order contributions can also become significant in the presence of static detuning errors, which can arise from miscalibration or drift in the carrier frequency of an external control field.

In this paper, we present a protocol for amplitude noise QNS in regimes where dephasing noise or detuning errors make significant contributions to qubit dynamics. Our protocol relies on a reduced-complexity optimization procedure that takes into account realistic device constraints to arrive at control waveforms that reduce the impact of higher order perturbative terms while retaining a high degree of spectral concentration. While the optimized waveforms sacrifice some spectral concentration with respect to the Slepians, they effectively suppress low-frequency dephasing noise and cancel static detuning errors. The optimized waveforms are well approximated by a family of analytic functions, which we term the dephasing-robust waveforms. In numerically simulated QNS experiments, we show that the dephasing robust waveforms reduce the magnitudes of higher order terms in the perturbative expansion, allowing for reliable estimation of the amplitude noise spectrum when combined with the tomographic procedure of Refs. [23,24].

The paper is organized as follows. Section II summarizes the perturbative treatment of the qubit dynamics under single-axis control, multiplicative amplitude noise, and dephasing originally presented in Ref. [23]. This section also introduces the noise spectrum and control filter functions (FFs), which capture how applied control modifies the response of the qubit to noise in the frequency domain. In Sec. III, we present the tomographic procedure, including perturbative corrections up to fourth order. We then cast the problem of finding a control waveform that minimizes the dominant higher-order corrections into an optimization, using techniques from linear programming to scalably incorporate physical constraints. Lastly, we introduce the analytic dephasing-robust waveforms, which closely approximate the results of our numerical optimization. Finally, in Sec. IV, we compare the dephasing-robust waveforms to the Slepians by numerically simulating amplitude noise QNS on a qubit subject to low-frequency dephasing noise or detuning error at various strengths. As the strength of the amplitude noise decreases relative to the dephasing or detuning error, we demonstrate that the dephasing robust waveforms enable accurate reconstruction of the amplitude noise spectrum.

II. BACKGROUND

A. Time domain: Control, noise, and dynamics

We consider a single qubit subject to time-dependent, temporally correlated noise and controlled by modulating the amplitude $\Omega(t)$ and phase $\phi(t)$ of an external driving field. By transforming into a frame rotating with the carrier frequency of the drive at resonance with the qubit energy splitting and making the rotating wave approximation, we can write the control Hamiltonian of the qubit in units of $\hbar = 1$ as

$$H_c(t) = \frac{\Omega(t)}{2}[\cos\phi(t)\sigma_1 + \sin\phi(t)\sigma_2]. \qquad (1)$$

Here, we denote the Pauli operators by $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3) \equiv (\sigma_x, \sigma_y, \sigma_z)$. The actual qubit evolution differs from the ideal dynamics generated by $H_c(t)$ due to the presence of dephasing and amplitude control noise. In the rotating frame, the noise is described by the Hamiltonian [3]:

$$H_N(t) = \beta_z(t)\sigma_3 + \beta_\Omega(t)H_c(t). \qquad (2)$$

Here, the dephasing $\beta_z(t)$ and amplitude noise $\beta_\Omega(t)$ are independent, stationary, Gaussian stochastic processes. Because $\beta_z(t)$ and $\beta_\Omega(t)$ are scalar and commute at all times, we stress that they are classical noise processes. Quantum noise processes, on the other hand, are associated with noncommuting operators acting on an external environment or bath.

Note that the dephasing enters $H_N(t)$ as an additive frequency fluctuation whereas the amplitude noise acts multiplicatively on the control Hamiltonian, inducing fluctuations in the Rabi frequency with magnitude $\Omega(t)\beta_\Omega(t)$. While $\beta_z(t)$ has units of frequency, $\beta_\Omega(t)$ is dimensionless. Static detuning errors, which generate coherent rotations about σ_3, contribute to the mean of the dephasing process $\beta_z(t)$. For later convenience, we write the noise Hamiltonian in matrix form

$$H_N(t) = \vec{B}(t)\,\mathbf{N}(t)\,\vec{\sigma}^T, \qquad (3)$$

where

$$\vec{B}(t) \equiv [\beta_\Omega(t), \beta_\Omega(t), \beta_z(t)]$$

and

$$\mathbf{N}(t) \equiv \begin{pmatrix} \frac{1}{2}\Omega(t)\cos\phi(t) & 0 & 0 \\ 0 & \frac{1}{2}\Omega(t)\sin\phi(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (4)$$

are the noise vector and noise matrix, respectively.

The complete qubit dynamics, including both control and noise, is described by the rotating-frame Hamiltonian $H(t) = H_c(t) + H_N(t)$. To isolate the dynamical contribution of the noise, we make one additional transformation into the toggling frame or interaction picture associated with $H_c(t)$. If $U_c(t) = \mathcal{T}_+\exp[-i\int_0^t ds\, H_c(s)]$ is the ideal gate implemented by the noiseless control, the qubit dynamics in the toggling frame are generated by the error Hamiltonian:

$$\tilde{H}(t) = U_c(t)^\dagger H_N(t)U_c(t). \qquad (5)$$

Since it is a unitary conjugation of Eq. (3), the error Hamiltonian can be expressed in matrix form as

$$\tilde{H}(t) = \vec{B}(t)\,\mathbf{Y}(t)\,\vec{\sigma}^T, \qquad (6)$$

where the control matrix is defined as $\mathbf{Y}(t) \equiv \mathbf{N}(t)\mathbf{R}(t)$ with $[\mathbf{R}(t)]_{ij} \equiv \frac{1}{2}\mathrm{Tr}[U_c(t)^\dagger\sigma_i U_c(t)\sigma_j]$ [27]. The unitary generated by the error Hamiltonian or error propagator, $\tilde{U}(t) = \mathcal{T}_+\exp[-i\int_0^t ds\, \tilde{H}(s)]$, evolves the qubit in the toggling frame and is related to the rotating-frame propagator by $U(t) \equiv$

$\mathcal{T}_+ \exp[-i \int_0^t ds\, H(s)] = U_c(t)\tilde{U}(t)$. Observe that the rotating and toggling frames are equivalent at a time t such that $U_c(t) = I$. Following Ref. [14], we parametrize the error propagator as

$$\tilde{U}(t) = \exp[-i\vec{a}(t)\cdot\vec{\sigma}] = e^{-i\sum_m \vec{a}^{(m)}(t)\cdot\vec{\sigma}}, \qquad (7)$$

where $\vec{a}(t) \equiv [a_1(t), a_2(t), a_3(t)]$ is the real, time-dependent error vector. On the right-hand side of this expression, the mth term in the summation depends on the mth-order term of a perturbative Magnus expansion of the propagator [28]. For a toggling-frame Hamiltonian along multiple Pauli components, it is generally impossible to determine an exact expression for $\vec{a}(t)$. For sufficiently weak noise [14], the error vector is approximated by the leading order term in the Magnus expansion, so $\tilde{U}(t) \approx \exp[-i \int_0^t ds\, \tilde{H}(s)]$ and

$$\vec{a}(t) \approx \vec{a}^{(1)}(t) \equiv \int_0^t ds\, \vec{B}(s)\, \mathbf{Y}(s), \qquad (8)$$

where the second equality follows from Eq. (6).

Thus far, we have considered the most general form of the control Hamiltonian in Eq. (1). In the remainder of this paper, we restrict the drive phase to $\phi(t) \in \{0, \pi\}$ so the control generates rotations about the x axis and

$$H_c(t) = \frac{\Omega(t)}{2}\sigma_1, \qquad (9)$$

where we have redefined the amplitude control waveform $\Omega(t)$ to take both positive and negative values. For this simplified control,

$$\mathbf{N}(t) = \begin{pmatrix} \frac{1}{2}\Omega(t) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{R}(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Theta(t) & -\sin\Theta(t) \\ 0 & \sin\Theta(t) & \cos\Theta(t) \end{pmatrix}, \qquad (10)$$

$$\mathbf{Y}(t) = \begin{pmatrix} \frac{1}{2}\Omega(t) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sin\Theta(t) & \cos\Theta(t) \end{pmatrix},$$

where $\Theta(t) = \int_0^t ds\, \Omega(s)$. From Eqs. (4), (8), and (10), the components of the error vector to leading order in the Magnus expansion are [23,27]

$$a_1^{(1)}(t) = \frac{1}{2}\int_0^t ds\, \Omega(s)\beta_\Omega(s), \qquad (11)$$

$$a_2^{(1)}(t) = \int_0^t ds\, \sin\Theta(s)\beta_z(s), \qquad (12)$$

$$a_3^{(1)}(t) = \int_0^t ds\, \cos\Theta(s)\beta_z(s). \qquad (13)$$

Due to the sparsity of the control matrix, each component of the error vector is the integrated product of a single control matrix element and a single noise source. Observe that the amplitude noise enters the qubit dynamics solely through the $a_1^{(1)}(t)$ component.

B. Frequency Domain: Spectra and Filter Functions

Because we are ultimately interested in the spectral properties of the noise, we transform from the time domain to the frequency domain, where the noise and control are represented by spectra and FFs, respectively. If the dephasing and amplitude noise sources are wide-sense stationary, their autocovariances can be parametrized in terms of the *lag time*, $\tau \equiv t_2 - t_1$, so

$$\langle \Delta\beta_n(t_1)\Delta\beta_n(t_2)\rangle = \langle \Delta\beta_n(\tau)\Delta\beta_n(0)\rangle,$$

where $n \in \{\Omega, z\}$, $\Delta O \equiv O - \langle O\rangle$, and $\langle\cdot\rangle$ denotes the ensemble average over realizations of the amplitude and dephasing noise processes. By the Weiner-Kinchin theorem, the power spectral density or spectrum of the noise is the Fourier transform of the noise autocovariance with respect to the lag time:

$$S_n(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau\, \langle \Delta\beta_n(\tau)\Delta\beta_n(0)\rangle e^{-i\omega\tau}. \qquad (14)$$

The objective of our noise spectroscopy procedure is estimating $S_\Omega(\omega)$, the spectrum associated with the amplitude noise.

Estimating $S_\Omega(\omega)$ requires experimentally measurable quantities that depend on the noise spectra. We take the approach of Ref. [23] in which the amplitude noise spectrum is inferred from quantities derived from the error vector components. For notational convenience, define

$$\vec{\beta}(t) \equiv [\beta_\Omega(t), \beta_z(t), \beta_z(t)], \qquad (15)$$

$$\vec{y}(t) \equiv [\Omega(t)/2, \sin\Theta(t), \cos\Theta(t)], \qquad (16)$$

so the error vector components in Eqs. (11)–(13) are given by $a_i^{(1)}(t) = \int_0^t ds\, y_i(s)\beta_i(s)$ for $i \in \{1, 2, 3\}$. The variances of the error vector components are then [23,27]

$$\langle \Delta a_i^{(1)}(T)^2\rangle = \int_0^T dt_2 \int_0^T dt_1\, \langle \Delta\beta_i(t_1)\Delta\beta_i(t_2)\rangle\, y_i(t_1)y_i(t_2)$$

$$= \frac{1}{\pi}\int_0^\infty d\omega\, S_i(\omega)\, F_i(\omega, T). \qquad (17)$$

This expression is an overlap integral between the noise spectrum,

$$S_i(\omega) \equiv \begin{cases} S_\Omega(\omega), & i = 1 \\ S_z(\omega), & i = 2, 3, \end{cases}$$

and the associated *control* FF [3,29,30]:

$$F_i(\omega, T) \equiv \left|\int_0^T dt\, e^{i\omega t}y_i(t)\right|^2. \qquad (18)$$

The control FFs are frequency-domain representations of the control generated by $H_c(t)$. For later convenience, we follow Ref. [3] and group the FFs according to whether they describe the qubit's response to amplitude or dephasing noise in Eq. (17), forming the amplitude and dephasing FFs

$$F_\Omega(\omega, T) \equiv F_1(\omega, T) = \frac{1}{4}\left|\int_0^T dt\, e^{i\omega t}\Omega(t)\right|^2, \qquad (19)$$

$$F_Z(\omega, T) \equiv F_2(\omega, T) + F_3(\omega, T)$$

$$= \left|\int_0^T dt\, e^{i\omega t}\sin\Theta(t)\right|^2 + \left|\int_0^T dt\, e^{i\omega t}\cos\Theta(t)\right|^2. \qquad (20)$$

Observe that the amplitude noise spectrum enters the qubit dynamics through $\langle \Delta a_1^{(1)}(T)^2 \rangle$, while the dephasing noise spectrum is present in both $\langle \Delta a_2^{(1)}(T)^2 \rangle$ and $\langle \Delta a_3^{(1)}(T)^2 \rangle$, hence the two contributions to the dephasing FF in Eq. (20).

C. Quantum noise spectroscopy

QNS protocols characterize the environmental noise of a quantum system by measuring system observables under various operational settings and correlating the measured values to the strength of different noise sources. The filter function formalism (FFF) offers a particularly clear insight into how this process works. In the FFF, to leading order in a perturbative expansion, the average value of a measured system quantity is expressed as an overlap integral between the spectrum of a noise source and a control FF,

$$\langle M(T) \rangle = C \int_0^\infty d\omega \, S(\omega) F(\omega, T), \qquad (21)$$

where C is some proportionality constant.

Note that the variance in Eq. (17) takes precisely this form. Applying different controls to the quantum system modifies the shape of $F(\omega, T)$ in the frequency domain. For QNS, the FF is often engineered to be localized within a chosen passband or frequency interval, e.g., $B \equiv [\omega_0 - \delta\omega, \omega_0 + \delta\omega]$. For sufficiently small $\delta\omega$, $S(\omega)$ is approximately constant within B. With the additional assumption that the support of $F(\omega, T)$ is minimal outside B, the quantity $\langle M(T) \rangle$ is approximately proportional to $S(\omega_0)$,

$$\langle M(T) \rangle \approx S(\omega_0) C \int_B d\omega \, F(\omega, T). \qquad (22)$$

Since C and $F(\omega, T)$ are known, we can infer $S(\omega_0)$. Choosing a set of control waveforms that concentrate $F(\omega, T)$ in different passbands then enables us to characterize $S(\omega)$ over a range of frequencies.

Note that this procedure depends critically on $F(\omega, T)$ being spectrally concentrated or localized in the target frequency band, B. If $F(\omega, T)$ exhibits spectral leakage or out-of-band spectral components, the linear approximation of the overlap integral in Eq. (22) is no longer valid and $\langle M(t) \rangle$ will depend on the value of the spectrum outside of B. References [23,24,27] minimized spectral leakage for QNS applications by using control waveforms derived from Slepian sequences, a family of discrete sequences that maximize spectral concentration for spectral estimation based on classical time series [26]. In absence of suitable control to localize the FF, the spectrum can still be estimated through the frequency comb techniques [16], linear inversion, or equivalent approaches that discretize the overlap integral over a range of frequencies [31,32]. Spectral leakage is still a detriment in these settings, however, as delocalized FFs can produce nonsparse, potentially ill-conditioned linear systems that amplify errors in the spectral estimate. As such, maximizing spectral concentration and minimizing spectral leakage is an important principle of filter design for QNS.

III. DEPHASING-ROBUST CHARACTERIZATION OF AMPLITUDE NOISE

Since $\langle \Delta a_1^{(1)}(T)^2 \rangle$ is an overlap integral between the amplitude noise spectrum and the amplitude FF, the filter-shaping strategies of the previous section can, in principle, be applied to infer $S_\Omega(\omega)$. Accomplishing this in a realistic system, however, requires a means of experimentally measuring $\langle \Delta a_1^{(1)}(T)^2 \rangle$. For noise that is zero mean, Ref. [23] devised a strategy to measure $\langle \Delta a_1^{(1)}(T)^2 \rangle$ that holds in a weak noise limit. As the strength of the noise increases, however, our ability to reliably determine $\langle \Delta a_1^{(1)}(T)^2 \rangle$ is complicated by terms that are higher order in the noise strength. In particular, higher order dephasing-dependent terms can make it impossible to isolate amplitude noise-dependent quantities, biasing estimates of $S_\Omega(\omega)$.

A. Bias from dephasing noise

For weak, zero-mean noise and amplitude control that generates a net identity (ensuring the toggling and rotating frames are equivalent), Ref. [23] determined $\langle \Delta a_1^{(1)}(T)^2 \rangle$ from a tomographic procedure depending on three experimentally measurable survival probabilities,

$$\langle \Delta a_1^{(1)}(T)^2 \rangle \approx \tfrac{1}{2}[1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)]$$
$$\equiv \mathcal{P}. \qquad (23)$$

Here, $\mathbb{P}(\uparrow_i, T) = \langle |\langle \uparrow_i | U(T) | \uparrow_i \rangle|^2 \rangle$ is the probability that the qubit remains in state $| \uparrow_i \rangle$ after evolving under noise and control for a time T and $| \uparrow_i \rangle$ denotes the $+1$ eigenstate of σ_i. This estimate holds in a weak-noise limit in which dynamical contributions beyond second order in $\beta_z(t)$ and $\Omega(t)\beta_\Omega(t)$ are negligible. Beyond this limit, terms that are higher order in the error vector components are no longer negligible in Eq. (23). In Appendix A, we extend the measurement strategy to noise with nonzero mean and include perturbative terms up to fourth order in $\beta_z(t)$ and $\Omega(t)\beta_\Omega(t)$. In the main text, we focus on dephasing noise with nonzero mean to treat static detuning errors, but take the amplitude noise to be zero mean for simplicity. With higher-order perturbative terms and nonzero-mean dephasing, Eq. (23) becomes

$$\langle \Delta a_1^{(1)}(T)^2 \rangle - \tfrac{1}{3}\langle a_1^{(1)}(T)^4 \rangle + \langle a_1^{(2)}(T)^2 \rangle$$
$$- \tfrac{1}{3}\langle a_2^{(1)}(T)^2 a_1^{(1)}(T)^2 \rangle - \tfrac{1}{3}\langle a_3^{(1)}(T)^2 a_1^{(1)}(T)^2 \rangle$$
$$\approx \tfrac{1}{2}[1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)]. \qquad (24)$$

Here, the superscript (2) denotes the second-order Magnus contribution to the error vector in Eq. (7), an explicit expression for which is provided in Appendix A. The first two terms of Eq. (24) depend solely on amplitude noise, while the remaining terms depend either on dephasing noise or both amplitude and dephasing noise. When the dephasing noise is strong relative to the amplitude noise or when the control increases the sensitivity of the qubit to dephasing noise or detuning error, the terms on the second line can dominate over the other higher order terms and become comparable in magnitude to $\langle \Delta a_1^{(1)}(T)^2 \rangle$. Therefore, naïvely employing Eq. (23) for amplitude noise spectroscopy will result in an estimate of $S_\Omega(\omega)$ that is biased by dephasing.

To better elucidate the relationship between noise strength, noise spectra, amplitude control, and the relative magnitudes of the terms in Eq. (24), we analyze the higher order terms in the frequency domain. First, define the overlap integrals

$$I_\Omega(T) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, F_\Omega(\omega, T) S_\Omega(\omega), \quad (25)$$

$$I_Z(T) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, F_Z(\omega, T) \big[S_z(\omega) + 2\pi \mu_z^2 \delta(\omega) \big], \quad (26)$$

where $\langle \beta_z(t) \rangle \equiv \mu_z$ is static in time due to stationarity. Observe that $\langle \Delta a_1^{(1)}(T)^2 \rangle = I_\Omega(T)$. Since $\beta_\Omega(t)$ and $\beta_z(t)$ are Gaussian, their fourth-order moments factor into products of one- and two-point correlation functions. Using this factorization, we show in Appendix A, that the higher order terms take the form

$$\langle a_1^{(1)}(T)^4 \rangle = 3 \, I_\Omega(T)^2, \quad (27)$$

$$\langle a_1^{(1)}(T)^2 a_2^{(1)}(T)^2 \rangle + \langle a_1^{(1)}(T)^2 a_3^{(1)}(T)^2 \rangle = I_\Omega(T) I_Z(T), \quad (28)$$

$$\begin{aligned}
\langle a_1^{(2)}(T)^2 \rangle = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \, G_Z(\omega, \omega', T) \\
\times \left[\frac{S_z(\omega) S_z(\omega')}{(2\pi)^2} + \frac{\mu_z^2}{2\pi} S_z(\omega) \delta(\omega') \right. \\
\left. + \frac{\mu_z^2}{2\pi} S_z(\omega') \delta(\omega) + \frac{\mu_z^4}{3} \delta(\omega) \delta(\omega') \right]. \quad (29)
\end{aligned}$$

Here, $G_Z(\omega, \omega', T)$ is a higher order FF given by

$$\begin{aligned}
G_Z(\omega, \omega', T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \\
\times \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)] \\
\times \left[e^{i\omega(t_1-t_2)} e^{i\omega'(t_3-t_4)} + e^{i\omega(t_1-t_3)} e^{i\omega'(t_2-t_4)} \right. \\
\left. + e^{i\omega(t_1-t_4)} e^{i\omega'(t_2-t_3)} \right]. \quad (30)
\end{aligned}$$

While Eq. (27) depends quadratically on the overlap integral between the amplitude noise spectrum and the amplitude FF, Eq. (28) is the product of the overlap integrals associated with amplitude noise and dephasing noise. Dephasing noise also enters Eq. (29), which is a two-dimensional overlap integral involving the dephasing noise spectrum and a higher order FF. We can express Eq. (24) in terms of the overlap integrals as

$$\begin{aligned}
\frac{1}{2} [1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)] \\
\approx I_\Omega(T) - I_\Omega(T)^2 - \frac{1}{3} I_\Omega(T) I_Z(T) + \langle a_1^{(2)}(T)^2 \rangle, \quad (31)
\end{aligned}$$

which highlights the source of each higher-order correction. If Eq. (23) is used to estimate $\langle \Delta a_1^{(1)}(T)^2 \rangle = I_\Omega(T)$, then Eq. (28) contributes a dephasing-dependent *multiplicative* bias proportional to $I_Z(T)$. The magnitude of this bias can be estimated from the measured survival probabilities:

$$\begin{aligned}
I_Z(T) = \langle \Delta a_2^{(1)}(T)^2 \rangle + \langle \Delta a_3^{(1)}(T)^2 \rangle \\
\approx 1 - \mathbb{P}(\uparrow_1, T). \quad (32)
\end{aligned}$$

The remaining higher-order term in Eq. (30) contributes a dephasing-dependent *additive* bias.

The additive and multiplicative biases can contribute to the estimator when the dephasing noise is stronger than the amplitude noise on the interval $[0, T]$, i.e.,

$$\int_0^T dt \, |\beta_z(t)|^2 > \int_0^T dt \, |\Omega(t) \beta_\Omega(t)|^2, \quad (33)$$

or when the dephasing FFs $F_Z(\omega, T)$ and $G_Z(\omega, \omega', T)$ have significant spectral support in regions of the frequency domain where $S_z(\omega)$ and $S_z(\omega) S_z(\omega')$ are, respectively, large. In this later case, the magnitudes of $I_Z(T)$ and $|\langle a_1^{(2)}(T)^2 \rangle|^{1/2}$ can be larger than $I_\Omega(T)$ even when the dephasing noise strength is comparable to or weaker than the amplitude noise strength. A particularly relevant example involves the $1/f$ dephasing noise present in solid-state platforms such as superconducting qubits and semiconductor qubits. Since $\mu_z \neq 0$ in the presence of detuning error, detuning contributes additional bias through Eq. (26) when the dephasing FF has dc support, i.e., $F_Z(0, T) \neq 0$. Similarly, if $G_Z(\omega, \omega', T) \neq 0$ at $\omega = 0$ or $\omega' = 0$, the detuning error can also contribute bias through Eq. (29).

B. Optimized waveforms

When the FFs have low-frequency or dc support, the higher-order dephasing-dependent terms can make a significant contribution to Eq. (24). As a consequence, $\langle \Delta a_1^{(1)}(T)^2 \rangle$ can no longer be isolated using the measurement strategy of Ref. [23], biasing estimates of the amplitude noise spectrum. Extending amplitude noise QNS to settings with detuning errors or dephasing noise that is predominantly low frequency requires a different control strategy. To engineer amplitude control waveforms that suppress low-frequency dephasing noise and detuning errors while maintaining spectral concentration of the amplitude filter, we turn to optimal control. The problem of filter design with bounded controls can be cast as a scalable optimization problem. Our objective is to design control waveforms of the form in Eq. (9) suited for amplitude noise spectroscopy in the presence of low-frequency dephasing noise or detuning errors. As such, the amplitude FF should be spectrally concentrated and the dephasing FF should act as a high-pass filter, capable of filtering out low-frequency and dc spectral components. To accomplish this, we exploit the properties of the family of Slepian or discrete prolate spheroidal sequences (DPSSs) [25]. For completeness, we have provided a brief summary of the DPSSs and their properties in Appendix B. The DPSS of order k is denoted by $\{v_n^{(k)}(N, W) \,|\, n = 0, \ldots, N-1\}$, where N is the length of the sequence and $W \in (0, 1/2)$ is the bandwidth parameter determining the size of the passband in the frequency domain. Owing to the spectral concentration of their discrete-time Fourier transforms (DTFTs), the lowest-order DPSSs have been used extensively in classical spectral estimation. References [23,27] harnessed the spectral concentration of the DPSS for amplitude noise spectroscopy in the quantum setting by employing the piecewise constant amplitude waveforms

$$\Omega(t) = \begin{cases} \Omega \, v_0^{(k)}(N, W), & t \in [0, \Delta t) \\ \Omega \, v_1^{(k)}(N, W), & t \in [\Delta t, 2\Delta t) \\ \quad \vdots & \quad \vdots \\ \Omega \, v_{N-1}^{(k)}(N, W), & t \in [(N-1)\Delta t, N\Delta t), \end{cases} \quad (34)$$

where Ω is the control amplitude in units of angular frequency and Δt is a time increment. This waveform produces an amplitude FF that is spectrally concentrated in the passband $[-2\pi W/\Delta t, 2\pi W/\Delta t]$. Through analog modulation techniques, the passband can be shifted to different locations along the frequency axis [27]. For example, sine modulation entails modifying the waveform above so $\Omega(t) = \Omega \sin(\omega_0 n \Delta t) v_n^{(k)}(N, W)$ for $t \in [n\Delta t, (n+1)\Delta t)$. This produces an amplitude FF with a positive-frequency pass-

band $B_{\mathrm{DPSS}}(\omega_0) \equiv [-2\pi W/\Delta t + \omega_0, 2\pi W/\Delta t + \omega_0]$. Cosine modulation, in which $\sin(\omega_0 n \Delta t)$ is replaced in the amplitude control waveform by $\cos(\omega_0 n \Delta t)$, likewise produces an amplitude FF with positive-frequency passband $B_{\mathrm{DPSS}}(\omega_0)$.

Inspired by the gradient ascent in function space (GRAFS) approach of Refs. [11,33], we parametrize the amplitude control waveform as a piecewise constant linear combination of cosine and sine modulated DPSS,

$$
\Omega(t) = \begin{cases}
\sum_{k=0}^{K-1} \Omega_{\mathrm{c}}^{(k)} v_0^{(k)}(N, W), & t \in [0, \Delta t) \\
\sum_{k=0}^{K-1} \left\{ \Omega_{\mathrm{c}}^{(k)} \cos[\omega_0 \Delta t] + \Omega_{\mathrm{s}}^{(k)} \sin[\omega_0 \Delta t] \right\} v_1^{(k)}(N, W), & t \in [\Delta t, 2\Delta t) \\
\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \vdots & \vdots \\
\sum_{k=0}^{K-1} \left\{ \Omega_{\mathrm{c}}^{(k)} \cos[\omega_0 (N-1)\Delta t] + \Omega_{\mathrm{s}}^{(k)} \sin[\omega_0 (N-1)\Delta t] \right\} v_{N-1}^{(k)}(N, W), & t \in [(N-1)\Delta t, N\Delta t).
\end{cases} \tag{35}
$$

Here, the contributions of the cosine- and sine-modulated DPSSs of order k are determined, respectively, by the real amplitudes $\Omega_{\mathrm{s}}^{(k)}$ and $\Omega_{\mathrm{c}}^{(k)}$. Since for all $\phi \in \mathbb{R}$ there exist $\Omega_{\mathrm{s}}^{(k)}$ and $\Omega_{\mathrm{c}}^{(k)}$ such that $\sin(\omega_0 m \Delta t + \phi) = \Omega_{\mathrm{s}}^{(k)} \sin(\omega_0 m \Delta t) + \Omega_{\mathrm{c}}^{(k)} \cos(\omega_0 m \Delta t)$, we do not lose any expressive capability by restricting the phase to zero in Eq. (35). Due to the linearity of the Fourier transform, this waveform produces an amplitude FF spectrally concentrated in $B_{\mathrm{DPSS}}(\omega_0)$, equivalent to the passband of the constituent cosine- and sine-modulated DPSSs.

While spectral concentration of the amplitude filter is enforced by the DPSS, we achieve the desired high-pass dephasing FF by minimizing a control-dependent objective function over the $2K$-dimensional space of the amplitudes $\vec{\Omega} \equiv [\Omega_{\mathrm{c}}^{(0)}, \Omega_{\mathrm{s}}^{(0)}, \ldots, \Omega_{\mathrm{c}}^{(K-1)}, \Omega_{\mathrm{s}}^{(K-1)}]$. Since $\Omega(t)$ is restricted to low-order DPSSs for which $2K \ll N$, this optimization offers a significant dimensionality reduction over approaches such as GRAPE that optimize over control parameters in each of the N timesteps [33]. To suppress low-frequency dephasing noise, we choose an objective function given by the overlap integral in Eq. (25) with a $1/f$ dephasing noise spectrum,

$$
I_Z(\vec{\Omega}, T) = \frac{1}{\pi} \int_0^\infty d\omega \, F_Z(\omega, \vec{\Omega}, T) \frac{1}{\omega + \delta\omega}, \tag{36}
$$

where $\delta\omega > 0$ is a small regularization term to make the integral well-defined. Minimizing this objective function corresponds to suppressing the fourth order, dephasing-dependent terms in Eq. (28), reducing the bias from dephasing noise. While we do not explicitly incorporate the fourth-order contribution from Eq. (29) into the objective function, we found that the waveforms resulting from our optimization reduce the magnitude of this term, as we will demonstrate numerically.

In realistic experimental implementations, the control amplitude is bounded by the maximum Rabi rate, $\Omega_{\max}$. This imposes an L_∞ constraint on the control waveform in Eq. (35),

$$
\left| \sum_{k=0}^{K-1} v_m^{(k)}(N, W) \left\{ \Omega_{\mathrm{c}}^{(k)} \cos[\omega_0 m \Delta t] + \Omega_{\mathrm{s}}^{(k)} \sin[\omega_0 m \Delta t] \right\} \right|
$$
$$
\leqslant \Omega_{\max}, \tag{37}
$$

for $m \in \{0, \ldots, N-1\}$. The set of $\vec{\Omega}$ satisfying these constraints defines the true feasible region of physically realizable, piecewise constant control waveforms. Directly imposing all $2N$ inequality constraints in our optimization, however, is not scalable and removes the low-dimension advantage of the Slepian functional basis. Reference [33] instead accounts for the maximum Rabi rate by imposing box inequality bounds on the individual basis function coefficients,

$$
K \max_m \left| v_m^{(k)}(N, W) \right| \left| \Omega_{\mathrm{s}}^{(k)} \right| \leqslant \frac{\Omega_{\max}}{2},
$$
$$
K \max_m \left| v_m^{(k)}(N, W) \right| \left| \Omega_{\mathrm{c}}^{(k)} \right| \leqslant \frac{\Omega_{\max}}{2},
$$

for $k \in \{0, \ldots, K-1\}$ and using the L-BFGS-B optimization algorithm, which natively handles such constraints. This approach reduces the number of constraints from $2N$ to $4K$. However, as depicted in Fig. 1(a) for the two-dimensional case with $K = 1$, the feasible region defined by the box constraints disregards large portions of the true feasible region. Note that for this example, the corners of the box constraint do not touch the edges of the true feasible region. An optimization algorithm will fail to find solutions that lie within the true feasible region if they fall outside the box constraints. This is particularly problematic for high-amplitude waveforms, which lie at the boundary of the true feasible region.

To reduce the number of constraints without overly restricting the feasible region, we developed a dimensionality reduction procedure based on linear programming. This procedure efficiently approximates the true feasible region with a reduced set I of linear inequality constraints numbering $|I| \ll 2N$. Intuitively, the dimensionality reduction identifies the most critical constraints as those with a significant effect on the size of the feasible region if added or removed. It retains these critical constraints while removing the constraints which have a minor impact on the feasible region. The critical constraints that remain in set I take the form

$$
\forall i \in I \quad \sum_{k=0}^{K-1} \left(b_{\mathrm{s},i}^{(k)} \Omega_{\mathrm{s}}^{(k)} + b_{\mathrm{c},i}^{(k)} \Omega_{\mathrm{c}}^{(k)} \right) \leqslant \Omega_{\max}, \tag{38}
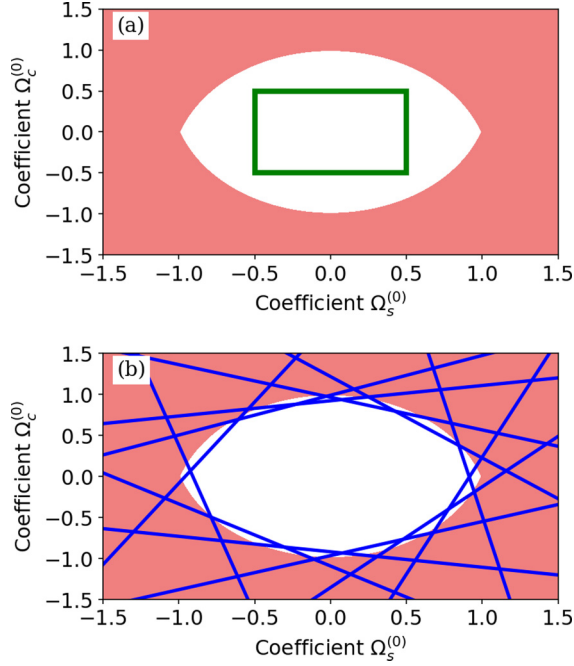$$

FIG. 1. Amplitude constraints for optimal control in a toy example with $K = 1$ and $\Omega_{\max} = 1$. Plot (a) compares the feasible region produced by box constraints [the interior of the dark (green) rectangle] to the true feasible region (white), showing that large portions of the true feasible region are excluded by the box constraints. In (b), the linear-programming reduced feasible region [area enclosed by the dark (blue) lines] is a much closer approximation of the true feasible region.

where $b_{s,i}^{(k)}, b_{c,i}^{(k)} \in \mathbb{R}$. Further details on the optimization and dimensionality reduction procedure are given in Appendix C, which may be of independent interest for linear programming optimization problems with a large number of affine constraints.

Crucially, if the waveform in Eq. (35) satisfies this set of reduced linear constraints, then it also satisfies the set of original constraints in Eq. (37). That is, the feasible region generated by the reduced linear constraints is a subset of the true feasible region. A comparison of the feasible regions generated by the box constraints and the reduced linear constraints along with the true feasible region for $K = 1$ is shown in Fig. 1. In this example, the $|I| = 12$ reduced linear constraints generate a feasible region that closely approximates the true feasible region generated by the $N = 40\,000$ constraints in Eq. (37), a dramatic improvement over the box constraints. This advantage continues into higher dimensions with $K > 1$.

In addition to the waveform amplitude being bounded by the maximum Rabi rate, the control must generate a net identity for Eq. (24) to hold. We can ensure that the control implements an identity gate by requiring that the amplitude waveform integrates to zero, i.e., $\int_0^T dt\, \Omega(t) = 0$. For our piecewise-constant waveforms in Eq. (35), this condition is imposed by

$$\sum_{k=0}^{K-1} \left[\Omega_s^{(k)} c_s^{(k)} + \Omega_c^{(k)} c_c^{(k)} \right] = 0, \qquad (39)$$

where

$$c_s^{(k)} = \sum_{m=0}^{N-1} \sin[\omega_0 m \Delta t] v_m^{(k)}(N, W),$$

$$c_c^{(k)} = \sum_{m=0}^{N-1} \cos[\omega_0 m \Delta t] v_m^{(k)}(N, W).$$

We further add a nonlinear constraint that the dc component of the dephasing FF is exactly zero, $F_Z(\omega = 0, \vec{\Omega}, T) = 0$, which eliminates the contribution of detuning error in Eq. (26). This constraint can be discretized and expanded as

$$0 = \left| \sum_{i=0}^{N-1} \sin \Theta_i \right|^2 + \left| \sum_{i=0}^{N-1} \cos \Theta_i \right|^2, \qquad (40)$$

where

$$\Theta_i = \Delta t \sum_{m=0}^{i-1} \sum_{k=0}^{K-1} v_m^{(k)}(N, W) \big\{ \Omega_c^{(k)} \cos[\omega_0 m \Delta t]$$
$$+ \Omega_s^{(k)} \sin[\omega_0 m \Delta t] \big\}.$$

Equations (39) and (40) add two additional constraints to the $|I|$ linear constraints required to bound the Rabi rate.

We implemented the nonlinear optimization using parameters relevant to superconducting transmon qubits. In Eq. (35), we took $N = 20\,000$ and $\Delta t = 5$ ns, so the total duration of the waveform is $T = N \Delta t = 100\ \mu$s. For the DPSS, we selected the bandwidth parameter so $NW = 1$. Typical applications use only the first $2NW$ DPSS, as higher order Slepians are less spectrally concentrated. In our optimization, we found that using the first $2NW + 1$ DPSS or, equivalently, taking $K = 3$ in Eq. (35), produced superior results at the expense of slightly increased spectral leakage in the amplitude FF. Using a gradient-based interior-point solver [34–36], we minimized the objective function $I_z(\vec{\Omega}, T)$ over $\vec{\Omega}$ subject to the linear constraints in Eqs. (38) and (39). We performed this optimization for 200 different modulation frequencies $\omega_0 \in \{2\pi/T, 2 \cdot 2\pi/T, \ldots, 200 \cdot 2\pi/T\}$, so $\omega_0/2\pi$ ranged from 0.01 MHz to 2.00 MHz. For each frequency ω_0, the modulated DPSS basis functions in Eq. (35) define a new family of parameterized waveforms, a new feasible region (and reduced feasible region), and a new optimization problem:

$$\min_{\Omega_c^{(k)}, \Omega_s^{(k)}} \quad 2 \int_0^\infty d\omega\, F_Z(\omega, \vec{\Omega}, T) \frac{1}{\omega + \delta\omega}$$

$$\text{s.t.} \quad \forall i \in I \quad \sum_{k=0}^{K-1} \left(b_{s,i}^{(k)} \Omega_s^{(k)} + b_{c,i}^{(k)} \Omega_c^{(k)} \right) \leqslant \Omega_{\max},$$

$$\sum_{k=0}^{K-1} \left(\Omega_s^{(k)} c_s^{(k)} + \Omega_c^{(k)} c_c^{(k)} \right) = 0,$$

$$F_Z(\omega = 0, \vec{\Omega}, T) = 0. \qquad (41)$$

A sample waveform resulting from this optimization at modulation frequency $\omega_0/2\pi = 0.1$ MHz is displayed in Fig. 2 along with the associated amplitude and dephasing FFs. For comparison, Fig. 2 also shows the FFs generated by the $k = 0$ DPSS (Slepian) waveform in Eq. (34) combined with
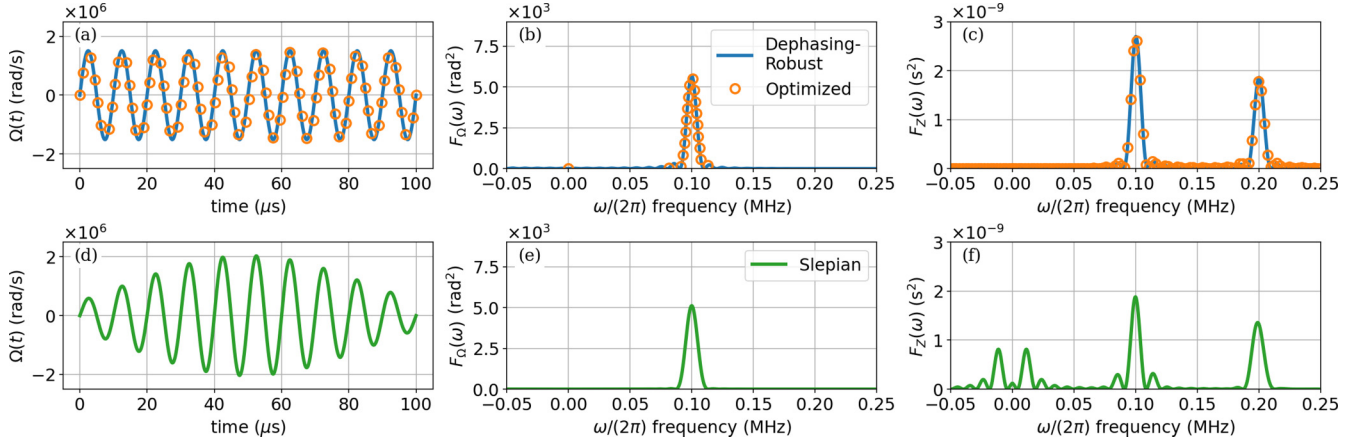
FIG. 2. Comparison of waveforms and FFs for amplitude noise QNS. The optimized waveforms (orange circles) and dephasing-robust waveforms corresponding to the first root of J_0 (blue solid line) are plotted (a) with their associated amplitude FFs (b) and dephasing FFs (c). The optimized and dephasing-robust waveforms and FFs almost completely overlap. Shown in (d) is the $k = 0$ DPSS (Slepian) waveform for $N = 2 \times 10^4$ and $NW = 1$ along with the corresponding amplitude (e) and dephasing (f) FFs. All waveforms have total time duration $T = 100\,\mu\text{s}$ and are modulated at frequency $\omega_0/2\pi = \lambda/2\pi = 0.1\,\text{MHz}$ to shift the passband of the amplitude FF.

sine modulation at frequency $\omega_0/2\pi = 0.1\,\text{MHz}$. The amplitude FFs of both the optimized [Fig. 2(b)] and the Slepian [Fig. 2(e)] are spectrally concentrated in passbands centered at 0.1 MHz. Although the optimized amplitude FF displays more spectral leakage in the form of oscillatory sidelobes outside the passband, the optimized dephasing FF in Fig. 2(c) is largely suppressed at frequencies below 0.1 MHz. This is a substantial improvement over the dephasing FF generated by the Slepian waveform, which has significant support at low frequencies. This demonstrates the effectiveness of the optimization in reducing the low-frequency components of the dephasing FF.

C. Analytic dephasing-robust waveforms

The optimized waveform for $\omega_0/2\pi = 0.1\,\text{MHz}$ obtained in the previous section and plotted in Fig. 2(a) appears to be a simple oscillatory function with a relatively constant amplitude. This pattern continues over a range of modulation frequencies. Specifically, we find that the numerically optimized waveforms are well-approximated by analytic functions, which we term dephasing-robust waveforms. These waveforms are defined by

$$\Omega_{\text{d-r}}(t) \equiv \Omega_0 \sin(\lambda t), \tag{42}$$

where the amplitude Ω_0 and control modulation frequency λ are constrained to take particular values that enforce suppression of low-frequency dephasing noise, namely, Ω_0 and λ must satisfy

$$\lambda = \frac{2\pi M}{T}, \tag{43}$$

$$J_0\left(\frac{\Omega_0}{\lambda}\right) = 0, \tag{44}$$

where M is an integer, T is the total duration of the waveform, and J_0 is the order-0 Bessel function of the first kind. Interestingly, a sinusoidal driving field satisfying these same conditions was employed in Ref. ([37]) to create dressed

qubits robust to detuning and static amplitude offsets. The first condition guarantees that the dephasing-robust waveforms generate a net identity over the time interval $[0, T]$, as $\int_0^T dt\, \Omega_{\text{d-r}}(t) = 0$. The second condition is equivalent to requiring that the ratio Ω_0/λ be a root of J_0. The Bessel function J_0 has infinitely many roots, with the first three being approximately 2.40, 5.52, and 8.65. As we will show, the first and second conditions imply that $F_Z(0, T) = 0$, ensuring that the mean-dependent contributions in Eqs. (26) and (29) vanish. Figure 2(a) shows that the dephasing-robust waveform corresponding to the first root is nearly identical to the waveform obtained by numerical optimization.

As desired for amplitude noise QNS, the amplitude FFs generated by the dephasing-robust waveforms are localized in tunable passbands set by λ. The amplitude FF follows from Eq. (19) and takes the form

$$F_\Omega(\omega, T) = \left(\frac{\Omega_0 \lambda \sin(\omega T/2)}{\omega^2 - \lambda^2}\right)^2, \tag{45}$$

which is spectrally concentrated in the passband $B_{\text{d-r}}(\lambda) = [\lambda - 2\pi/T, \lambda + 2\pi/T]$.

The most notable property of the dephasing-robust waveforms is that spectral concentration of the amplitude FF is preserved, while low-frequency components of the dephasing FF are suppressed. Using the periodic structure of the dephasing-robust waveform, we show in Appendix D that the dephasing filter can be expressed as

$$F_Z(\omega, T) = \frac{\sin^2(M\pi\omega/\lambda)}{\sin^2(\pi\omega/\lambda)}\left[\left|\int_0^{2\pi/\lambda} dt \cos\Theta(t)e^{-i\omega t}\right|^2 + \left|\int_0^{2\pi/\lambda} dt \sin\Theta(t)e^{-i\omega t}\right|^2\right]. \tag{46}$$
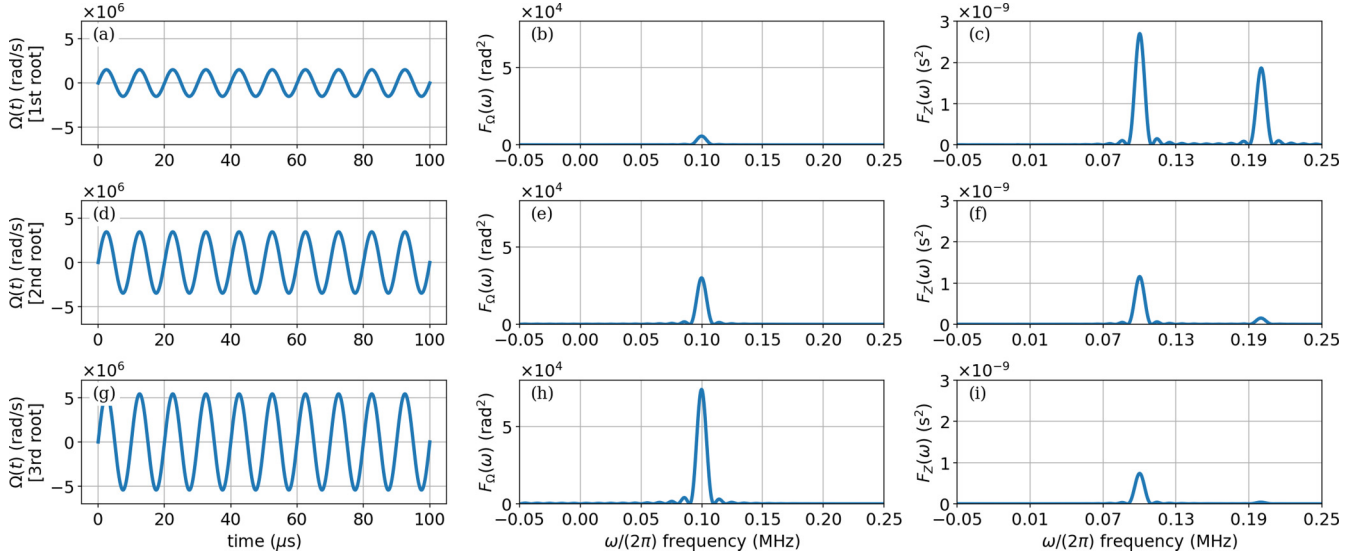
FIG. 3. Leading order dephasing-robust waveforms and FFs. Dephasing-robust waveform solutions corresponding to the first (top row), second (middle row), and third (bottom row) roots of J_0, all with modulation frequency $\lambda = 0.10$ MHz and total time duration $T = 100\,\mu s$. Plotted from left to right are the amplitude control waveforms [(a), (d), (g)], the amplitude FFs [(b), (e), (h)], and dephasing FFs [(c), (f), (i)]. Note that the dephasing FFs are suppressed at frequencies below λ.

The significance of Eq. (44) becomes apparent when we examine this expression at $\omega = 0$, noting that

$$\int_0^{2\pi/\lambda} dt\, \cos\Theta(t) = \frac{2\pi \cos(\Omega_0/\lambda)J_0(\Omega_0/\lambda)}{\lambda}, \quad (47)$$

$$\int_0^{2\pi/\lambda} dt\, \sin\Theta(t) = \frac{2\pi \sin(\Omega_0/\lambda)J_0(\Omega_0/\lambda)}{\lambda}. \quad (48)$$

The condition $J_0(\Omega_0/\lambda) = 0$, thus, guarantees $F_Z(0, T) = 0$, which cancels detuning error.

The term immediately to the right of the equality in Eq. (46), which is known as the Fejér kernel [38], is a periodic function with peaks centered at integer multiples of λ. For fixed λ, the height of the peaks increases and the width of the peaks decreases with increasing M. When $M \gg 1$, the Fejér kernel is approximated by a frequency comb or sum of delta functions centered at integer multiples of λ [16]. In this regime, we show in Appendix D that the dephasing FF takes the form

$$F_Z(\omega, T) \approx 2\pi T \sum_{k \in \mathbb{Z}} \delta(\omega - k\lambda) \left| J_k\left(\frac{\Omega_0}{\lambda}\right) \right|^2, \quad (49)$$

where J_k is the order-k Bessel function of the first kind. Since $F_Z(0, T) = 0$, the dephasing FF is a high-pass filter that suppresses low frequency dephasing noise up to the cutoff frequency λ.

Along with the formal derivation of Eq. (49) presented in Appendix D, there is an intuitive explanation for the structure of the dephasing FF based on an analogy with DD sequences. We note that the dephasing-robust waveforms in Eq. (42) consist of an alternating train of half-sine window pulses, which for the lowest Bessel root correspond to a rotation of approximately $\pm 1.53\pi$. This is reminiscent of DD sequences with finite-width $\pm\pi$ pulses [39,40]. Using Eq. (20), the zero-frequency component of the dephasing FF can be expressed

as

$$F_Z(\omega = 0, T) = \left| \int_0^T dt\, \exp i\Theta(t) \right|^2. \quad (50)$$

For an alternating train of $\pm\pi$ finite-width pulses, $\Theta(t)$ smoothly transitions between $[0, \pi]$, meaning that $\exp i\Theta(t)$ is always within the first two quadrants of the complex plane. As such, $\exp i\Theta(t)$ has a positive imaginary component for all t and $F_Z(\omega = 0, T)$ in Eq. (50) will be nonzero. Over-rotating the pulses beyond π, as is done by the dephasing-robust waveforms, allows $\Theta(t)$ to take values beyond $[0, \pi]$ and $\exp i\Theta(t)$ to take values in all four quadrants of the complex plane. With a carefully chosen over-rotation, contributions from quadrants III and IV cancel those from quadrants I and II, which causes the dephasing FF to vanish at zero frequency. With the dc component of the dephasing FF removed, the structure of the waveform as an evenly spaced alternating pulse train produces a high-pass filter similarly to dynamical decoupling sequences.

D. Comparison of waveforms

Since the dephasing-robust and optimized waveforms are nearly identical for the first root of J_0, it is unsurprising that the amplitude and dephasing FFs in Figs. 2(b) and 2(c) also closely match. Figure 3 shows the dephasing-robust waveforms and FFs for the first three roots of J_0 with $\lambda/2\pi = 0.1$ MHz. For all roots, the amplitude FFs are spectrally concentrated about λ. The high-pass nature of the dephasing FF with λ as the low-frequency cutoff is also evident for all roots. Interestingly, the high-frequency support of $F_Z(\omega, T)$ beyond λ differs between the roots with the second nonzero harmonic at 2λ being considerably smaller for the second and third roots. Importantly, as evidenced by the low-frequency support of $F_Z(\omega, T)$ for the DPSS waveforms and the absence of low-frequency support of $F_Z(\omega, T)$
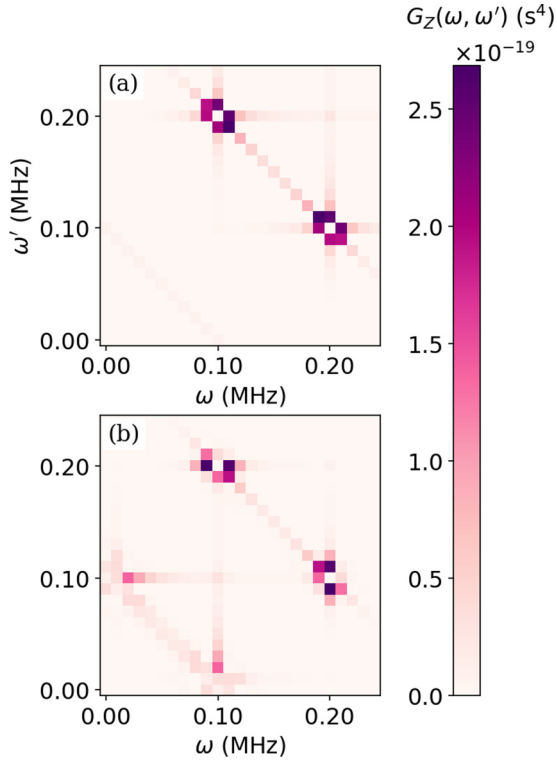
FIG. 4. The higher-order dephasing filter function $G_z(\omega, \omega', T)$ for the dephasing-robust (a) and Slepian (b) waveforms with $\lambda = 0.10$ MHz and $T = 100 \, \mu$s. The unwanted peak at $(0, \lambda)$ in the bottom-left corner of the Slepian FF demonstrates a sensitivity to a combination of low-frequency and static dephasing, while the absence of this peak for the dephasing-robust indicates that the waveforms are robust against higher-order low-frequency dephasing noise.

for the dephasing-robust waveforms, the three leading-order roots offer superior suppression of low-frequency dephasing noise.

While we have thus far focused on the noise filtering properties of $F_Z(\omega, T)$, the strength of the dephasing-dependent higher order terms in Eq. (24) also depends on the higher order FF $G_Z(\omega, \omega', T)$ defined in Eq. (30). Unlike $F_Z(\omega, T)$, $G_Z(\omega, \omega', T)$ forms a two-dimensional surface in the frequency domain. The higher-order perturbative term in Eq. (29) illustrates that low-frequency dephasing noise is suppressed when $G_Z(\omega, \omega', T)$ has minimal low-frequency support over both ω and ω'. Similarly, detuning error is suppressed when $G_Z(\omega, \omega', T) = 0$ at $\omega = 0$ and $\omega' = 0$. Figures 4(a) and 4(b), respectively, show $G_Z(\omega, \omega', T)$ generated by the dephasing-robust waveform in Fig. 2(a) and the DPSS waveform in Fig. 2(d). To produce these plots, we evaluated the quadruple integral that defines $G_Z(\omega, \omega', T)$ using an efficient iterated fast Fourier transform procedure described in Appendix E. For the dephasing-robust waveform, $G_Z(\omega, \omega', T)$ is dominated by two peaks at $(\omega, \omega') = (\lambda, 2\lambda)$ and $(\omega, \omega') = (2\lambda, \lambda)$ and has minimal spectral support when $\omega, \omega' < \lambda$. Similar to $F_Z(\omega, T)$, the modulation frequency λ acts as a cutoff below which low-frequency dephasing noise is suppressed. The DPSS waveform, on the other hand, has more support at low and dc frequency with additional peaks

at $(\omega, \omega') = (0, \lambda)$ and $(\omega, \omega') = (\lambda, 0)$, making it vulnerable to detuning error when the dephasing noise spectrum has support at $\omega \approx \lambda$. The high-pass nature of $G_Z(\omega, \omega', T)$ for the dephasing-robust waveforms again translates into improved filtering of low-frequency dephasing noise and detuning error.

IV. APPLICATION TO AMPLITUDE NOISE QNS

The dephasing-robust waveforms closely match the results of our numerical optimization and, therefore, offer near-optimal suppression of low-frequency dephasing and detuning error while at the same time preserving spectral concentration in the amplitude FF. The high-pass filtering characteristics of the FFs $F_Z(\omega, T)$ and $G_Z(\omega, \omega', T)$ ensure that the dephasing-robust waveforms minimize the dynamical contributions of both low-frequency dephasing and detuning error. The resulting reduction in dephasing-induced bias ensures that the tomographic measurement strategy can reliably isolate the dynamical contribution of amplitude noise. To quantify the practical difference this makes for noise characterization, we numerically simulate amplitude noise QNS in the presence of low-frequency dephasing noise and detuning errors using both the dephasing-robust and DPSS waveforms.

To assess the robustness of the waveforms in the presence of low-frequency dephasing versus detuning, we implement two different numerical QNS experiments in which the qubit evolves under amplitude noise and either (1) static detuning error or (2) low-frequency dephasing. To fairly compare the performance of the waveforms for a range of modulation frequencies, we consider a flat amplitude noise spectrum with a high-frequency cutoff ω_h,

$$S_\Omega(\omega) = \begin{cases} A_\Omega, & \omega \leqslant \omega_h \\ 0, & \omega > \omega_h. \end{cases}$$

Static detuning error is introduced by taking $\beta_z(t) = \Delta$ in the noise Hamiltonian of Eq. (2). For the dephasing, we simulate noise with a $1/f$ spectrum, characteristic of solid state devices [41], with both low- and high-frequency cutoffs for efficient numerical simulation,

$$S_z(\omega) = C \begin{cases} \frac{A_z}{\omega_l}, & |\omega| \leqslant \omega_l \\ \frac{A_z}{\omega}, & \omega_l < |\omega| \leqslant \omega_h \\ 0, & |\omega| > \omega_h. \end{cases}$$

Here, the low-frequency cutoff ω_l ensures that the strength of the dephasing noise is finite at $\omega = 0$ and the dimensionless scale factor C determines the noise strength. In our simulations, $A_\Omega = 1.04 \times 10^{-11}$ rad^2/Hz, $\omega_h/2\pi = 2$ MHz, $\omega_l/2\pi = 0.01$ MHz, and $A_z = 10^8$ Hz^2. While this value of ω_l is large for $1/f$ noise in a solid-state device, the resulting dephasing noise spectrum is still dominated by contributions at low frequency. To examine the impact of the dephasing noise strength on the reconstruction, we vary the scale factor C between 299.1 and 3.18. The resulting $1/e$ decay time under free evolution (T_2^*) ranges from 4 μs to 100 μs. The detuning error Δ varies between 0.01 MHz and 0.19 MHz.

For control, we employ the dephasing-robust and DPSS waveforms similar to those depicted in Fig. 2 with an

expanded range of modulation frequencies. Concretely, since typical laboratory control hardware generates piecewise constant waveforms, we discretize the dephasing robust waveform in Eq. (42) into N increments of duration Δt so the resulting amplitude waveform is

$$\Omega(t) = \begin{cases} 0, & t \in [0, \Delta t) \\ \Omega_0 \sin(\lambda \Delta t), & t \in [\Delta t, 2\Delta t) \\ \vdots & \vdots \\ \Omega_0 \sin[\lambda(N-1)\Delta t], & t \in [(N-1)\Delta t, N\Delta t). \end{cases} \tag{51}$$

For the waveform parameters, we use $N = 10\,000$, $\Delta t = 10\,\text{ns}$, and $T = N\Delta t = 100\,\mu\text{s}$. The modulation frequency takes values $\lambda \in \{\Delta\omega, \ldots, L\Delta\omega\}$ with $\Delta\omega/2\pi = 0.01\,\text{MHz}$ and $L = 200$. If we use only the dephasing-robust waveforms generated by the first root of J_0, we would need to vary the amplitude Ω_0 so $\Omega_0/\lambda \approx 2.40$ for all λ. Instead, we use up to the 159th root of J_0, using larger roots for smaller λ, so $\Omega_0/2\pi \approx 5\,\text{MHz}$ for all modulation frequencies.

For the DPSS, we use waveforms of the lowest ($k = 0$) order combined with sine modulation:

$$\Omega(t) = \begin{cases} 0, & t \in [0, \Delta t) \\ \Omega_{\max} v_1^{(0)}(N, W) \sin(\lambda \Delta t), & t \in [\Delta t, 2\Delta t) \\ \vdots & \vdots \\ \Omega_{\max} v_{N-1}^{(0)}(N, W) \sin[\lambda(N-1)\Delta t], & t \in [(N-1)\Delta t, N\Delta t). \end{cases} \tag{52}$$

The $k = 0$ DPSSs, which are the most spectrally concentrated, generate a net identity gate when combined with sine modulation since

$$v_n^{(0)}(N, W) = v_{N-n}^{(0)}(N, W),$$
$$\sin(\lambda n \Delta t) = -\sin[\lambda(N-n)\Delta t],$$

ensuring $\int_0^T dt\,\Omega(t) = 0$. In Eq. (52), $\Omega_{\max}/2\pi = 5\,\text{MHz}$, while the other parameters are identical to the dephasing-robust waveform, i.e., $N = 10,000$, $\Delta t = 10\,\text{ns}$, $T = N\Delta t = 100\,\mu\text{s}$, and $\lambda \in \{\Delta\omega, \ldots, L\Delta\omega\}$, with $\Delta\omega/2\pi = 0.01\,\text{MHz}$ and $L = 200$. We select the bandwidth parameter W of the DPSS so $NW = 1$, which creates an amplitude FF concentrated in the passband $B_{\text{DPSS}}(\lambda) = [\lambda - 2\pi/T, \lambda + 2\pi/T]$, equivalent to the passband of a dephasing-robust waveform at modulation frequency λ. Although our simulations use DPSSs and dephasing-robust waveforms with the same passbands, the DPSS waveforms are more spectrally concentrated with a spectral concentration ratio (see Appendix B) of 0.981. The spectral concentration of the dephasing-robust waveforms is 0.904, for comparison.

To reconstruct the amplitude noise spectrum, we use an inversion approach similar to Ref. [32]. By discretizing the overlap integral in Eq. (17) into increments $\Delta\omega$ up to the maximum frequency $L\Delta\omega$, we obtain

$$\langle a_1^{(1)}(T)^2 \rangle \approx \frac{1}{\pi} \sum_{\ell=1}^{L} \int_{(\ell-\frac{1}{2}-\frac{1}{2}\delta_{\ell,1})\Delta\omega}^{(\ell+\frac{1}{2})\Delta\omega} d\omega\, F_\Omega(\omega, T)\, S_\Omega(\omega)$$
$$\approx \vec{F}_\Omega(T) \cdot \vec{S}_\Omega, \tag{53}$$

where $\vec{F}_\Omega(T)$ and $\vec{S}_\Omega$ are L-dimensional vectors with elements

$$[\vec{F}_\Omega(T)]_{\ell=1} = \frac{1}{\pi} \int_0^{\frac{3}{2}\Delta\omega} d\omega\, F_\Omega(\omega, T),$$

$$[\vec{F}_\Omega(T)]_{\ell>1} = \frac{1}{\pi} \int_{(\ell-\frac{1}{2})\Delta\omega}^{(\ell+\frac{1}{2})\Delta\omega} d\omega\, F_\Omega(\omega, T),$$

$$[\vec{S}_\Omega]_\ell = S_\Omega(\ell\Delta\omega) \tag{54}$$

for $\ell \in \{1, \ldots, L\}$. Note that the integration region defining the first element of $\vec{F}_\Omega$ in Eqs. (54) is larger than the subsequent regions to ensure it encloses the peak of the amplitude FF centered at $\lambda = \Delta\omega$. Equation (53) shows that measuring the $\langle a_1^{(1)}(T)^2 \rangle$ generated by the waveforms at each modulation frequency $\lambda \in \{\Delta\omega, \ldots, L\Delta\omega\}$ produces a linear system that can be inverted to obtain $\vec{S}_\Omega$. Since $S_\Omega(\omega)$ is non-negative, the reconstruction problem is an instance of non-negative least-squares regression, which is solvable through standard numerical techniques.

In Fig. 5, we use this reconstruction technique to compare the performance of the dephasing-robust and DPSS waveforms in the presence of detuning error of variable strength. The accuracy of the DPSS reconstruction in Fig. 5(b) clearly degrades as the strength of the detuning error increases. In contrast, the dephasing-robust reconstruction in Fig. 5(a) remains relatively accurate. The difference in performance between the two waveforms is directly related to the higher order terms that bias the tomographic measurement strategy in Eq. (31), as we demonstrate in Fig. 6. Recall that the tomographic measurement strategy approximates the variance $\langle \Delta a_1^{(1)}(T)^2 \rangle$ as a linear combination of survival probabilities, $\mathcal{P} \equiv \frac{1}{2}[1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)]$. Figure 6(a) shows that for the DPSS waveforms, $\langle \Delta a_1^{(1)}(T)^2 \rangle$ and $\mathcal{P}$ diverge as the detuning error increases. For the dephasing-robust waveforms, $\langle \Delta a_1^{(1)}(T)^2 \rangle$ and $\mathcal{P}$ are closely matched over the range of detuning errors. This occurs because the dephasing-robust waveforms reduce the magnitudes of the higher-order detuning-dependent terms as compared to the DPSS, as shown in Figs. 6(c) and 6(d). While the dephasing-robust waveforms successfully mitigate
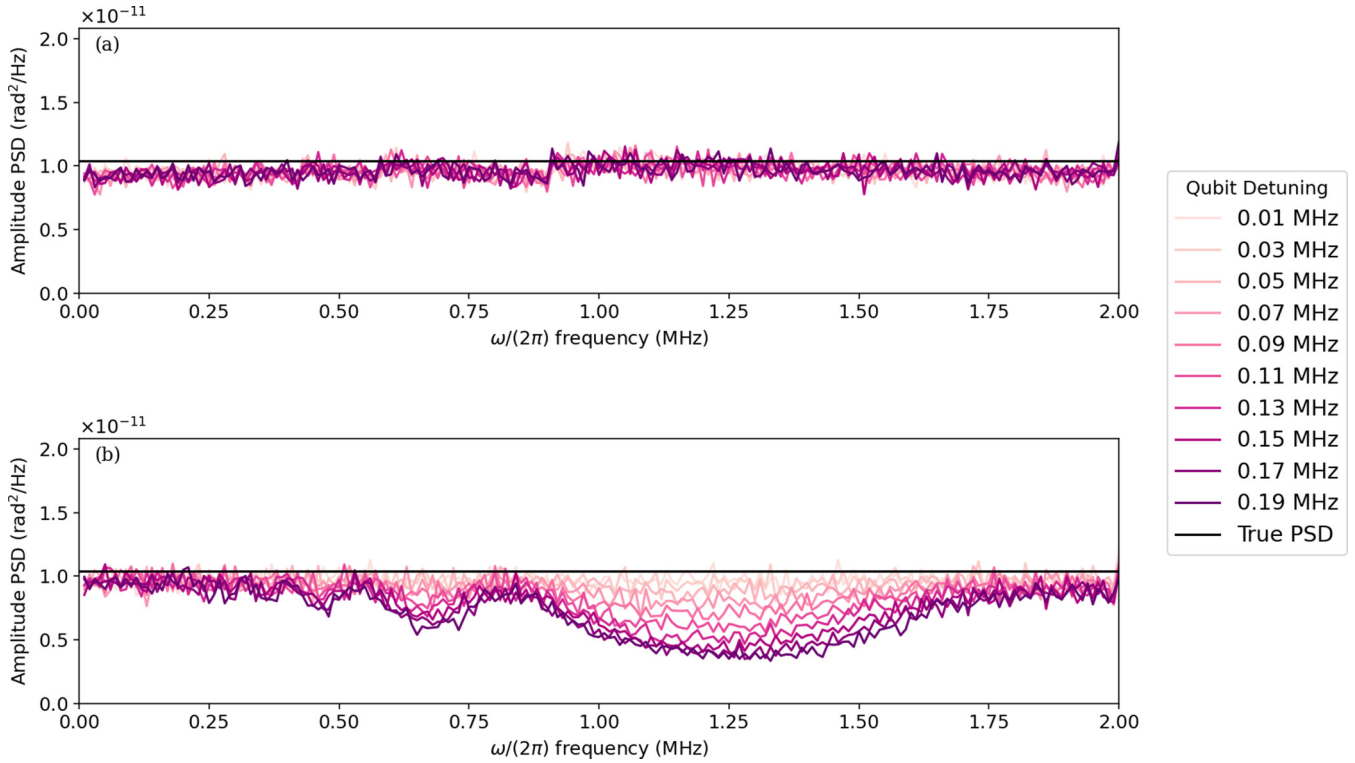
FIG. 5. Amplitude noise QNS with detuning errors of different magnitudes. Reconstructions of the amplitude noise spectrum are depicted for the dephasing-robust waveforms (a) and the DPSS waveforms (b) in the presence of detuning errors ranging from $\Delta = 0.01$ MHz (light pink) to 0.19 MHz (dark purple).
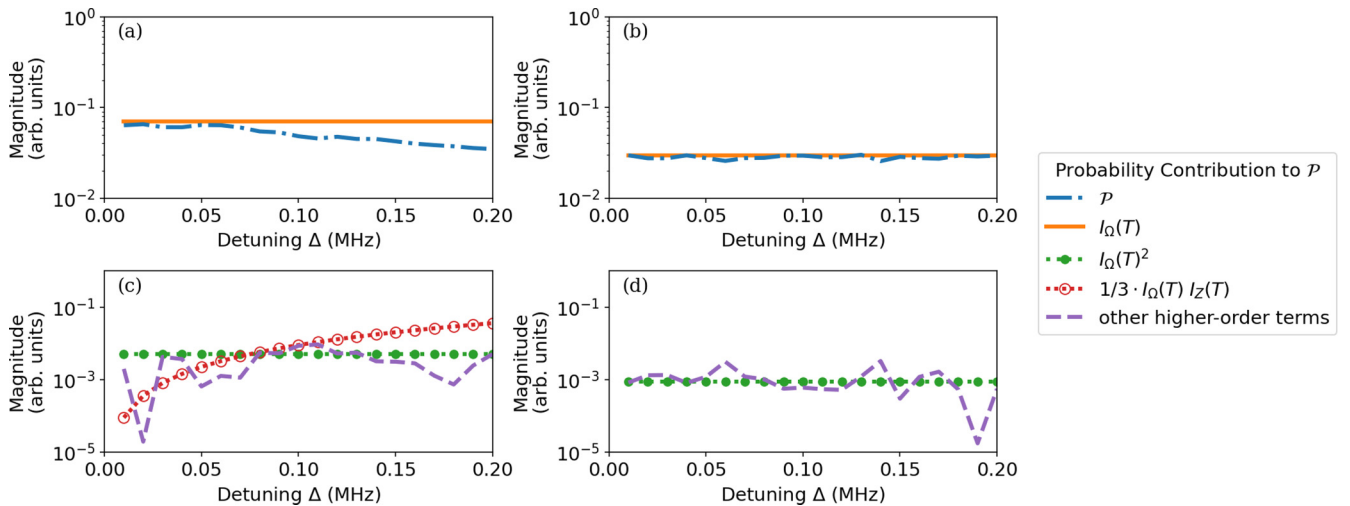

FIG. 6. Contributions to the survival probability in Eq. (24) with detuning error of various strengths for the DPSS (a), (c) and dephasing-robust waveforms (b), (d). Both the DPSS and dephasing-robust waveforms have modulation frequency $\lambda/2\pi = 1.00$ MHz with all other parameters identical to those used for the reconstructions in Fig. 5. Plots (a) and (b) show the discrepancy between $\mathcal{P} \equiv \frac{1}{2}[1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)]$ (orange, solid line) and $I_\Omega(T)$ (blue, dot-dashed line) for the DPSS waveform and the dephasing-robust waveform, respectively. The discrepancy is determined by the higher-order dephasing-dependent contributions to Eq. (24), which are plotted in (c) for the DPSS waveform and in (d) for the dephasing-robust waveform. The higher-order amplitude term $I_\Omega(T)^2$ is plotted as green solid circles, the amplitude-dephasing cross term $1/3 \cdot I_\Omega(T) I_Z(T)$ is plotted as red hollow circles, and the remaining contributions including higher-order dephasing and statistical jitter are plotted as dashed purple lines. For the dephasing-robust waveform, the $1/3 \cdot I_\Omega(T) I_Z(T)$ contribution is exactly canceled because the dc component of the dephasing FF vanishes, i.e., $F_Z(0, T) = 0$.
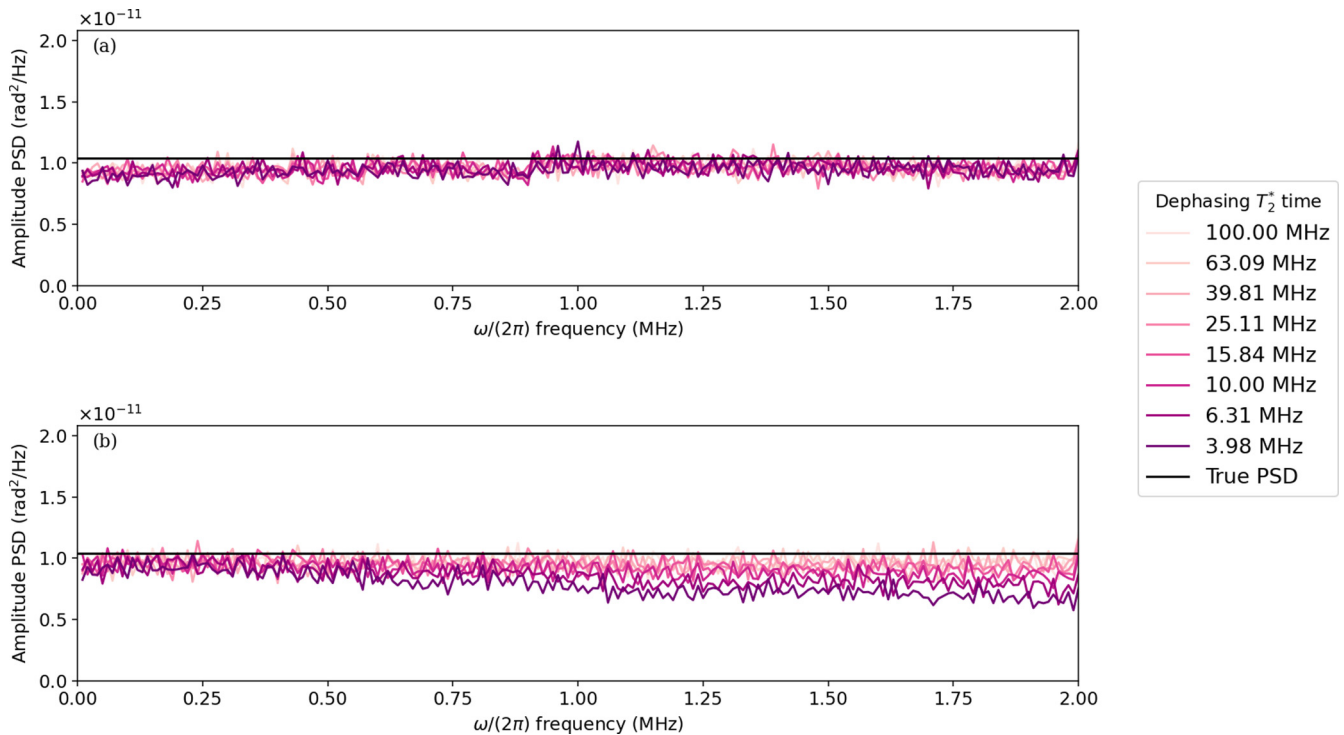
FIG. 7. Amplitude noise QNS with variable strength dephasing. Reconstruction of the amplitude noise spectrum using the dephasing-robust waveforms (a) and the DPSS waveforms (b) in the presence of dephasing noise with T_2^* times ranging from $4-100$ μs (dark purple to light pink).

the higher-order detuning-dependent terms, the amplitude noise-dependent term $I_\Omega(T)^2$ is one of the few higher-order contributions that remains. This term largely accounts for the slight difference between the reconstruction and true amplitude noise spectrum in Fig. 5(a).

Figure 7 compares the performance of the dephasing-robust and DPSS waveforms in the presence of low-frequency dephasing noise of variable strength. The reconstruction is carried out for dephasing noise with a range of T_2^* times differing by over an order of magnitude, with smaller T_2^* indicating stronger noise. Due to bias from low-frequency dephasing, the DPSS reconstruction deviates from the true amplitude noise spectrum for the lowest T_2^* times. Interestingly, the Slepian reconstruction also exhibits more errors at higher modulation frequencies. Because of their filtering properties, the dephasing-robust waveforms produce an estimate of the amplitude noise spectrum that remains relatively accurate for all strengths of dephasing noise and modulation frequencies. Note, however, that the DPSS reconstruction is only qualitatively affected by the dephasing noise when the T_2^* time is extremely short relative to the experiment duration of $T = 100$ μs. As such, we conclude detuning error is more detrimental to amplitude noise QNS than pure $1/f$ dephasing noise.

While this analysis focuses on the relationship between the strength of dephasing noise or detuning error and the higher-order terms that bias amplitude noise QNS, it is important to note that the strength of the amplitude noise plays an important role as well as the interplay between dephasing and detuning. Recall that Eq. (31), which shows the higher-order corrections to the tomographic measurement strategy, demonstrates a negative multiplicative bias,

$I_\Omega(T)I_Z(T)/3$, dependent on $F_Z(\omega, T)$ as well as a positive additive bias dependent on $G_Z(\omega, \omega', T)$. The observation that the DPSS waveforms underestimate the true amplitude noise spectrum in Figs. 5(b) and 7(b), as well as the relative size of $I_\Omega(T)I_Z(T)/3$ in Fig. 6(c), indicate that the negative bias was the dominant source of error in these simulations. As the negative bias is proportional to $I_\Omega(T)$, we expect it to introduce the same proportionate error for a wide range of amplitude noise strengths. The additive bias due to $G_Z(\omega, \omega, T)$, in contrast, would become more significant for weaker amplitude noise strengths. Additionally, because certain terms contributing to the additive bias in Eq. (30) are only nonzero when both dephasing noise and detuning errors are present, we may observe different qualitative effects in this scenario. These effects are likely to be much less impactful for the dephasing-robust waveforms because the dc components of $G_Z(\omega, \omega', T)$ are largely suppressed, as shown in Fig. 4.

V. CONCLUSION

In this paper, we used a combination of numerical and analytical methods to determine a family of control waveforms that enable accurate spectroscopy of amplitude control noise in the presence of strong, low-frequency dephasing noise and detuning errors. These waveforms are unique in that they suppress low-frequency dephasing noise and detuning errors while simultaneously producing a spectrally concentrated amplitude FF. To establish the mechanism through which dephasing and detuning can bias estimates of the amplitude noise spectrum, we extended the theoretical framework of Refs. [23,27] to include both zero-mean noise and higher order perturbative terms. This higher order perturbative

analysis enabled us to formulate an objective function quantifying the dynamical contribution of dephasing and cast the task of minimizing dephasing-dependent bias as an optimal control problem. Our numerical optimization utilized techniques based on linear programming to scalably incorporate realistic constraints on control waveforms, such as limitations on the maximum amplitude. The numerically optimized waveforms closely matched the family of analytic dephasing-robust waveforms. Analysis of the primary and higher-order dephasing FFs verified that the dephasing-robust waveforms filter the dynamical contributions of low-frequency dephasing noise and detuning error.

In numerically simulated amplitude noise QNS experiments, the filtering properties of the dephasing-robust waveforms enabled accurate characterization of amplitude noise in the presence of low-frequency dephasing noise and detuning errors. In the case of detuning errors, the dephasing-robust robust waveforms offered significant improvement over existing protocols based on DPSS waveforms. For $1/f$ dephasing noise, on the other hand, the performance of dephasing-robust and DPSS waveforms was comparable for all but the lowest T_2^* times we considered. The performance gains offered by the dephasing-robust waveforms in the presence of a combination of detuning error and dephasing noise, or dephasing noise with a spectrum other than $1/f$, is an area of future study.

Another extension of this paper involves our linear programming-inspired optimization procedure. The waveform and amplitude constraints that our method scalably incorporates are relevant to a broad range of laboratory control hardware and more general optimal control problems. Immediate applications include filter design for noise mitigation and QNS in other settings with bounded-strength controls.

APPENDIX A: MEASUREMENTS

1. Survival probabilities

Consider an experiment in which the qubit is prepared in $|\uparrow_i\rangle$ and evolves under the rotating frame Hamiltonian $H(t) = H_N(t) + H_{\text{ctrl}}(t)$ from $t = 0$ to $t = T$. At time T, the qubit is measured projectively in the $|\uparrow_i\rangle$, $|\downarrow_i\rangle$ basis, with $|\uparrow_i\rangle$, $|\downarrow_i\rangle$ denoting the $+1$, -1 eigenstates of σ_i, respectively. If this experiment is repeated, the expected fraction of experiments in which $|\uparrow_i\rangle$ is measured is given by the survival probability,

$$\mathbb{P}(\uparrow_i, T) = \langle \langle \uparrow_i | U(T) | \uparrow_i \rangle^2 \rangle, \tag{A1}$$

where $\langle \cdot \rangle$ is the ensemble average over realizations of the dephasing and amplitude noise. If the amplitude waveform in Eq. (9) generates an identity gate over time $t \in [0, T]$, the rotating frame and toggling frame propagators at time T are equivalent, i.e., $U(T) = U_c(T)\tilde{U}(T) = \tilde{U}(T)$. By substituting $U(T) = \tilde{U}(T)$ in Eq. (A1) and using the parametrization of $\tilde{U}(T)$ in terms of the error vector $\vec{a}(T) \equiv [a_1(T), a_2(T), a_3(T)]$ in Eq. (7), we obtain

$$\mathbb{P}(\uparrow_i, T) = \langle \langle \uparrow_i | \tilde{U}(T) | \uparrow_i \rangle^2 \rangle = \langle \langle \uparrow_i | e^{-i\vec{a}(T) \cdot \vec{\sigma}} | \uparrow_i \rangle^2 \rangle. \tag{A2}$$

2. Derivation of Eq. (23) with higher order terms

Evaluating the survival probabilities in Eq. (A2) for $i \in \{1, 2, 3\}$ and taking a Taylor expansion to fourth order in the error vector components produces

$$\mathbb{P}(\uparrow_1, T) = \left\langle \frac{a_2(T)^2 + a_3(T)^2 + 2a_1(T)^2 + [a_2(T)^2 + a_3(T)^2]\cos[2\sqrt{a_1(T)^2 + a_2(T)^2 + a_3(T)^2}]}{2[a_1(T)^2 + a_2(T)^2 + a_3(T)^2]} \right\rangle$$

$$= 1 - \langle a_2(T)^2 \rangle - \langle a_3(T)^2 \rangle + \frac{1}{3}[\langle a_2(T)^4 \rangle + \langle a_3(T)^4 \rangle + \langle a_2(T)^2 a_1(T)^2 \rangle + \langle a_3(T)^2 a_1(T)^2 \rangle + 2\langle a_2(T)^2 a_3(T)^2 \rangle], \tag{A3}$$

$$\mathbb{P}(\uparrow_2, T) = \left\langle \frac{a_1(T)^2 + a_3(T)^2 + 2a_2(T)^2 + [a_1(T)^2 + a_3(T)^2]\cos[2\sqrt{a_1(T)^2 + a_2(T)^2 + a_3(T)^2}]}{2[a_1(T)^2 + a_2(T)^2 + a_3(T)^2]} \right\rangle$$

$$= 1 - \langle a_1(T)^2 \rangle - \langle a_3(T)^2 \rangle + \frac{1}{3}[\langle a_1(T)^4 \rangle + \langle a_3(T)^4 \rangle + \langle a_2(T)^2 a_1(T)^2 \rangle + \langle a_3(T)^2 a_2(T)^2 \rangle + 2\langle a_3(T)^2 a_1(T)^2 \rangle], \tag{A4}$$

$$\mathbb{P}(\uparrow_3, T) = \left\langle \frac{a_1(T)^2 + a_2(T)^2 + 2a_3(T)^2 + [a_1(T)^2 + a_2(T)^2]\cos[2\sqrt{a_1(T)^2 + a_2(T)^2 + a_3(T)^2}]}{2[a_1(T)^2 + a_2(T)^2 + a_3(T)^2]} \right\rangle$$

$$= 1 - \langle a_2(T)^2 \rangle - \langle a_1(T)^2 \rangle + \frac{1}{3}[\langle a_2(T)^4 \rangle + \langle a_1(T)^4 \rangle + \langle a_2(T)^2 a_3(T)^2 \rangle + \langle a_3(T)^2 a_1(T)^2 \rangle + 2\langle a_1(T)^2 a_2(T)^2 \rangle]. \tag{A5}$$

From the Magnus expansion of the error propagator, $\tilde{U}(T) = e^{-i\sum_{m=1}^{\infty} M^{(m)}(T)}$, we can determine a perturbative expansion of the error vector by projecting the terms of the Magnus expansion into the Pauli basis so $\vec{a} = \sum_{m=1}^{\infty} \vec{a}^{(m)}(T)$, where $a_i^{(m)}(T) =$

$\text{Tr}[\sigma_i M^{(m)}(T)] \sim T^m$ for $i \in \{1, 2, 3\}$. The first two terms in the error vector expansion have components

$$a_i^{(1)}(T) = \frac{1}{2}\text{Tr}\left\{\sigma_i \int_0^T dt\, \tilde{H}(t)\right\}, \tag{A6}$$

$$a_i^{(2)}(T) = -\frac{i}{4}\text{Tr}\left\{\sigma_i \int_0^T dt \int_0^t dt'\, [\tilde{H}(t), \tilde{H}(t')]\right\}. \tag{A7}$$

By taking $a_i(T) = a_i^{(1)}(T) + a_i^{(2)}(T) + \dots$ and keeping terms up to order T^4 in Eqs. (A3)–(A5), we find

$$\frac{1}{2}[1 + \mathbb{P}(\uparrow_1, T) - \mathbb{P}(\uparrow_2, T) - \mathbb{P}(\uparrow_3, T)]$$

$$= \langle a_1^{(1)}(T)^2 \rangle + 2\langle a_1^{(1)}(T) a_1^{(2)}(T)\rangle + \langle a_1^{(2)}(T)^2 \rangle - \frac{1}{3}\left[\langle a_1^{(1)}(T)^4 \rangle + \langle a_2^{(1)}(T)^2\, a_1^{(1)}(T)^2 \rangle + \langle a_3^{(1)}(T)^2\, a_1^{(1)}(T)^2 \rangle\right]$$

$$= \langle \Delta a_1^{(1)}(T)^2 \rangle + \langle a_1^{(1)}(T)\rangle^2 + 2\langle a_1^{(1)}(T) a_1^{(2)}(T)\rangle + \langle a_1^{(2)}(T)^2 \rangle - \frac{1}{3}\left[\langle a_1^{(1)}(T)^4 \rangle + \langle a_2^{(1)}(T)^2\, a_1^{(1)}(T)^2 \rangle + \langle a_3^{(1)}(T)^2\, a_1^{(1)}(T)^2 \rangle\right]. \tag{A8}$$

Because they contain odd moments of $\beta_\Omega(t)$, the second and third terms vanish when the amplitude noise is zero mean. In this case, we recover Eq. (24) from the main text.

3. Explicit expressions of the terms in Eq. (A8)

The second term in Eq. (A8) arises when the amplitude noise has a nonzero mean. Its form follows in a straightforward manner from Eq. (11):

$$\langle a_1^{(1)}(T)\rangle^2 = \frac{1}{4}\left[\int_0^T dt\, \Omega(t) \langle \beta_\Omega(t)\rangle\right]^2 = \frac{\mu_\Omega^2}{4}\left[\int_0^T dt\, \Omega(t)\right]^2. \tag{A9}$$

Because the mean of a stationary process is time independent, we have taken $\mu_\Omega \equiv \langle \beta_\Omega(t)\rangle$ in the last line.

The third term in Eq. (A8), which depends on a third order moment of $\beta_\Omega(t)$ and $\beta_z(t)$, is also nonvanishing when the amplitude noise has nonzero mean. Using the Gaussianity and independence of $\beta_\Omega(t)$ and $\beta_z(t)$, we can factorize the third-order moment into a product of one- and two-point correlation functions:

$$\langle \beta_\Omega(t_1)\beta_z(t_2)\beta_z(t_3)\rangle = \langle \beta_\Omega(t_1)\beta_z(t_2)\rangle\langle \beta_z(t_3)\rangle + \langle \beta_\Omega(t_1)\beta_z(t_3)\rangle\langle \beta_z(t_2)\rangle + \langle \beta_z(t_2)\beta_z(t_3)\rangle\langle \beta_\Omega(t_1) - 2\langle \beta_\Omega(t_1)\rangle\langle \beta_z(t_2)\rangle\langle \beta_z(t_3)\rangle$$

$$= \langle \beta_z(t_2)\beta_z(t_3)\rangle\langle \beta_\Omega(t_1)\rangle.$$

Using this expression and Eq. (A7) to evaluate $a_1^{(2)}(T)$, we find

$$\langle a_1^{(1)}(T) a_1^{(2)}(T)\rangle = \frac{\mu_\Omega}{2}\left[\int_0^T dt_1\, \Omega(t_1)\right]\int_0^T dt_2 \int_0^{t_2} dt_3\, \sin[\Theta(t_2) - \Theta(t_3)]\langle \beta_z(t_2)\beta_z(t_3)\rangle. \tag{A10}$$

The remaining higher-order terms in Eq. (A8) depend on fourth-order moments of $\beta_\Omega(t)$ and $\beta_z(t)$. Like the third-order moment above, the fourth-order moments factor into products of one- and two-point correlation functions. For $i_1, i_2, i_3, i_4 \in \{\Omega, z\}$, an arbitrary fourth moment can be written as

$$\langle \beta_{i_1}(t_1)\beta_{i_2}(t_2)\beta_{i_3}(t_3)\beta_{i_4}(t_4)\rangle = \langle \beta_{i_1}(t_1)\beta_{i_2}(t_2)\rangle\langle \beta_{i_3}(t_3)\beta_{i_4}(t_4)\rangle + \langle \beta_{i_1}(t_1)\beta_{i_3}(t_3)\rangle\langle \beta_{i_2}(t_2)\beta_{i_4}(t_4)\rangle$$

$$+ \langle \beta_{i_1}(t_1)\beta_{i_4}(t_4)\rangle\langle \beta_{i_3}(t_3)\beta_{i_2}(t_2)\rangle - 2\langle \beta_{i_1}(t_1)\rangle\langle \beta_{i_2}(t_2)\rangle\langle \beta_{i_3}(t_3)\rangle\langle \beta_{i_4}(t_4)\rangle. \tag{A11}$$

From this expression, we can determine the higher order terms in the frequency domain.

Consider the fifth, sixth, and seventh terms on the right-hand side of Eq. (A8), which depend on $\vec{a}^{(1)}(T)$ and can be written as

$$\langle a_i^{(1)}(T)^2\, a_j^{(1)}(T)^2 \rangle = \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4\, y_i(t_1)y_i(t_2)y_j(t_3)y_j(t_4)\left[\langle \beta_i(t_1)\beta_i(t_2)\rangle\langle \beta_j(t_3)\beta_j(t_4)\rangle\right.$$

$$\left. + \langle \beta_i(t_1)\beta_j(t_3)\rangle\langle \beta_i(t_2)\beta_j(t_4)\rangle + \langle \beta_i(t_1)\beta_j(t_4)\rangle\langle \beta_i(t_2)\beta_j(t_3)\rangle - 2\langle \beta_i(t_1)\rangle\langle \beta_i(t_2)\rangle\langle \beta_j(t_3)\rangle\langle \beta_j(t_4)\rangle\right], \tag{A12}$$

where $i \in \{1, 2, 3\}$ and $j = 1$.

When $i \neq j$, the independence of $\beta_\Omega(t)$ and $\beta_z(t)$ implies that the last three terms cancel, leaving

$$\langle a_i^{(1)}(T)^2\, a_j^{(1)}(T)^2 \rangle_{i \neq j} = \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4\, y_i(t_1)y_i(t_2)y_j(t_3)y_j(t_4)\, \langle \beta_i(t_1)\beta_i(t_2)\rangle\langle \beta_j(t_3)\beta_j(t_4)\rangle \tag{A13}$$

$$= \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4\, y_i(t_1)y_i(t_2)y_j(t_3)y_j(t_4)\left[\langle \Delta\beta_i(t_1)\Delta\beta_i(t_2)\rangle\langle \Delta\beta_j(t_3)\Delta\beta_j(t_4)\rangle\right.$$

$$\left. + \langle \Delta\beta_i(t_1)\Delta\beta_i(t_2)\rangle\langle \beta_j(t_3)\rangle\langle \beta_j(t_4)\rangle + \langle \beta_i(t_1)\rangle\langle \beta_i(t_2)\rangle\langle \Delta\beta_j(t_3)\Delta\beta_j(t_4)\rangle + \langle \beta_i(t_1)\rangle\langle \beta_i(t_2)\rangle\langle \beta_j(t_3)\rangle\langle \beta_j(t_4)\rangle\right]. \tag{A14}$$

When $\beta_j(t) = \beta_\Omega(t)$ is zero mean, this expression reduces to

$$\langle a_i^{(1)}(T)^2 a_j^{(1)}(T)^2 \rangle_{i \neq j} = \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4 \, y_i(t_1) y_i(t_2) y_j(t_3) y_j(t_4) \left[\langle \Delta\beta_i(t_1) \Delta\beta_i(t_2) \rangle + \langle \beta_i(t_1) \rangle \langle \beta_i(t_2) \rangle \right]$$
$$\times \langle \Delta\beta_j(t_3) \Delta\beta_j(t_4) \rangle.$$

From Eqs. (11)–(13) and (15), we then find

$$\langle a_1^{(1)}(T)^2 a_2^{(1)}(T)^2 \rangle + \langle a_1^{(1)}(T)^2 a_3^{(1)}(T)^2 \rangle$$

$$= \int_0^T dt_1 \int_0^T dt_2 \Omega(t_1)\Omega(t_2)\langle \Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_2)\rangle \int_0^T dt_3 \int_0^T dt_4 \sin\Theta(t_3)\sin\Theta(t_4)\left[\langle \Delta\beta_z(t_3)\Delta\beta_z(t_4)\rangle + \mu_z^2\right]$$

$$+ \int_0^T dt_1 \int_0^T dt_2 \Omega(t_1)\Omega(t_2)\langle \Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_2)\rangle \int_0^T dt_3 \int_0^T dt_4 \cos\Theta(t_3)\cos\Theta(t_4)\left[\langle \Delta\beta_z(t_3)\Delta\beta_z(t_4)\rangle + \mu_z^2\right]$$

$$= \left\{ \frac{1}{2\pi}\int_{-\infty}^\infty d\omega \, S_\Omega(\omega)F_\Omega(\omega,T)\right\}\left\{\frac{1}{2\pi}\int_{-\infty}^\infty d\omega\left[S_z(\omega) + 2\pi\mu_z^2\delta(\omega)\right]\left[\left|\int_0^T dt e^{i\omega t}\sin\Theta(t)\right|^2 + \left|\int_0^T dt e^{i\omega t}\cos\Theta(t)\right|^2\right]\right\}$$

$$= \left\{\frac{1}{2\pi}\int_{-\infty}^\infty d\omega \, S_\Omega(\omega)F_\Omega(\omega,T)\right\}\left\{\frac{1}{2\pi}\int_{-\infty}^\infty d\omega\left[S_z(\omega) + 2\pi\mu_z^2\delta(\omega)\right]F_Z(\omega,T)\right\} = I_\Omega(T)I_Z(T), \tag{A15}$$

as given in the main text.

When $i = j = 1$, on the other hand, we have

$$\langle a_1^{(1)}(T)^4 \rangle = \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4 \, \Omega(t_1)\Omega(t_2)\Omega(t_3)\Omega(t_4)\left[\langle \beta_\Omega(t_1)\beta_\Omega(t_2)\rangle\langle\beta_\Omega(t_3)\beta_\Omega(t_4)\rangle\right.$$

$$+ \langle \beta_\Omega(t_1)\beta_\Omega(t_3)\rangle\langle\beta_\Omega(t_2)\beta_\Omega(t_4)\rangle + \langle\beta_\Omega(t_1)\beta_\Omega(t_4)\rangle\langle\beta_\Omega(t_3)\beta_\Omega(t_2)\rangle - 2\mu_\Omega^4\Big]$$

$$= \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4 \, \Omega(t_1)\Omega(t_2)\Omega(t_3)\Omega(t_4)\Big\{\langle \Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_2)\rangle\langle\Delta\beta_\Omega(t_3)\Delta\beta_\Omega(t_4)\rangle$$

$$+ \langle \Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_3)\rangle\langle\Delta\beta_\Omega(t_2)\Delta\beta_\Omega(t_4)\rangle + \langle\Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_4)\rangle\langle\Delta\beta_\Omega(t_2)\Delta\beta_\Omega(t_3)\rangle + [\langle\beta_\Omega(t_1)\beta_\Omega(t_2)\rangle$$

$$+ \langle\beta_\Omega(t_1)\beta_\Omega(t_3)\rangle + \langle\beta_\Omega(t_1)\beta_\Omega(4)\rangle + \langle\beta_\Omega(t_2)\beta_\Omega(t_3)\rangle + \langle\beta_\Omega(t_2)\beta_\Omega(t_4)\rangle + \langle\beta_\Omega(t_3)\beta_\Omega(t_4)\rangle]\mu_\Omega^2 - 5\mu_\Omega^4\Big\}. \tag{A16}$$

If the amplitude noise is zero mean, the last line in this expression vanishes, producing

$$\langle a_1^{(1)}(T)^4\rangle = \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \int_0^T dt_4 \, \Omega(t_1)\Omega(t_2)\Omega(t_3)\Omega(t_4)\left[\langle\Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_2)\rangle\langle\Delta\beta_\Omega(t_3)\Delta\beta_\Omega(t_4)\rangle\right.$$

$$+ \langle\Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_3)\rangle\langle\Delta\beta_\Omega(t_2)\Delta\beta_\Omega(t_4)\rangle + \langle\Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_4)\rangle\langle\Delta\beta_\Omega(t_3)\Delta\beta_\Omega(t_2)\rangle]$$

$$= 3\int_0^T dt_1 \int_0^T dt_2 \Omega(t_1)\Omega(t_2)\langle\Delta\beta_\Omega(t_1)\Delta\beta_\Omega(t_2)\rangle \int_0^T dt_3 \int_0^T dt_4 \Omega(t_3)\Omega(t_4)\langle\Delta\beta_\Omega(t_3)\Delta\beta_\Omega(t_4)\rangle$$

$$= 3\left[\frac{1}{2\pi}\int_{-\infty}^\infty d\omega S_\Omega(\omega)F_\Omega(\omega,T)\right]^2 = 3\,I_\Omega(T)^2,$$

as given in the main text.

Next, we consider the fourth term on the right-hand side of Eq. (A8), which depends on $a_1^{(2)}(T)$. Using Eq. (A11) to evaluate the higher order moment and taking $\langle \beta_z(t)\rangle = \mu_z$ for stationary noise, we obtain

$$\langle a_1^{(2)}(T)^2\rangle = \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)] \langle\beta_z(t_1)\beta_z(t_2)\beta_z(t_3)\beta_z(t_4)\rangle$$

$$= \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)]\left[\langle\beta_z(t_1)\beta_z(t_2)\rangle\langle\beta_z(t_3)\beta_z(t_4)\rangle\right.$$

$$+ \langle\beta_z(t_1)\beta_z(t_3)\rangle\langle\beta_z(t_2)\beta_z(t_4)\rangle + \langle\beta_z(t_1)\beta_z(t_4)\rangle\langle\beta_z(t_2)\beta_z(t_3)\rangle - 2\mu_z^4\Big]$$

$$= \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)]\Big\{\langle\Delta\beta_z(t_1)\Delta\beta_z(t_2)\rangle\langle\Delta\beta_z(t_3)\Delta\beta_z(t_4)\rangle$$

$$+ \langle\Delta\beta_z(t_1)\Delta\beta_z(t_3)\rangle\langle\Delta\beta_z(t_2)\Delta\beta_z(t_4)\rangle + \langle\Delta\beta_z(t_1)\Delta\beta_z(t_4)\rangle\langle\Delta\beta_z(t_2)\Delta\beta_z(t_3)\rangle + [\langle\beta_z(t_1)\beta_z(t_2)\rangle + \langle\beta_z(t_1)\beta_z(t_3)\rangle$$

$$+ \langle\beta_z(t_1)\beta_z(4)\rangle + \langle\beta_z(t_2)\beta_z(t_3)\rangle + \langle\beta_z(t_2)\beta_z(t_4)\rangle + \langle\beta_z(t_3)\beta_z(t_4)\rangle]\mu_z^2 - 5\mu_z^4\Big\}. \tag{A17}$$

If the dephasing noise is zero-mean, the final line vanishes and we have

$$\left\langle a_1^{(2)}(T)^2 \right\rangle_{\mu_z=0} = \frac{1}{(2\pi)^2} \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)]$$
$$\times \left[\int_{-\infty}^{\infty} d\omega \, e^{i\omega(t_1-t_2)} S(\omega) \int_{-\infty}^{\infty} d\omega' \, e^{i\omega'(t_3-t_4)} S(\omega') + \int_{-\infty}^{\infty} d\omega \, e^{i\omega(t_1-t_3)} S(\omega) \int_{-\infty}^{\infty} d\omega' \, e^{i\omega'(t_2-t_4)} S(\omega') \right.$$
$$\left. + \int_{-\infty}^{\infty} d\omega \, e^{i\omega(t_1-t_4)} S(\omega) \int_{-\infty}^{\infty} d\omega' \, e^{i\omega'(t_2-t_3)} S(\omega') \right].$$

Collecting like terms produces

$$\left\langle a_1^{(2)}(T)^2 \right\rangle_{\mu_z=0} = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \, G_Z(\omega, \omega', T) \, S_z(\omega) S_z(\omega'),$$

where

$$G_Z(\omega, \omega', T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)][e^{i\omega(t_1-t_2)} e^{i\omega'(t_3-t_4)}$$
$$+ e^{i\omega(t_1-t_3)} e^{i\omega'(t_2-t_4)} + e^{i\omega(t_1-t_4)} e^{i\omega'(t_2-t_3)}].$$

When the dephasing noise is nonzero-mean, the terms in the final line of Eq. (A17) make a contribution,

$$\left\langle a_1^{(2)}(T)^2 \right\rangle_{\mu_z \neq 0} = \mu_z^2 \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)] \left[\langle \beta_z(t_1)\beta_z(t_2) \rangle \right.$$
$$\left. + \langle \beta_z(t_3)\beta_z(t_4) \rangle + \langle \beta_z(t_1)\beta_z(t_3) \rangle + \langle \beta_z(t_2)\beta_z(t_4) \rangle + \langle \beta_z(t_1)\beta_z(t_4) \rangle + \langle \beta_z(t_2)\beta_z(t_3) \rangle - 5\mu_z^2 \right]$$
$$= \mu_z^2 \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)] \left[\langle \Delta\beta_z(t_1)\Delta\beta_z(t_2) \rangle \right.$$
$$\left. + \langle \Delta\beta_z(t_3)\Delta\beta_z(t_4) \rangle + \langle \Delta\beta_z(t_1)\Delta\beta_z(t_3) \rangle + \langle \Delta\beta_z(t_2)\Delta\beta_z(t_4) \rangle + \langle \Delta\beta_z(t_1)\Delta\beta_z(t_4) \rangle + \langle \Delta\beta_z(t_2)\Delta\beta_z(t_3) \rangle + \mu_z^2 \right]$$
$$= \frac{\mu_z^2}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin[\Theta(t_1) - \Theta(t_2)] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)][e^{i\omega(t_1-t_2)} e^{i\omega'(t_3-t_4)}$$
$$+ e^{i\omega(t_1-t_3)} e^{i\omega'(t_2-t_4)} + e^{i\omega(t_1-t_4)} e^{i\omega'(t_2-t_3)}] \left[S_z(\omega)\delta(\omega') + S_z(\omega')\delta(\omega) + \frac{2\pi\mu_z^2}{3}\delta(\omega)\delta(\omega') \right]$$
$$= \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \, G_Z(\omega, \omega', T) \left[\frac{\mu_z^2}{2\pi} S_z(\omega)\delta(\omega') + \frac{\mu_z^2}{2\pi} S_z(\omega')\delta(\omega) + \frac{\mu_z^4}{3}\delta(\omega)\delta(\omega') \right]. \tag{A18}$$

By taking $\left\langle a_1^{(2)}(T)^2 \right\rangle_{\mu_z=0} + \left\langle a_1^{(2)}(T)^2 \right\rangle_{\mu_z \neq 0}$, we recover Eq. (29) in the main text.

APPENDIX B: SLEPIAN (DPSS) SEQUENCES AND SPECTRAL CONCENTRATION

For a detailed treatment of the Slepian or DPSSs and their applications in classical signal processing and spectral estimation, see Refs. [25,26,38]. We briefly summarize material from these references here. The DPSSs $\{v_n^{(k)}\} \equiv \{v_0^{(k)}, \ldots, v_N^{(k)}\}$ are a family of N, length-N sequences that are solutions to a $N \times N$ Toeplitz matrix eigenvalue equation:

$$\sum_{m=0}^{N-1} \frac{\sin 2\pi W(n-m)}{\pi(n-m)} v_m^{(k)} = \lambda_k(N, W) \, v_n^{(k)}. \tag{B1}$$

Here, $k \in \{0, \ldots, N-1\}$ indexes the order of the DPSS and $W < 1/2$ is a bandwidth parameter related to spectral concentration in the frequency domain. The frequency domain representations of the DPSS, known as the discrete prolate spheroidal wave functions, depend on the DTFT of the DPSS:

$$U^{(k)}(N, W; \omega) = \epsilon_k \sum_{n=0}^{N-1} v_n^{(k)} e^{i\omega[n-(N-1)/2]\Delta t}.$$

Here, $\epsilon_k = 1 \, (i)$ for k even (odd).

For a general length-N sequence, $\{a_1, \ldots, a_N\} \equiv \{a_n\}$, with DTFT $\tilde{a}(\omega) = \sum_{n=0}^{N-1} a_n e^{i\omega[n-(N-1)/2]\Delta t}$, spectral concentration in a frequency band $B(0) \equiv [-\delta\omega, \delta\omega]$ is quantified by the ratio

$$\mathcal{C}_B[\{a_n\}] = \frac{\int_{-\delta\omega}^{\delta\omega} d\omega \, |\tilde{a}(\omega)|^2}{\int_{-\omega_N}^{\omega_N} d\omega \, |\tilde{a}(\omega)|^2},$$

where $\omega_N \equiv \pi/\Delta t$ is the Nyquist frequency and $\delta\omega < \omega_N$. For $W = \delta\omega\Delta t/2\pi$, the spectral concentration of the kth order DPSS in the band $B(0)$ is given by

$$\mathcal{C}_B[\{v_n^{(k)}\}] = \frac{\int_{-\delta\omega}^{\delta\omega} d\omega \, |U^{(k)}(N, W; \omega)|^2}{\int_{-\omega_N}^{\omega_N} d\omega \, |U^{(k)}(N, W; \omega)|^2} = \lambda_k(N, W),$$

where $\lambda_k(N, W)$ is the eigenvalue from Eq. (B1). Of all length-N sequences, the DPSS of order $k = 0$ with eigenvalue $\lambda_0(N, W) \approx 1$ has the maximum spectral concentration in $B(0)$. The first $\lfloor NW \rfloor$ DPSS orders are also highly spectrally concentrated.

For the waveforms we use in amplitude noise QNS, we can define a similar measure for the spectral concentration of the amplitude filter in a band $B(\omega_0) \equiv [\omega_0 - \delta\omega, \omega_0 + \delta\omega]$,

$$\mathcal{R} \equiv \frac{\int_{B(\omega_0)} d\omega \, F_\Omega(\omega, T)}{\int_{-\infty}^{\infty} d\omega \, F_\Omega(\omega, T)}.$$

Because the amplitude FF is defined as the *continuous* Fourier transform of the amplitude waveform, the bound on the lower integral is $(-\infty, \infty)$ rather than $(-\omega_N, \omega_N)$.

APPENDIX C: DETAILS OF NUMERICAL OPTIMIZATION

The dephasing-robust waveforms were originally obtained numerically through the use of the constrained nonlinear optimization solver Casadi and backend IPOPT [34–36]. Here we describe how we expressed the problem of filter design with bounded controls as a scalable numerical optimization problem. The procedure may be of independent interest for linear programming optimization problems with a large number of affine constraints. The true feasible region is defined by the following set of $2N$ linear constraints:

$$\pm \sum_{k=0}^{K} v_m^{(k)}(N, W) \left[\Omega_c^{(k)} \cos[\omega_0 m \Delta t] + \Omega_s^{(k)} \sin[\omega_0 m \Delta t] \right]$$
$$\leqslant \Omega_{\max}, \quad \forall m \in \{0, \dots, N-1\}. \tag{C1}$$

For notational convenience, we first define the $2K$ variables $x^{(k)}$ and $4KN$ coefficients $a_m^{(k)}$ by

$$x^{(k)} = \Omega_c^{(k)},$$
$$x^{(k+K)} = \Omega_s^{(k)},$$
$$a_m^{(k)} = v_m^{(k)}(N, W) \cos[\omega_0 m \Delta t]/\Omega_{\max},$$
$$a_m^{(k+K)} = v_m^{(k)}(N, W) \sin[\omega_0 m \Delta t]/\Omega_{\max},$$
$$a_{m+N}^{(k)} = -v_m^{(k)}(N, W) \cos[\omega_0 m \Delta t]/\Omega_{\max},$$
$$a_{m+N}^{(k+K)} = -v_m^{(k)}(N, W) \sin[\omega_0 m \Delta t]/\Omega_{\max}.$$

We now recast the problem into a standard form, which can be done by dividing each inequality by $\Omega_{\max}$ and combining the coefficients to produce

$$\sum_{k=0}^{2K} a_m^{(k)} x^{(k)} \leqslant 1 \quad \forall m \in \{0, \dots, 2N-1\}. \tag{C2}$$

We are inspired by a classic routine in linear programming for identifying redundant constraints, which have the defining property that they can be pruned from the system without changing the feasible region [42], namely, we attempt to maximize the violation of one selected constraint m', subject to all the other constraints. The selected constraint is redundant with respect to the other constraints if and only if the maximum

violation is negative:

$$\max_{x^{(k)}} \quad v = -1 + \sum_{k=0}^{2K} a_{m'}^{(k)} x^{(k)},$$
$$\text{s.t.} \quad \forall m \neq m' \quad \sum_{k=0}^{2K} a_m^{(k)} x^{(k)} \leqslant 1. \tag{C3}$$

If the result of this optimization is a positive number v, then there exists an assignment $x^{(k)}$ such that

$$\forall m \neq m' \quad \sum_{k=0}^{2K} a_m^{(k)} x^{(k)} \leqslant 1, \tag{C4}$$

but $\sum_{k=0}^{2K} a_{m'}^{(k)} x^{(k)} > 1$. Such an assignment lies outside the feasible region of the constraint set Eq. (C2), but within the feasible region of the constraint set Eq. (C4). This proves that the constraint m' has an effect on the feasible region if removed, and is therefore not redundant.

In contrast, if the result of the optimization is a negative number v, then there does not exist an assignment $x^{(k)}$ which simultaneously satisfies Eq. (C4) while not satisfying $\sum_{k=0}^{2K} a_{m'}^{(k)} x^{(k)} \leqslant 1$. By applying De Morgan's laws of formal logic, this is equivalent to the statement that Eq. (C4) logically implies $(\sum_{k=0}^{2K} a_{m'}^{(k)} x^{(k)} \leqslant 1)$. Therefore, incorporating the constraint m' does not remove any points from the feasible region. As additional constraints can never increase the size of the feasible region, this proves that the feasible regions of constraint set Eq. (C2), and constraint set Eq. (C4) are equal. Therefore, the constraint m' is redundant with respect to all the other constraints and can be pruned from the system without affecting the feasible region. This can be desirable due to the gain in efficiency of having a problem with fewer constraints. For the edge cases where the result of the optimization is zero [positive infinity], then the constraint m' is redundant (not redundant), respectively.

We modify this routine to also prune *approximately* redundant constraints which have a small effect on the feasible region. The algorithm first creates an overapproximation to the feasible region whose distance under the L_∞ norm (i.e., the maximum absolute distance) from the true feasible region is bounded by ϵ, and then contracts the overapproximation by a factor of ϵ, thus producing an underapproximation to the true feasible region.

For our target application, all our constraints in our applications have a natural length-scale because they are affine constraints with a nonzero constant term. This allows us to define a small violation in terms of the constant term, namely, a violation is small if the maximum violation of Eq. (C3) is less than a chosen $\epsilon > 0$. This approach would be ill-defined if we included constraints with no constant term such as

$$\sum_{k=0}^{2K} a_m^{(k)} x^{(k)} \leqslant 0.$$

Any violation to this constraint can be made arbitrarily large or small by simply multiplying $a_m^{(k)} \, \forall \, k$ by a suitable constant > 0. This constant rescaling trick cannot be applied to constraints with a constant term, as the conversion to standard form would remove any such scaling. As such, the question of

Algorithm 1. Prune approximatively redundant affine constraints.

Require: $\epsilon > 0$, affine constraints set $\mathcal{C}$
 1: Normalize the affine constraints such that the constant term is 1
 2: Randomly shuffle order of constraints
 3: active set $\mathcal{A} \leftarrow \emptyset$
 4: **for each** $c \in \mathcal{C}$ **do**
 5:     violation $v \leftarrow$ maximize violation of c subject to $\mathcal{A}$ $\triangleright$ using linear programming subroutine
 6:     **if** $v > \epsilon$ **then**
 7:         $\mathcal{A} \leftarrow \mathcal{A} \cup c$ $\triangleright$ include non-redundant constraints
 8:     **end if**
 9: **end for**
10: Tighten all constraints in $\mathcal{A}$ by factor of $(1 + \epsilon)$
11: Randomly shuffle order of active set $\mathcal{A}$
12: **for each** $c \in \mathcal{A}$ **do**
13:     violation $v \leftarrow$ maximize violation of c subject to $\mathcal{A} \setminus c$ $\triangleright$ using linear programming subroutine
14:     **if** $v \leqslant 0$ **then**
15:         $\mathcal{A} \leftarrow \mathcal{A} \setminus c$ $\triangleright$ prune redundant constraints
16:     **end if**
17: **end for**
18: **return** $\mathcal{A}$

---

whether a violation of this constraint is large or small depends not on the projective space of equivalent representations of this constraint but rather on the particular way that the constraint is written. Therefore, the notion of a small violation would be ill-defined.

In addition to using this property to identify *approximately* redundant constraints, we also exploit this property to later contract the feasible region by rescaling the constant term. This is used to convert an overapproximation of the feasible region into an underapproximation in an unbiased manner, namely, after we have identified a core subset of constraints M, we then tighten the constraints by a factor of $(1 + \epsilon)$ as follows:

$$\sum_{k=0}^{2K} a_m^{(k)} x^{(k)} \leqslant (1 + \epsilon)^{-1} \quad \forall m \in \{0, \ldots, 2N - 1\}. \quad \text{(C5)}$$

The heuristic runtime of Algorithm 1 scales as $O(m\,f(n, |\mathcal{A}|))$, where m is the number of linear inequality constraints, $\mathcal{A}$ is the largest active set encountered by the algorithm, and $f(n, m)$ is the time to solve a linear programming problem with n variables and m inequality constraints. Significantly, $|\mathcal{A}|$ does not asymptotically depend on m for well-behaved instances (such as in our use case when the constraints are continuously parameterized), making the overall runtime of Algorithm 1 linear in m.

For our application in Sec. III B, the naïve formulation involved a nonlinear program in six variables with 40 000 linear inequality constraints. After the reduction with $\epsilon = 0.10$, we had six variables with 200 linear inequality constraints. This substantially improved the efficiency of solving the optimization problem Eq. (41), which was otherwise intractable to run at scale.

APPENDIX D: DERIVATION OF THE DEPHASING-ROBUST DEPHASING FF

Recall that the dephasing FF is defined by

$$F_Z(\omega, T) = \left| \int_0^T dt \cos \Theta(t) e^{-i\omega t} \right|^2 + \left| \int_0^T dt \sin \Theta(t) e^{-i\omega t} \right|^2, \quad \text{(D1)}$$

where $\Theta(t) = \int_0^t ds\, \Omega(s)$ is the angle of rotation about σ_1 generated by the control. For $\Omega(t) = \Omega_0 \sin(\lambda t)$, the angle rotation becomes $\Theta(t) = \Omega_0 (1 - \cos \lambda t)/\lambda$. Observe that $\Theta(t)$, like $\Omega(t)$, is periodic over $\tau_\lambda \equiv 2\pi/\lambda$. If the total evolution time is an integer multiple of τ_λ, i.e., $T = M\tau_\lambda$, we can use the periodic nature of $\Theta(t)$ to write the cosine term in Eq. (D1) as

$$\left| \int_0^T dt \cos \Theta(t) e^{-i\omega t} \right|^2 = \left| \sum_{m=0}^{M-1} \int_{m\tau_\lambda}^{(m+1)\tau_\lambda} dt \cos \Theta(t) e^{-i\omega t} \right|^2 = \left| \sum_{m=0}^{M-1} e^{-im\omega\tau_\lambda} \right|^2 \left| \int_0^{\tau_\lambda} dt \cos \Theta(t) e^{-i\omega t} \right|^2. \quad \text{(D2)}$$

Summing the geometric series in the first term produces [16]

$$\left| \sum_{m=0}^{M-1} e^{-im\omega\tau_\lambda} \right|^2 = \frac{\sin^2(M\pi\omega/\lambda)}{\sin^2(\pi\omega/\lambda)}. \quad \text{(D3)}$$

By substituting this expression into Eq. (D2) and performing a similar calculation on the sine term in Eq. (D1), we find

$$\left| \int_0^T dt \cos \Theta(t) e^{-i\omega t} \right|^2 = \frac{\sin^2(M\pi\omega/\lambda)}{\sin^2(\pi\omega/\lambda)} \left| \int_0^{\tau_\lambda} dt \cos \Theta(t) e^{-ik\lambda t} \right|^2, \tag{D4}$$

$$\left| \int_0^T dt \sin \Theta(t) e^{-i\omega t} \right|^2 = \frac{\sin^2(M\pi\omega/\lambda)}{\sin^2(\pi\omega/\lambda)} \left| \int_0^{\tau_\lambda} dt \sin \Theta(t) e^{-ik\lambda t} \right|^2. \tag{D5}$$

Subsituting these expressions into Eq. (D1) produces the exact form of the dephasing FF given in Eq. (46) of the main text.

To obtain Eq. (49), the approximate form of the dephasing FF in terms of Bessel functions, we approximate Eq. (D3) as a frequency comb in the limit of large M [16]:

$$\left| \sum_{m=0}^{M-1} e^{-im\omega\tau_\lambda} \right|^2 = \frac{\sin^2(M\pi\omega/\lambda)}{\sin^2(\pi\omega/\lambda)} \overset{M\gg 1}{\approx} M\lambda \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda). \tag{D6}$$

When $M \gg 1$, cosine and sine terms in Eqs. (D4) and (D5) then take the approximate forms

$$\left| \int_0^T dt \cos \Theta(t) e^{-i\omega t} \right|^2 \approx M\lambda \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left| \int_0^{\tau_\lambda} dt \cos \Theta(t) e^{-ik\lambda t} \right|^2, \tag{D7}$$

$$\left| \int_0^T dt \sin \Theta(t) e^{-i\omega t} \right|^2 \approx M\lambda \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left| \int_0^{\tau_\lambda} dt \sin \Theta(t) e^{-ik\lambda t} \right|^2. \tag{D8}$$

The integral terms in Eqs. (D7) and (D8) can be written in terms of exponential functions:

$$\int_0^{\tau_\lambda} dt \cos \Theta(t) e^{-ik\lambda t} = \frac{1}{2} \int_0^{\tau_\lambda} dt e^{i\Theta(t)} e^{-ik\lambda t} + \frac{1}{2} \int_0^{\tau_\lambda} dt e^{-i\Theta(t)} e^{-ik\lambda t}, \tag{D9}$$

$$\int_0^{\tau_\lambda} dt \sin \Theta(t) e^{-ik\lambda t} = \frac{1}{2i} \int_0^{\tau_\lambda} dt e^{i\Theta(t)} e^{-ik\lambda t} - \frac{1}{2i} \int_0^{\tau_\lambda} dt e^{-i\Theta(t)} e^{-ik\lambda t}. \tag{D10}$$

By substituting $\tau_\lambda = 2\pi/\lambda$ and $\Theta(t) = \Omega_0(1 - \cos\lambda t)/\lambda$ into the positive exponential term, we find

$$\int_0^{\tau_\lambda} dt e^{i\Theta(t)} e^{-ik\lambda t} = e^{i\Omega_0/\lambda} \int_0^{2\pi/\lambda} dt e^{-i(\frac{\Omega_0}{\lambda}\cos\lambda t + k\lambda t)} = \frac{e^{i\Omega_0/\lambda}}{\lambda} \int_0^{2\pi} d\tau e^{-i(\frac{\Omega_0}{\lambda}\cos\tau + k\tau)} = \frac{e^{i\Omega_0/\lambda} e^{ik\pi/2}}{\lambda} \int_{\pi/2}^{5\pi/2} d\tau e^{-i(\frac{\Omega_0}{\lambda}\sin\tau + k\tau)}$$

$$= \frac{e^{i\Omega_0/\lambda} e^{ik\pi/2}}{\lambda} \left[\int_0^{2\pi} d\tau e^{-i(\frac{\Omega_0}{\lambda}\sin\tau + k\tau)} + \int_{2\pi}^{5\pi/2} d\tau e^{-i(\frac{\Omega_0}{\lambda}\sin\tau + k\tau)} - \int_0^{\pi/2} d\tau e^{-i(\frac{\Omega_0}{\lambda}\sin\tau + k\tau)} \right].$$

The last two terms in this expression cancel, leaving

$$\int_0^{\tau_\lambda} dt e^{i\Theta(t)} e^{-ik\lambda t} = \frac{e^{i\Omega_0/\lambda} e^{ik\pi/2}}{\lambda} \int_0^{2\pi} d\tau\, e^{-i(\frac{\Omega_0}{\lambda}\sin\tau + k\tau)} = \frac{e^{i\Omega_0/\lambda} e^{-ik\pi/2}}{\lambda} \int_{-\pi}^{\pi} d\tau\, e^{i(\frac{\Omega_0}{\lambda}\sin\tau - k\tau)} = \frac{2\pi e^{i\Omega_0/\lambda} e^{-ik\pi/2}}{\lambda} J_k\left(\frac{\Omega_0}{\lambda}\right), \tag{D11}$$

where the last equality follows from the integral representation of a Bessel function of the first kind. Performing a similar calculation on the negative exponential term yields

$$\int_0^{\tau_\lambda} dt e^{-i\Theta(t)} e^{-ik\lambda t} = \frac{2\pi e^{-i\Omega_0/\lambda} e^{-ik\pi/2}}{\lambda} J_{-k}\left(\frac{\Omega_0}{\lambda}\right) = \frac{2\pi e^{-i\Omega_0/\lambda} e^{-ik\pi/2}}{\lambda} (-1)^k J_k\left(\frac{\Omega_0}{\lambda}\right). \tag{D12}$$

By combining Eqs. (D1) and (D7)–(D12), the dephasing FF becomes

$$F_Z(\omega, T) \approx M\lambda \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left[\left| \int_0^{\tau_\lambda} dt \cos \Theta(t) e^{-ik\lambda t} \right|^2 + \left| \int_0^{\tau_\lambda} dt \sin \Theta(t) e^{-ik\lambda t} \right|^2 \right]$$

$$= M\lambda \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left[\frac{1}{2} \left| \int_0^{\tau_\lambda} dt e^{i\Theta(t)} e^{-ik\lambda t} \right|^2 + \frac{1}{2} \left| \int_0^{\tau_\lambda} dt e^{-i\Theta(t)} e^{-ik\lambda t} \right|^2 \right]$$

$$= M\lambda \left(\frac{2\pi}{\lambda}\right)^2 \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left| J_k\left(\frac{\Omega_0}{\lambda}\right) \right|^2 = 2\pi T \sum_{k\in\mathbb{Z}} \delta(\omega - k\lambda) \left| J_k\left(\frac{\Omega_0}{\lambda}\right) \right|^2.$$

Note that the same approximate dephasing FF can be obtained for the waveform $\Omega(t) = \Omega_0 \cos(\lambda t)$.

APPENDIX E: HIGHER ORDER FF COMPUTATION

In Appendix A, an equation for the higher-order dephasing FF $G_Z(\omega, \omega', T)$ was derived:

$$G_Z(\omega, \omega', T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin\left[\Theta(t_1) - \Theta(t_2)\right] \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin[\Theta(t_3) - \Theta(t_4)]$$

$$\times \, [e^{i\omega(t_1 - t_2)} e^{i\omega'(t_3 - t_4)} + e^{i\omega(t_1 - t_3)} e^{i\omega'(t_2 - t_4)} + e^{i\omega(t_1 - t_4)} e^{i\omega'(t_2 - t_3)}]. \tag{E1}$$

This formula involves a quadruply nested integral, and it is not obvious how to factor the expression into smaller, more manageable integrals. Directly using this expression to evaluate the higher-order FF at single pair of frequencies ω, ω' has a runtime complexity of $O(N^4)$, where N is the number of segments in the piecewise-constant waveform. To sample the higher-order FF at a grid of frequencies (with density proportional to N) has complexity $O(N^6)$, which is extremely computationally intensive. Here we describe a more efficient approach for sampling $G_Z(\omega, \omega', T)$ at a grid of frequencies with runtime $O(N^2 \log_2(N)^2)$.

First, we use the sum and difference trigonometric identities to rewrite $\sin[\Theta(t_1) - \Theta(t_2)]$ into a (sum of) products:

$$\sin[\Theta(t_1) - \Theta(t_2)] = \sin\Theta(t_1)\cos\Theta(t_2) - \cos\Theta(t_1)\sin\Theta(t_2). \tag{E2}$$

Using this to rewrite Eq. (E1), we get

$$G_Z(\omega, \omega', T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \int_0^T dt_3 \int_0^{t_3} dt_4 [\sin\Theta(t_1)\cos\Theta(t_2) - \cos\Theta(t_1)\sin\Theta(t_2)]$$

$$\times \, [\sin\Theta(t_3)\cos\Theta(t_4) - \cos\Theta(t_3)\sin\Theta(t_4)][e^{i\omega(t_1 - t_2)} e^{i\omega'(t_3 - t_4)} + e^{i\omega(t_1 - t_3)} e^{i\omega'(t_2 - t_4)} + e^{i\omega(t_1 - t_4)} e^{i\omega'(t_2 - t_3)}].$$

This expression can be expanded into a sum of 12 integrands. For the remainder of this section, we shall focus on the following subexpression, which has the arguments in the exponentials split according to the groupings t_1, t_4 and t_2, t_3 (which we note is different than the groupings t_1, t_2 and t_3, t_4 induced by the bounds on integration). The other subexpressions proceed similarly:

$$f(\omega, \omega') = \int_0^T dt_1 \int_0^{t_1} dt_2 \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin\Theta(t_1)\cos\Theta(t_2)\sin\Theta(t_3)\cos\Theta(t_4) e^{i\omega(t_1 - t_4)} e^{i\omega'(t_2 - t_3)}. \tag{E3}$$

For convenience, we define a helper function $H(\omega_1, \omega_2)$, where we emphasize the notational change from (ω, ω') to (ω_1, ω_2):

$$H(\omega_1, \omega_2) = \int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin\Theta(t_1)\cos\Theta(t_2) e^{i\omega_1 t_1} e^{i\omega_2 t_2}. \tag{E4}$$

We can now begin to rewrite Eq. (E3):

$$f(\omega, \omega') = \int_0^T dt_1 \int_0^{t_1} dt_2 \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin\Theta(t_1)\cos\Theta(t_2)\sin\Theta(t_3)\cos\Theta(t_4) e^{i\omega(t_1 - t_4)} e^{i\omega'(t_2 - t_3)}$$

$$= \int_0^T dt_1 \int_0^{t_1} dt_2 \int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin\Theta(t_1)\cos\Theta(t_2)\sin\Theta(t_3)\cos\Theta(t_4) e^{i\omega t_1} e^{-i\omega t_4} e^{i\omega' t_2} e^{-i\omega' t_3}$$

$$= \left(\int_0^T dt_1 \int_0^{t_1} dt_2 \, \sin\Theta(t_1)\cos\Theta(t_2) e^{i\omega t_1} e^{i\omega' t_2}\right)\left(\int_0^T dt_3 \int_0^{t_3} dt_4 \, \sin\Theta(t_3)\cos\Theta(t_4) e^{-i\omega' t_3} e^{-i\omega t_4}\right)$$

$$= H(\omega, \omega') \cdot H(-\omega', -\omega).$$

As such, we see that the arguments in the exponentials depending on different time orderings is no great barrier to factoring the subexpression into doubly nested integrals. Indeed, all 12 subexpressions in our expanded formula for $G_Z(\omega, \omega', T)$ yield to such decompositions (some using different helper functions). Importantly, decomposing the subexpression in this manner reduces the complexity from evaluating a single quadruply nested integral to computing two doubly nested integrals, which are substantially faster to calculate [runtime $O(N^2)$ as opposed to $O(N^4)$]. Through this approach, we can compute a single element $G_Z(\omega, \omega', T)$ in runtime $O(N^2)$ by computing many doubly nested integrals.

Discrete Fourier transform to efficiently compute higher order FF

We have shown how to compute a single element $G_Z(\omega, \omega', T)$ in runtime $O(N^2)$ by computing many doubly nested integrals. Next, we demonstrate how to use the fast-Fourier transform to compute $G_Z(\omega, \omega', T)$ for many frequencies at once, with total runtime $O(N^2 \log_2(N)^2)$.

We focus on the computationally relevant discrete Fourier transform (DFT) application, where we discretize the function $\Theta(t)$ into a sequence of length N, and the frequencies ω, ω' at which we evaluate $G_Z(\omega, \omega', T)$ are integer multiples of $\Delta\omega = 2\pi/(TN)$. It suffices to demonstrate that $H(\omega_1, \omega_2)$ can be efficiently computed at all such frequencies simultaneously [in runtime $O(N^2)$]. Once this has been done, then computing $G_Z(\omega, \omega', T)$ is straightforward.

We begin by rewriting

$$H(\omega_1, \omega_2) = \int_0^T dt_1 \int_0^{t_1} dt_2 \sin\Theta(t_1)\cos\Theta(t_2)e^{i\omega_1 t_1}e^{i\omega_2 t_2} = \int_0^T dt_1 e^{i\omega_1 t_1} \sin\Theta(t_1) \int_0^{t_1} dt_2 \cos\Theta(t_2)e^{i\omega_2 t_2}.$$

Next, we define a few more helper functions:

$$h(t, \omega) \equiv \int_0^t ds \cos\Theta(s)e^{i\omega s} = \int_0^T ds\, g(t, s)e^{i\omega s}, \quad \text{where } g(t, s) \equiv \begin{cases} \cos\Theta(s) & s \leqslant t \\ 0 & s > t. \end{cases}$$

Then we discretize time and convert the integrals to finite sums:

$$H(\omega_1, \omega_2) \equiv (\Delta t)^2 \sum_{j_1=0}^{N-1} e^{i\omega_1 j_1 \Delta t} \sin\Theta(j_1\Delta t) \sum_{i_2=0}^{j_1} \cos\Theta(j_2\Delta t)e^{i\omega_2 j_2 \Delta t}, \tag{E5}$$

$$h(k, \omega) \equiv (\Delta t)\sum_{j=0}^{N-1} g(k, j)e^{i\omega j \Delta t}, \tag{E6}$$

$$\text{where } g(k, j) \equiv \begin{cases} \cos\Theta(j) & j \leqslant k \\ 0 & j > k. \end{cases} \tag{E7}$$

If k is regarded as a fixed parameter, then Eq. (E6) is readily seen to be the discrete fourier transform of $g(k, j)$ in Eq. (E7) with respect to j. Therefore for each fixed value of t, the DFT is able to compute $h(k, \omega)$ for all (discretized) values of $\omega = m \cdot \Delta\omega$ simultaneously in time $O(N\log_2(N))$. By repeating this for each (discretized) time t, we are able to compute all values of $h(k, \omega)$ in runtime $O(N^2\log_2(N)^2)$.

After we have precomputed all values of $h(t, \omega)$, we express the original helper function $H(\omega_1, \omega_2)$ as

$$H(\omega_1, \omega_2) \equiv (\Delta t)^2 \sum_{j_1=0}^{N-1} e^{i\omega_1 j_1 \Delta t} \sin\Theta(j_1\Delta t) \sum_{i_2=0}^{j_1} \cos\Theta(j_2\Delta t)e^{i\omega_2 j_2 \Delta t},$$

$$= (\Delta t)\sum_{j_1=0}^{N-1} e^{i\omega_1 j_1 \Delta t}[\sin\Theta(j_1\Delta t)\,h(j_1, \omega_2)], \tag{E8}$$

and we see that it can again be regarded as a Fourier transform. Explicitly, we now regard ω_2 as a fixed parameter and take the DFT with respect to t_1. This can be computed for all frequencies ω_1 in runtime $O(N\log_2(N))$ using the DFT. By repeating this for all values of ω_2 [without needing to recompute $h(k, \omega)$ from the previous step], we can thus compute $H(\omega_1, \omega_2)$ for all (discretized) frequencies ω_1, ω_2 in runtime $O(N^2\log_2(N)^2)$. We note that this is almost the same asymptotic complexity as it takes to compute a single element of H. Thus, by carefully rearranging our original expression we can leverage the DFT to efficiently compute the higher-order FFs.

---

[1] J. P. G. van Dijk, E. Kawakami, R. N. Schouten, M. Veldhorst, L. M. K. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano, Phys. Rev. Appl. **12**, 044054 (2019).

[2] A. Soare, H. Ball, D. Hayes, J. Sastrawan, M. C. Jarratt, J. J. McLoughlin, X. Zhen, T. J. Green, and M. J. Biercuk, Nat. Phys. **10**, 825 (2014).

[3] H. Ball and M. J. Biercuk, EPJ Quantum Technol. **2**, 11 (2015).

[4] H. Ball, W. D. Oliver, and M. J. Biercuk, npj Quantum Inf. **2**, 16033 (2016).

[5] J. Raftery, A. Vrajitoarea, Z. Gengyan, Z. Leng, S. Srinivasan, and A. Houck, arXiv:1703.00942.

[6] M. H. Levitt, Prog. Nucl. Magn. Reson. Spectrosc. **18**, 61 (1986).

[7] L. Viola and S. Lloyd, Phys. Rev. A **58**, 2733 (1998).

[8] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).

[9] K. Khodjasteh and L. Viola, Phys. Rev. Lett. **102**, 080501 (2009).

[10] T. Chalermpusitarak, B. Tonekaboni, Y. Wang, L. M. Norris, L. Viola, and G. A. Paz-Silva, PRX Quantum **2**, 030315 (2021).

[11] Y. Oda, D. Lucarelli, K. Schultz, B. D. Clader, and G. Quiroz, arXiv:2206.03504 (2022).

[12] C. J. Trout, K. Schultz, P. Titum, L. Norris, G. Quiroz, and B. D. Clader, arXiv:2205.06741.

[13] P. Szańkowski, G. Ramon, J. Krzywda, D. Kwiatkowski, and Ł. Cywiński, J. Phys.: Condens. Matter **29**, 333001 (2017).

[14] T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk, New J. Phys. **15**, 095004 (2013).

[15] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, Y. Cory, David G.and Nakamura, J.-S. Tsai, and W. D. Oliver, Nat. Phys. **7**, 565 (2011).

[16] G. A. Álvarez and D. Suter, Phys. Rev. Lett. **107**, 230501 (2011).

[17] G. A. Paz-Silva, L. M. Norris, F. Beaudoin, and L. Viola, Phys. Rev. A **100**, 042334 (2019).

[18] L. M. Norris, G. A. Paz-Silva, and L. Viola, Phys. Rev. Lett. **116**, 150503 (2016).

[19] G. Ramon, Phys. Rev. B **100**, 161302(R) (2019).

[20] P. Szańkowski, M. Trippenbach, and L. Cywiński, Phys. Rev. A **94**, 012109 (2016).

[21] G. A. Paz-Silva, L. M. Norris, and L. Viola, Phys. Rev. A **95**, 022121 (2017).

[22] U. von Lüpke, F. Beaudoin, L. M. Norris, Y. Sung, R. Winik, J. Y. Qiu, M. Kjaergaard, D. Kim, J. Yoder, S. Gustavsson, L. Viola, and W. D. Oliver, PRX Quantum **1**, 010305 (2020).

[23] V. Frey, S. Mavadia, L. M. Norris, W. de Ferranti, D. Lucarelli, L. Viola, and M. J. Biercuk, Nat. Commun. **8**, 2189 (2017).

[24] V. Frey, L. M. Norris, L. Viola, and M. J. Biercuk, Phys. Rev. Appl. **14**, 024021 (2020).

[25] D. Slepian, Bell Syst. Tech. J. **57**, 1371 (1978).

[26] D. J. Thomson, Proc. IEEE **70**, 1055 (1982).

[27] L. M. Norris, D. Lucarelli, V. M. Frey, S. Mavadia, M. J. Biercuk, and L. Viola, Phys. Rev. A **98**, 032315 (2018).

[28] S. Blanes, F. Casas, J. Oteo, and J. Ros, Phys. Rep. **470**, 151 (2009).

[29] A. G. Kofman and G. Kurizki, Phys. Rev. Lett. **93**, 130406 (2004).

[30] l. Cywiński, R. M. Lutchyn, C. P. Nave, and S. Das Sarma, Phys. Rev. B **77**, 174509 (2008).

[31] C. Ferrie, C. Granade, G. Paz-Silva, and H. M. Wiseman, New J. Phys. **20**, 123005 (2018).

[32] A. Murphy, J. Epstein, G. Quiroz, K. Schultz, L. Tewala, K. McElroy, C. Trout, B. Tien-Street, J. A. Hoffmann, B. D. Clader, J. Long, D. P. Pappas, and T. M. Sweeney, Phys. Rev. Res. **4**, 013081 (2022).

[33] D. Lucarelli, Phys. Rev. A **97**, 062346 (2018).

[34] A. Wächter and L. T. Biegler, Math. Program. **106**, 25 (2006).

[35] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, Math. Program. Comput. **11**, 1 (2019).

[36] O. Tange, *GNU Parallel 2018* (Ole Tange, Berkeley, CA, 2018).

[37] I. Hansen, A. E. Seedhouse, A. Saraiva, A. Laucht, A. S. Dzurak, and C. H. Yang, Phys. Rev. A **104**, 062415 (2021).

[38] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, UK, 1993).

[39] L. Viola and E. Knill, Phys. Rev. Lett. **90**, 037901 (2003).

[40] T. Ishikawa, A. Yoshizawa, Y. Mawatari, H. Watanabe, and S. Kashiwaya, Phys. Rev. Appl. **10**, 054059 (2018).

[41] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, Rev. Mod. Phys. **86**, 361 (2014).

[42] Y. Estinmgsih, Farikhin, and R. H. Tjahjana, J. Phys.: Conf. Ser. **1321**, 022073 (2019).