Reducing sequencing complexity in dynamical quantum error suppression by Walsh modulation

David Hayes
Joint Quantum Institute, University of Maryland, College Park, MD 20742 USA

Kaveh Khodjasteh and Lorenza Viola
Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755 USA

Michael J. Biercuk[*]
School of Physics and Centre for Engineered Quantum Systems, The University of Sydney, NSW 2006 Australia
(Dated: September 28, 2011)

We study dynamical error suppression from the perspective of reducing sequencing complexity, in order to facilitate efficient semi-autonomous quantum-coherent systems. With this aim, we focus on *digital* sequences where all interpulse time periods are integer multiples of a minimum clock period and compatibility with simple digital classical control circuitry is intrinsic, using so-called *Walsh functions* as a general mathematical framework. The Walsh functions are an orthonormal set of basis functions which may be associated directly with the control propagator for a digital modulation scheme, and dynamical decoupling (DD) sequences can be derived from the locations of digital transitions therein. We characterize the suite of the resulting *Walsh dynamical decoupling* (WDD) sequences, and identify the number of periodic square-wave (Rademacher) functions required to generate a Walsh function as the key determinant of the error-suppressing features of the relevant WDD sequence. WDD forms a unifying theoretical framework as it includes a large variety of well-known and novel DD sequences, providing significant flexibility and performance benefits relative to basic quasi-periodic design. We also show how Walsh modulation may be employed for the protection of certain nontrivial logic gates, providing an implementation of a dynamically corrected gate. Based on these insights we identify Walsh modulation as a digital-efficient approach for physical-layer error suppression.

PACS numbers: 03.67.Pp, 03.65.Yz, 37.10.Ty

I. INTRODUCTION

Dynamical quantum error correction has been proposed as a strategy by which arbitrarily accurate evolutions may in principle be implemented in a large class of open quantum systems. This approach involves application of open-loop control protocols at the physical level [1–16]; through time-dependent modulation of the system's dynamics, the effects of an environment which fluctuates sufficiently slowly are coherently averaged out. Dynamical decoupling (DD) is an experimentally validated [17–33] subclass of these protocols specifically tailored to the task of suppressing decoherence during the implementation of the identity operator – resulting in improved quantum storage.

DD takes physical inspiration from the spin echo in nuclear magnetic resonance (NMR) [34–36], and relies on deterministic (periodic [1] or aperiodic [8, 10]), or even random [37, 38] modulation of the idle system via pulsed control. A proliferation of analytical formalisms and new DD schemes has appeared in the literature, in particular for the paradigmatic case of a single qubit exposed to pure (classical and/or quantum) dephasing [1, 10, 20, 23, 39–44]. While an understanding of the performance of these sequences may be unified by the application of a *noise filtering* framework [7, 16, 45–47], each sequence brings particular requirements for the necessary pulse timings, often incorporating nonintuitive analytical expressions or numerical search to define pulse locations in a sequence. As a result, the generation of DD pulse sequences at the lab bench is usually accomplished using specially programmed microcontrollers or a PC under user control.

In this work, we address the problem of control complexity in dynamical error correction, introducing a set of *digital* DD protocols optimized for hardware compatibility and minimization of sequencing complexity. These protocols are based on the *Walsh functions* [48, 49], which take binary values and are composed of products of square waves, forming an orthonormal basis similar to the sines and cosines. The Walsh functions benefit from compact notation and a uniform mathematical basis for sequence construction. We describe the error-suppressing properties of Walsh modulation and Walsh dynamical decoupling (WDD), and introduce a quantitative metric for *sequencing complexity*, r, the number of Rademacher square-wave functions which must be multiplied (or added mod-2) in hardware to generate the control propagator for that sequence. The Rademacher functions in turn can be trivially generated via ele-

arXiv:1109.6002v1 [quant-ph] 27 Sep 2011

mentary digital circuits synchronous with a distributed clock signal. Thus, by construction Walsh modulation is highly compatible with simple digital sequencing circuitry and digital clocking as may be needed in large or semi-autonomous systems. This approach therefore provides *efficient* error suppression, important for real-world implementations beyond simple demonstration experiments.

The remainder of this manuscript is organized as follows. After introducing the relevant system and control setting in Section II, Section III is devoted to describing the mathematical formulation and key features of the Walsh functions, and their natural use in defining WDD schemes. In this section, our main results are established, including an exact relationship between the sequencing complexity r and the order of decoherence suppression in the perturbative limit. Here, we also describe the entire taxonomy of WDD sequences, discussing relationships between WDD and well-known sequences, as well as characterizing new DD sequences with digital timing. In Sec. IV, we discuss extensions of the WDD formalism beyond the simplest setting of a single qubit exposed to classical dephasing noise. In particular, we show how two-axis generalizations of WDD naturally recover and expand existing concatenated DD schemes, and discuss how Walsh modulation allows enhancement of the fidelity of nontrivial gate operations, examining a recent trapped-ion experiment as an example [50]. We follow this with an analysis of the explicit benefits of WDD over other optimized DD approaches in Sec. V, and close with a brief summary and outlook. A proof of WDD error suppression properties from Rademacher functions is included in Appendix A for completeness.

II. PHYSICAL SETTING

DD is applicable to a variety of non-Markovian error models (including arbitrary many-qubit systems interacting with quantum environments, [2, 51]) and can address non-ideal pulses or continuous control scenarios (as in Eulerian DD [5]). Our starting point here, however, is the simplest yet practically important case of a single qubit subject to classical phase noise and controlled with ideal π pulses. In this setting, the qubit is affected by an undesired noise Hamiltonian,

$$H_{\mathrm{noise}} = \beta(t)\sigma_z, \qquad (1)$$

where $\beta(t)$ is a stochastic process and σ_i denote the Pauli matrices for $i = x, y, z$. We assume that the system can be controlled by application of idealized sequences of qubit rotations, each corresponding to a π rotation around the x axis, i.e., to the unitary operator σ_x. Each sequence is characterized by the pulse timings $\{t_j\}_{j=1}^s$, where $t_0 = 0$ and $t_{s+1} = \tau$ denote the initial (preparation) and final (readout) times respectively. We use $\delta_j = t_j/\tau$ to denote the "normalized pulse locations" and

the "pulse pattern" $p = \{\delta_j\}_{j=1}^s$, to distinguish among various sequences with the same running time [52].

During the evolution, the π pulses implement the control propagator

$$U_c(t) = \sigma_x^{[y(t)+1]/2},$$

where $y(t)$ takes values ± 1 and switches instantaneously between these values at times corresponding to application of the π pulses. For brevity, we shall refer to $y(t)$ as the "sequence propagator" in what follows. The "filter function" $F_p(\omega\tau)$ associated with the pulse pattern $p = \{\delta_j\}$ and sequence duration τ is defined in terms of $\tilde{y}(\omega\tau)$, the Fourier transform of the sequence propagator [40, 46]:

$$F_p(\omega\tau) = \omega^2 |\tilde{y}(\omega\tau)|^2 \qquad (2)$$

$$= \Big| \sum_{j=0}^s (-1)^j (e^{i\delta_j \omega\tau} - e^{i\delta_{j+1}\omega\tau}) \Big|^2. \qquad (3)$$

The case of free evolution (also referred to as free induction decay, FID, in NMR terminology) is formally included by letting $s = 0$.

The presence of the noise Hamiltonian (1) implies a coherence loss at the readout due to phase randomization, resulting from the ensemble average with respect to noise realizations. The action of DD is to break the system's evolution into a sequence of interactions with the environment with *alternating signs*, resulting in significantly reduced phase accumulation at the end of the sequence [1, 2]. This may be effectively represented as the expected value of the convolution of the stochastic noise term with the sequence propagator [10, 16, 35, 40, 46]. In particular, the filter function provides a compact *exact* expression for the coherence decay under Gaussian noise; if the system is prepared in a superposition of eigenstates of σ_z, its coherence $W \equiv |\langle\sigma_+\rangle(\tau)|$ decays as $e^{-\chi_p(\tau)}$, where

$$\chi_p(\tau) = \frac{2}{\pi} \int_0^\infty \frac{S_\beta(\omega)}{\omega^2} F_p(\omega\tau) d\omega, \qquad (4)$$

and $S_\beta(\omega)$ is the power spectrum of the noise $\beta(t)$. While nominally the integration range in Eq. (4) is infinite, in practice $S_\beta(\omega)$, and thus the "spectral measure" $\lambda(\omega) \equiv S_\beta(\omega)/2\pi\omega^2$ [44], is significant only for frequencies smaller than an ultraviolet cut-off frequency ω_c. The filter function can also be employed for a qubit coupled to a purely-dephasing quantum bosonic environment with a similar expression for decoherence under DD sequences [1, 10]. The effect of continuous control can also be approximated within the filter function formalism as long as terms that are of second order and higher in H_{noise} are ignored [7, 11] (see also Sec. IV).

The filter function enters the integrand in Eq. (4) as a multiplicative factor of $\lambda(\omega)$, therefore, *as long as $F_p(\omega\tau)$ is small for $\omega < \omega_c$*, the coherence loss will also remain small, as desired. Determining the actual value of $\chi_p(\tau)$

requires a detailed knowledge of $S_\beta(\omega)$, yet we may compare the coherence associated with various DD sequences (including free evolution) by directly comparing their corresponding filter functions over the interval $[0, \omega_c\tau]$. For all sequences the filter function vanishes at zero, however we may differentiate the low-frequency behaviors (rates of growth) among various timing patterns. Let the low-frequency behavior of the filter function be given by

$$F_p(\omega\tau) \propto (\omega\tau)^{2(\alpha+1)}, \qquad (5)$$

corresponding to

$$\tilde{y}(\omega\tau) \propto (\omega\tau)^\alpha.$$

In the language of filter design [16], Eq. (5), defines a "highpass filter" with a "rolloff" of $6(\alpha + 1)$dB/octave. As long as the cutoff frequency ω_c is sufficiently small, the low-frequency behavior of the filter function translates to

$$1 - W \propto (\omega_c\tau)^{2\alpha+1}.$$

For example, the filter function for free evolution $F_{\{0,1\}}(\omega\tau) = 2\sin^2(\omega\tau/2)$ corresponds to $\alpha = 0$, while high-order DD sequences have larger positive α's. The frequency range over which the filter function suppresses noise is given by the "bandwidth" Ω_p, roughly defined as the largest frequency below which $F_p(\omega\tau) \leq 1$, a value approximately commensurate with the value ω_{F1} introduced in previous analyses of the filter function [16]. Large values of the bandwidth Ω_p improve the high-frequency robustness of a particular DD sequence and allow it to be used with comparatively lower pulse rates.

III. DYNAMICAL DECOUPLING BY DIGITAL MODULATION

Many previously developed schemes for DD rely on access to sequences defined in continuous time, which is a good approximation for most benchtop experiments conducted today. In the long term, however, where large systems with error rates deep below the fault-tolerance threshold are required, these approximations will cease to provide an accurate estimate of residual error rates. Previous work has shown that the benefits of optimized sequences such as UDD become diminished in such cases where we expect the use of digital control and discretized time [16]. We are thus motivated to find *digital modulation* schemes that are both intrinsically compatible with discrete time and with the digital control hardware that will inevitably be employed in sequencing.

We define two relevant timescales for such a modulation scheme; the total running time τ, and a *minimum interval* (minimum switching time [44]) given by

$$\tau_{\min} = \tau/2^m,$$

where 2^m is the largest possible number of free-evolution periods in an applied sequence. In practice, $\tau_{\min}$ is bounded from below by technological constraints such as modulation rates or hardware clock speeds. In the case of digital modulation, all interpulse periods must be defined as integer multiples of $\tau_{\min}$. This in turn places constraints on the allowed values of δ_j in a sequence. Irrational values of fractional pulse locations have intrinsic conflict with digital modulation, thus mandating alternate approaches.

In order to overcome this challenge, we have identified the Walsh functions as a mathematical basis that is compatible with digital sequencing hardware. In this section we discuss the Walsh functions and the WDD derived from them.

A. Walsh functions and WDD

The Walsh functions are a family of binary valued (± 1) piecewise-constant functions on the $[0, 1]$ interval [53]. They found a place in engineering in the 1960s, when they started being applied to problems ranging from communications and signal analysis to image processing, and noise filtering [48, 49]. The Walsh functions come in a variety of labeling conventions, including the Hadamard, "sequency", and Paley (or dyadic). In particular, the sequency ordering counts the number of "switchings" of the Walsh functions. In this paper, however, we focus on the Paley ordering, given in terms of the so-called *Rademacher functions* [54], which are defined as

$$\mathsf{R}_j(x) = \mathrm{sgn}\left[\sin\left(2^j\pi x\right)\right], \quad j \geq 0,$$

and correspond to periodic digital switchings between ± 1 over $[0, 1]$ with the "rate" 2^j. The Walsh function of Paley order n, $\mathsf{W}_n(x)$, is then defined as [55]

$$\mathsf{W}_n(x) = \prod_{j=1}^m \mathsf{R}_j(x)^{b_j}, \quad x \in [0, 1], \qquad (6)$$

where we denote by $(b_m b_{m-1}...b_1)_2$ the *binary* representation of n, that is, $n = b_m 2^{m-1} + b_{m-1}2^{m-2} + \ldots + b_1 2^0$. The actual number, r, of Rademacher functions used in constructing a Walsh function, is the *Hamming weight* (number of non-zero binary digits) of n. For illustration, the Walsh functions $\{\mathsf{W}_n(x)\}_{n=1}^{32}$ in the Paley ordering are shown in Fig. 1(a). The Walsh functions form an orthonormal basis over $[0, 1]$, that is,

$$\int_0^1 \mathsf{W}_n(x)\mathsf{W}_m(x)dx = \delta_{mn}, \qquad (7)$$

where δ_{mn} is the Kronecker delta. Any integrable function $f(x)$ defined over $[0, 1]$ has a convergent Walsh-Fourier expansion similar to the usual Fourier series:

$$f(x) = \sum_{n=0}^\infty a_n \mathsf{W}_n(x), \quad a_n = \int_0^1 f(x)\mathsf{W}_n(x)dx. \qquad (8)$$

This allows us to expand any sequence propagator in the Walsh basis. We note in passing that the Walsh basis
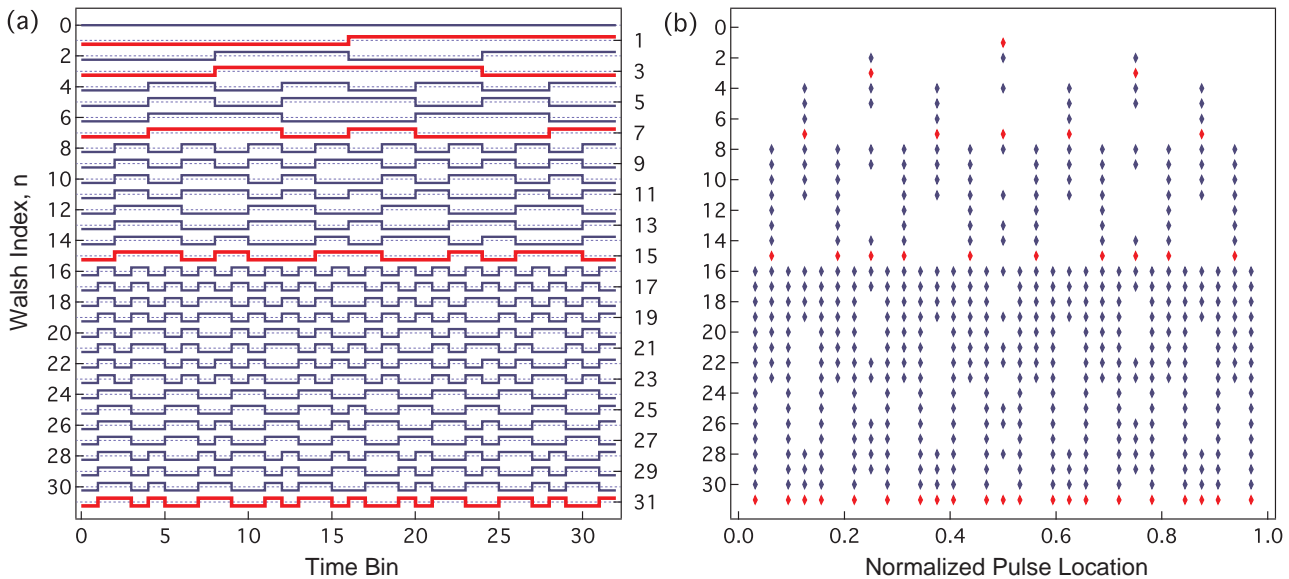
Figure 1: (color online) Walsh functions listed by Paley ordering. (a) The first 32 Walsh functions. (b) The normalized pulse locations, δ_j, corresponding to the digital transition points of Walsh functions. Sequences WDD_{2^r-1} are highlighted in both panels.

is technically over-complete while the Rademacher functions form an orthonormal system [54].

We define the *Walsh DD* (WDD) sequences directly based on the Walsh functions: that is, for a given running time τ, the sequence WDD_n is determined by the control propagator

$$y(t) = \mathsf{W}_n(t/\tau), \ \ t \in [0,\tau]. \tag{9}$$

Equivalently, we may specify the normalized pulse locations for WDD_n as the switching locations of $\mathsf{W}_n(x)$, whereas the constant pieces of the Walsh functions correspond to interpulse delays. By construction, the number of pulses used in WDD_n is given by the sequency s associated with W_n. More explicitly, if in binary notation $n = (b_m b_{m-1}...b_1)_2$ and $s = (g_m g_{m-1}...g_1)_2$, then $g_i = b_i + b_{i+1} \bmod 2$ [56], known as a Gray code.

Both the total running time, τ, and the minimum switching time, $\tau_{\min} = \tau/2^m$, impose constraints on the accessible WDD sequences. First, we must have $m = O(\log_2 n)$, the number of digits in the binary representation of n for a given WDD sequence. Once these parameters are fixed, there is a maximum value $n \approx O(\tau/\tau_{\min})$ for which WDD_n is viable. Reversing this argument, by focusing on a finite set of WDD_n and a fixed running time, we automatically accommodate a finite minimum switching time.

Figure 1(b) depicts the pulse locations plotted as δ_j for a normalized sequence duration $\tau = 1$ extracted from the $\mathsf{W}_{32}(t/\tau)$. Some familiar DD sequences are immediately identified: WDD_1 is the spin echo sequence and WDD_3 is the two-pulse Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence. We will return to these correspondences in Section III C.

B. Error suppression properties

A primary aim in the construction of DD sequences is to increase the order of error suppression, resulting in better cancellation of noise at sufficiently low frequencies [10, 40, 46, 57, 58]. In order to characterize the error suppression capabilities of WDD sequences, we begin by noting that the Walsh functions, constructed as products of r Rademacher functions [Eq. (6)] satisfy the following important equality (see Appendix for an explicit proof):

$$\int_0^1 [\mathsf{R}_{j_1}(x) \cdots \mathsf{R}_{j_r}(x)] x^k dx \equiv 0, \quad k = 0, \ldots, r-1, \tag{10}$$

where the $\{j_k\}$ indices label the locations of the r nonzero digits in the binary representation of n. Equivalently, the monomials x^k have *no* Walsh-Fourier component a_n as long as n has a Hamming weigth of at least k. Let us focus on the Fourier transform of the propagator function for WDD_n. Using Eqs. (2) and (9) we have:

$$\tilde{y}(\omega\tau) = \tau \int_0^1 \mathsf{W}_n(x) e^{i\omega\tau x} dx,$$

$$= \tau \int_0^1 \mathsf{R}_{j_1}(x) \cdots \mathsf{R}_{j_r}(x) \sum_{k=r}^{\infty} \frac{(i\omega\tau x)^k}{k!} dx,$$

where the powers of x below r have been eliminated by using Eq. (10). Together with Eq. (2), the above equation implies that

$$F_{\text{WDD}_n}(\omega\tau) \propto (\omega\tau)^{2(r+1)}. \tag{11}$$

Thus, WDD$_n$ *suppresses errors up to order* r. Expressed differently, all WDD$_n$ derived from Walsh functions composed of r Rademacher functions exhibit the same order of error suppression. The relationship between r, the Hamming weight of the Walsh order n, and the order of error suppression is one of the main results of this work.

The error-suppressing properties of WDD$_n$ can also be directly established upon obtaining a compact expression for the corresponding filter function (in analogy to [40]), which is possible by using Eq. (3) in combination with a change of variable $e^{i\omega\tau/2^m} \mapsto z$. For $n = (b_m \cdots b_0)_1$, the filter function for WDD$_n$ is given by

$$F_{\text{WDD}_n}(\omega\tau) = \left| (1-z) \prod_{j=1}^{m} (1 + (-1)^{b_j} z^{2^{j-1}}) \right|^2$$

$$= 4^{m+1} \sin^2(\omega\tau_{\min}/2)$$

$$\times \prod_{\{j|b_j=1\}} \sin^2(2^{j-2}\omega\tau_{\min}) \times \prod_{\{j|b_j=0\}} \cos^2(2^{j-2}\omega\tau_{\min}). \quad (12)$$

where the product chooses sin or cos factors based on the j-th binary digit b_j of n. Interestingly, the trigonometric factors appearing in Eq. (12) have complementary implications for the filter function and ultimately for error suppression. The smallest common period among these terms is at $\omega\tau_{\min} = 2\pi$. Each sine term is linear in $\omega\tau$ at $\omega = 0$ and each factor of $\sin^2(2^{j-2}\omega\tau_{\min})$ subsequently contributes a factor proportional to $(\omega\tau)^2$ to the filter function. These combine to give $F_{\text{WDD}_n}(\omega\tau) \propto (\omega\tau)^{2(r+1)}$, consistent with Eq. (11). Also, all sine terms have zeroes at $\omega\tau_{\min} = k\pi$ for integer k, that correspond to zeroes at $\omega = 2^m k\pi/\tau$. In contrast, they all carry large maxima occurring at $\omega = O(\tau_{\min}^{-1})$. The cosine terms on the other hand have no effect on the filter function around $\omega = 0$, but drop to zero at somewhat higher frequencies (at which the sine terms might be significantly large). This suggests that they can reduce the spikes of the sine terms and contribute to increasing the bandwidth Ω_{p}.

With the insights above, we find that sequences WDD$_{2^r-1}$ *have the lowest value of sequency* (pulse number) for a given order of error suppression (among the Walsh family). This is because the binary representation of a digital integer $n = 2^r - 1$ requires all $b_i = 1$. Figure 2 depicts the filter functions (in a log-log scale) associated with WDD$_{2^r-1}$ for $r = 1, \cdots, 5$ (as identified in Fig. 1), where we illustrate how the low-frequency rolloff is increased by increasing r. We compare the filter functions for these sequences to those of Uhrig DD (UDD), known to provide s-order error suppression given s pulses [10, 40].

In Fig. 2(b) we plot the filter functions for all WDD sequences with $n < 32$ and $r = 4$. We see that *all* have the same low-frequency rolloff, validating the assertion that the order of error suppression in a WDD sequence is determined by r.

The actual coherence error, $1 - W$, associated with a given DD sequence can be calculated by using the relevant power spectrum $S_\beta(\omega)$ for the noise in Eq. (4). The

presence of spikes in the filter function at $\omega = O(\tau_{\min}^{-1})$ (not shown) implies that we can only expect error suppression as long as $\omega_c \tau_{\min} < 1$. We can guarantee this by shrinking $\tau_{\min}$ through use of more frequent decoupling pulses, as far as the technological limitations on pulse rates permit this (see Ref. [44] for a discussion of the lower bounds on error suppression in DD at constrained minimum switching time).

Fig. 2(c) depicts the performance of WDD$_{2^r-1}$ sequences for a specific illustrative case where a $1/\omega^2$ noise power spectral density $S(\omega)$ persists up to a Gaussian high-frequency cutoff ω_c (given in units of inverse time interval $1/\tau$). We have scaled the noise to have a value comparable to that derived from the low-frequency behavior of nuclear spin diffusion in singlet-triplet qubits [59]. We observe, as expected, that not only does the calculated $1/e$ coherence time increase with n, but the slope of the error accumulation at short times also increases with n. This is a manifestation of the increasing order of error suppression with r described above, and shows the utility of high-order WDD sequences for quantum computing applications where minimizing the error probability is of utmost importance [60].

C. The Walsh sequence suite

It is clear from the previous discussion that a number of familiar integer-based DD sequences can be identified as special instances of WDD sequences. These sequences are described next for reference and are summarized in Table I. For example, periodic DD involves repetitive application of uniformly spaced π pulses [1]. A Walsh function composed of a single Rademacher function $R_k(x)$ ($r = 1$) corresponds to a PDD with $2^{r+1} - 1$ pulses. CPMG sequences are modifications of PDD in which the first and last free-evolution periods are half the duration of the interpulse period. We refer to a CPMG sequence with n pulses as CPMG$_n$. A Walsh function of the form $R_k(x)R_{k-1}(x)$, with $k \geq 2$, corresponds to $r = 2$ and CPMG$_{2^k}$.

We also observe that the sequences defined by WDD$_{2^r-1}$, identified in the previous section, are in fact concatenated DD (CDD) sequences. The latter are known to allow arbitrary orders of error suppression in DD for arbitrary noise models by recursive embedding of a sequence in a "larger" one [8, 61]. As long as a first-order DD sequence can be implemented in a system, concatenating it with itself results in higher orders of cancellation, at the expenses of increased sequence length. For a purely dephasing environment as examined here, the natural first-order DD sequence (the "base" for concatenation) is the well-known spin-echo sequence, characterized by two equal intervals separated by a π pulse; each higher level of concatenation corresponds then to embedding the spin echo in a larger spin echo [43, 46]. Using induction, we can show that each concatenation level corresponds to multiplying the sequence
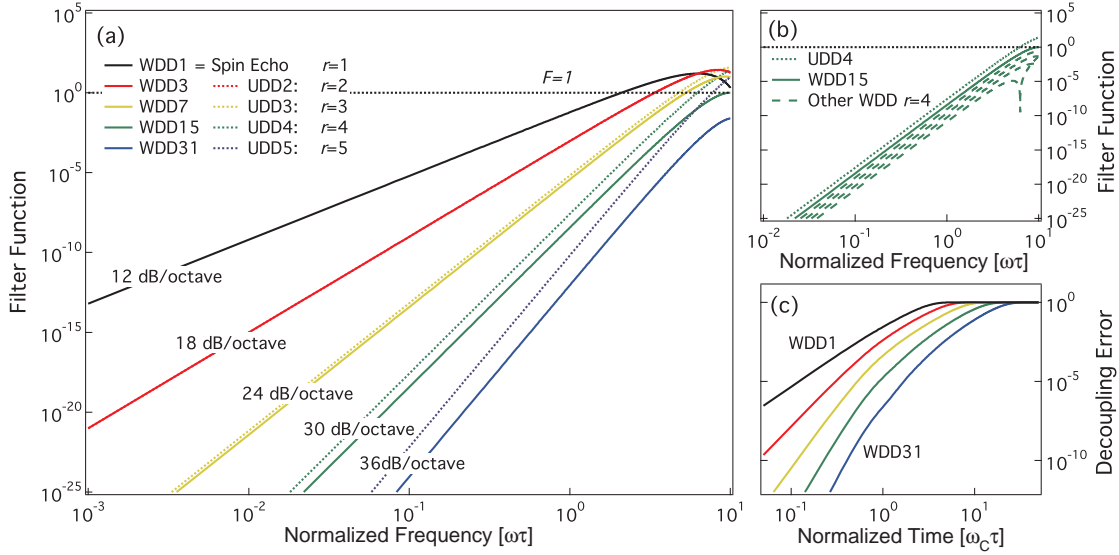
Figure 2: (color online) Error suppression properties of WDD sequences. (a) Filter functions calculated for sequences derived from $W_{2^r-1}(t/\tau)$ and corresponding to CDD_r, with fixed total sequence duration, τ. For each value of r, the number of fundamental Rademacher functions that are needed to produce the sequence increases, and the order of error-suppression, demonstrated graphically by the low-frequency rolloff of the filter function, increases by 6 dB/octave. Dashed lines correspond to the UDD sequence providing the same order of error suppression. (b) Filter functions for all WDD sequences with $n < 32$ and $r = 4$. UDD4 is represented by a dashed line and WDD15 by a solid line. Dashed lines correspond, left-to-right, to WDD23, WDD27, WDD29, WDD30. (c) Decoupling error as a function of dimensionless frequency normalized by the cutoff frequency, ω_c in $S_\beta(\omega) = \alpha \omega^{-2} \exp\left[-(\omega/\omega_c)^2\right]$. Here, $\alpha = 5 \times 10^{16}$, and frequency ranges are set as in Ref. [59]. Sequences are same as in (a), and increase in n from left to right.

n	WDD_n	Number of Pulses (Sequency)	Rolloff (dB/octave)
2^r	PDD	$2^{r+1} - 1$	$6 \times (1+1)$
$2^{r-1} + 2^r$	CPMG	2^r	$6 \times (2+1)$
$2^r - 1$	CDD	$\lceil \frac{2^{r+1}-2}{3} \rceil$	$6 \times (r+1)$

Table I: Familiar sequences among WDD_n. $\lceil x \rceil$ denotes the ceiling function, that is, the smallest integer not less than x. The exact expression for the number of pulses in CDD was obtained with the aid of [62].

propagator by a Rademacher function. This results in a product of all Rademacher functions up to order r, thus $b_1 = \ldots = b_r = 1$, which corresponds to WDD_{2^r-1}. In summary,

$$\mathrm{CDD}_r \leftrightarrow \mathsf{R}_1(x) \cdots \mathsf{R}_r(x) \leftrightarrow \mathrm{WDD}_{2^r-1}. \quad (13)$$

Again, the WDD sequences that correspond to CDD have a propagator corresponding to a product of Rademachers of all orders from 1 to r, giving r-order error suppression with the lowest value of sequency, s.

The WDD family significantly enlarges the sampling set of DD sequences relative to the much more constrained CDD. If the minimum switching time is constrained to $\tau_{\min} = \tau/2^m$, then all the 2^m WDD_n sequences with $n = 0, \cdots, 2^m - 1$ are viable. In contrast, the subset of viable CDD sequences for the same fixed

total time contains only m sequences, an exponentially smaller number. (However, it remains an even smaller subset of *all* possible digital DD sequences, which contains 2^{2^m} sequences. This potential advantage will be discussed in Sec. V.)

WDD thus forms a unifying mathematical framework for the generation of digital DD sequences, including many familiar sequences and a large variety of novel sequences. Using insights derived above, we may fully characterize the structure of arbitrary Walsh functions, and therefore arbitrary WDD modulations. By the structure of the Walsh functions themselves, *all* WDD sequences can be produced recursively from free evolution by combining the following two intuitive operations:

• *Repetition*, where the sequence propagator $y(t)$ is repeated identically to produce a longer sequence:

$$y(t) \mapsto \begin{cases} y(t/2), & t < \tau/2, \\ y((t-\tau)/2), & \tau/2 < t < \tau. \end{cases}$$

• *Concatenation*, where the sequence propagator $y(t)$ is repeated with the sign reversed, still yielding a longer sequence:

$$y(t) \mapsto \begin{cases} y(t/2), & t < \tau/2, \\ -y((t-\tau)/2), & \tau/2 < t < \tau. \end{cases}$$

The actual implementation of repetition and concatenation may involve inserting a π pulse in the middle of the formed sequence to account for the required sign change at $t = \tau/2$. In general, WDD_{2n} is constructed by repeating WDD_n, whereas WDD_{2n+1} is constructed by concatenating WDD_n. For instance, $\text{WDD}_{30} = (\text{WDD}_{15})(\text{WDD}_{15})$ [repetition] and $\text{WDD}_{31} = (\text{WDD}_{15})\pi(\text{WDD}_{15})$ [concatenation], but note the reversal of the role of the middle pulse in $\text{WDD}_{14} = (\text{WDD}_7)\pi(\text{WDD}_7)$ [repetition] or $\text{WDD}_{15} = (\text{WDD}_7)(\text{WDD}_7)$ [concatenation].

Each concatenation increases the rolloff slope by one order (by contributing an additional sine factor in F_{WDD_n}), whereas each repetition does not improve the rolloff (a cosine factor) and may instead increase the bandwidth of the filter, Ω_p. This produces a diversity of design features that can be used to improve coherence *times* (e.g. the $1/e$ decay time denoted T_2) in addition to coherence *values* (or, equivalently, error rates) by choosing the appropriate Walsh basis function.

IV. BEYOND CLASSICAL PHASE NOISE AND QUANTUM MEMORY

In this section, we consider different extensions of the idea of Walsh modulation to control settings more general than examined thus far.

A. Quantum phase noise

Decoherence in a quantum system is most generically described by the interactions with a *quantum* environment. We focus first, as before, on a single qubit (see below for multi-qubit extensions). In the absence of control and an internal system Hamiltonian, evolution under the noise Hamiltonian of Eq. (1) is then replaced by evolution under an open-system dephasing Hamiltonian of the form

$$H = H_{SB} + H_B, \quad H_{SB} = \sigma_z \otimes B_z,$$

where physically H_{SB} represents the interaction term, and B_z and H_B are generic operators acting on the environment, respectively. Typically, H_{SB} causes entanglement between the system and the environment, which eventually results in loss of phase coherence and mixed qubit states.

The WDD sequences can be applied to the above general quantum dephasing scenario. Under a sequence of π pulses, the evolution is identically generated by a piecewise-constant Hamiltonian $\pm\sigma_z \otimes B_z + H_B$. Here, the filter function formalism is not exactly applicable in general. An important exception, as noted, is provided by the case of a bosonic bath [1, 10, 43], and results accurate up to the second order in H_{SB} have been established in [63] for an arbitrary quantum dephasing environment under periodic DD (see also [64] for exact results on more

general classical noise models). Nonetheless, a description based on an effective Hamiltonian and the Magnus expansion can be used to approximate the evolution of the system provided that $\|B_z\|$ (operator norm of B_z) is sufficiently small. Making note of the concatenated and repetitive structure of WDD and following Ref. [61], we can show that the norm of the effective Hamiltonian for the evolution under WDD_n is given by:

$$\|H_{\text{WDD}_n}\| = O[\|B_z\|\tau^r \max(\|B_z\|, \|H_B\|)^r], \quad (14)$$

where the strength of the bare interaction $\|B_z\|$ is scaled by a factor proportional to τ^r, with r denoting, as before, the Hamming weight of n. This implies that the expansion of the propagator for the evolution in τ starts with the power τ^{r+1}. The fidelity loss can be defined as a distance between the ideal state (original state in DD) and the actual propagated state of the system at the end of the evolution (after tracing out the environmental degrees of freedom) [65]. In the limit of small actions τH_{WDD_n}, the fidelity loss scales with $\|\tau H_{\text{WDD}_n}\|^2$ or, equivalently, *the qubit fidelity loss scales at worst with* $\tau^{2(r+1)}$. While this result mirrors the error cancellation properties of WDD in the filter function formalism [Eq. (11)], we emphasize that asymptotic relationships such as Eq. (14) must be understood as scaling laws that may hide (possibly large) r-dependent prefactors. This fact highlights the importance of Eq. (4) as an exact expression for Gaussian phase noise for which a counterpart does not exist in a general quantum-dephasing setting.

B. Generic decoherence

As already mentioned, DD can be applied to a large class of open quantum systems: a control propagator that regularly traverses the elements of the so-called DD group in principle allows suppression of arbitrary non-Markovian decoherence, including multi-qubit noise in many-qubit systems [2]. When the noise operators are decomposed into algebraically independent components, one may reconstruct the generic decoupling procedure through concatenation as well [61]. While the group-theoretic DD design is applicable only to first-order error suppression, concatenation allows us to combine DD sequences for different noise axes to produce a high order *and* generic decoupling procedure. This idea is the basis of a variety of recent DD schemes, such as quadratic DD (QDD) [66] and its multi-qubit variants [51, 67].

The WDD sequences can be similarly concatenated along different axes to allow suppression of general noise. For example, consider a generalized classical noise Hamiltonian acting on a qubit:

$$H_{\text{noise}} = \beta_x(t)\sigma_x + \beta_y(t)\sigma_y + \beta_z(t)\sigma_z,$$

where $\beta_i(t)$ are stochastic processes. Applying a WDD sequence with σ_x pulses effectively removes the $\beta_y(t)\sigma_y + \beta_z(t)\sigma_z$ noise component while it leaves $\beta_x(t)\sigma_x$ intact.

Embedding this sequence itself in a WDD sequence with σ_y (or σ_z) pulses removes the remaining $\beta_x(t)\sigma_x$ component as well. We can generalize this idea to define *generic Walsh dynamical DD* (GWDD) for a qubit. The control propagator $U_c(t)$ for the sequence GWDD_n is given by

$$U_c(t) = \sigma_x^{[x(t)+1]/2}\sigma_y^{[y(t)+1]/2},$$

where

$$
\begin{aligned}
x(t) &= \mathsf{R}_{j_1}(t)\mathsf{R}_{j_3}(t)\cdots,\mathsf{R}_{j_{2r-1}}(t),\\
y(t) &= \mathsf{R}_{j_2}(t)\mathsf{R}_{j_4}(t)\cdots\mathsf{R}_{j_{2r}}(t),
\end{aligned}
$$

and, as in the single-axis case, the $\{j_k\}$ indices label the locations of non-zero digits in the binary representation of n. In practice, the final sequence corresponds to applying σ_x or σ_y pulses at rates given by the the corresponding Rademacher functions, taking into account the algebraic simplifications such as $\sigma_x\sigma_x = I$ or $\sigma_x\sigma_y = \sigma_z$ and ignoring all the resulting $\pm$ signs. We note that for $n = 2^{2r} - 1$, GWDD_n reproduces the generic CDD sequence of level r described in Ref. [8], while repetitions of GWDD_n include truncated periodic CDD protocols (so-called PCDD) such as investigated in [68, 69]. Again the Walsh functions form a unifying mathematical basis for sequence construction, reproducing familiar decoupling protocols.

For a multi-qubit system where qubits interact with the environment *linearly* [2, 12], GWDD is also applicable, except that each unitary pulse operation has to be replaced with a collective version that affects all qubits simultaneously and equally. We remark in passing that the idea of Rademacher products (concatenation in particular) can in principle be extended to generic finite-dimensional control systems provided that the role of σ_x and σ_y is replaced with the generators of the corresponding DD group [70].

C. Non-identity operations

Our focus so far has been on error suppression while preserving arbitrary quantum states, hence effectively implementing the identity operator with higher fidelity than free evolution for a given duration. The periodic properties of the filter functions associated with WDD sequences also allow us to achieve suppression of frequency noise while effecting *non-trivial* (non-identity) quantum gates, close in spirit to *dynamically corrected gates* (DCGs) [12–14].

The starting point in DCG constructions is the separation of the action of a gate on a noisy system into ideal and error parts, both represented as unitary operators. Concretely, let Q and U_Q denote the ideal unitary gate that is to be implemented and the actual unitary propagator corresponding to the evolution during the control that aims to implement Q, respectively. We can write

$$U_Q = Q\exp(-iE_Q),$$

where E_Q is the error per gate (or *error action*) associated with Q and the goal is to minimize E_Q for any desired Q. The basic intuition is to use the separation of the ideal gate action and the error to mix and match error parts so that altogether the errors cancel out in a perturbative manner.

DCGs have algebraic connections to DD sequences but they remove the need for instantaneous ideal pulses and are used to implement non-identity unitary actions on the system. The general theory for constructing DCGs to higher orders for arbitrary systems appears in [14], but for our purpose we focus on an example that applies to a recent experiment on correcting errors due to laser frequency jitter in a multi-qubit entangling gate mediated by laser light [50].

Consider first the spin-echo sequence $XfXf$, where X denotes an ideal σ_x gate and f a free evolution interval. The latter can be interpreted as a *primitive* implementation of the identity gate I. Thus, we may write the spin echo as

$$I^{(1)} = XI^{(0)}XI^{(0)}, \tag{15}$$

where $I^{(0)}$ refers to a zeroth-order approximation for the identity action through free evolution, and $I^{(1)}$ refers instead to an improved approximation of the identity action with a higher order of error suppression. By recursively applying Eq. (15), or by repeating an improved identity gate, we obtain the WDD sequences corresponding to progressively better and longer approximations of the identity gate. Consider next a *specialized* scenario in which the X gates are still ideal but, instead of free evolution, we apply a target gate Q with an associated error E_Q. We also assume that the intended gate action Q and the X gates commute. We begin with the sequence $XQXQ$ whose actual propagator is given by

$$
\begin{aligned}
XQ\exp(-iE_Q)&XQ\exp(-iE_Q) = \\
&Q^2X\exp(-iE_Q)X\exp(-iE_Q), \tag{16}
\end{aligned}
$$

where $X\exp(-iE_Q)X\exp(-iE_Q)$ can be interpreted as a spin-echo sequence for free evolution with an effective Hamiltonian E_Q. Thus, if E_Q consists of terms that *anticommute* with X (i.e. $XE_QX = -E_Q$), we expect that the sequence $XQXQ$ suppresses error in a manner similar to $XIXI$, except that the resulting action will now be the non-trivial gate Q^2. Furthermore, we may concatenate or repeat this sequence to obtain higher-order suppression of errors at the expense of a longer gate sequence.

The model which we just described for a simple high-order DCG applies to the experimental setup in Ref. [50], in which a spin-motional entangling gate is applied to trapped ion qubits. The fundamental concept of this gate is that by state-selectively exciting the harmonic-oscillator motional modes of ions in a shared trapping potential, it is possible to entangle the internal spins of the qubits. The actual unitary propagator associated

with a single gate is given by

$$U_Q(t) = e^{S_N(\alpha(t)a^\dagger - \alpha^*(t)a)}Q, \qquad (17)$$

where Q is a geometric phase gate, described above, that entangles the qubits. This gate relies on disentangling the spin and motion at the end of the gate by detuning the driving force from a motional resonance by δ and setting the gate time $t = j2\pi/\delta$, where j is a positive integer. In this case the drive and motion desynchronize at t, and one precisely implements Q. Residual spin-motional entanglement corresponds to an error and we thus define the error per gate as

$$-iE_Q = S_N\left[\alpha(t)a^\dagger - \alpha^*(t)a\right].$$

The coefficient $\alpha(t)$ is the time dependent displacement due to the laser detuning δ, that is, $\alpha(t) = (\Omega/2)\int_0^t ds \exp(-i\delta s)$. Such error may be due, in practice, to the fact that δ typically carries an additive frequency error Δ for which we have

$$\alpha(t) = \frac{\Omega}{2}\int_0^t ds \exp[-i(\delta + \Delta)s].$$

In this case we find that the harmonic oscillator does not produce a "closed loop" in phase space, thus yielding an error due to residual spin-motion entanglement.

Our goal is to achieve an improved approximation of the entangling gate in Eq. (17) by cancellation of $\alpha(t)$ to a high order at the end of the evolution. Besides varying the detuning δ, the evolution of $\alpha(t)$ can also be controlled by flipping the phase of the laser-mediated optical dipole force used to drive the ion motion (a nearly instantaneous action), which effectively corresponds to switching the sign of the interaction. Consider the evolution of the system punctuated by such phase flips occuring at the times $p = \{t_j\}_{j=1}^n$, with $t_0 = 0$ and $t_{n+1} = \tau$ to denote the total gate/evolution duration. The expression for $\alpha(\tau)$ is then given by

$$\alpha(\tau) = \frac{\Omega}{2}\sum_{j=0}^n (-1)^j \int_{t_j}^{t_{j+1}} ds \exp[-i(\delta + \Delta)s],$$

and we note that

$$\int_{t_j}^{t_{j+1}} ds \exp[-i(\delta + \Delta)s] = -i\frac{e^{-i(\delta+\Delta)t_j} - e^{-i(\delta+\Delta)t_{j+1}}}{\delta + \Delta}.$$

In reality, the mismatch Δ may follow a probability distribution $P(\Delta)$ and we can write the expected value of $|\alpha(\tau)|^2$ as

$$\overline{|\alpha(\tau)|^2} = \frac{\Omega^2}{4}\int_{-\infty}^\infty \frac{F_p((\delta+\Delta)\tau)}{(\delta+\Delta)^2}P(\Delta)d\Delta, \qquad (18)$$

where we have used the expression for the filter function in Eqs. (2) and (3). This expression resembles the equation for the decoherence error associated with a DD
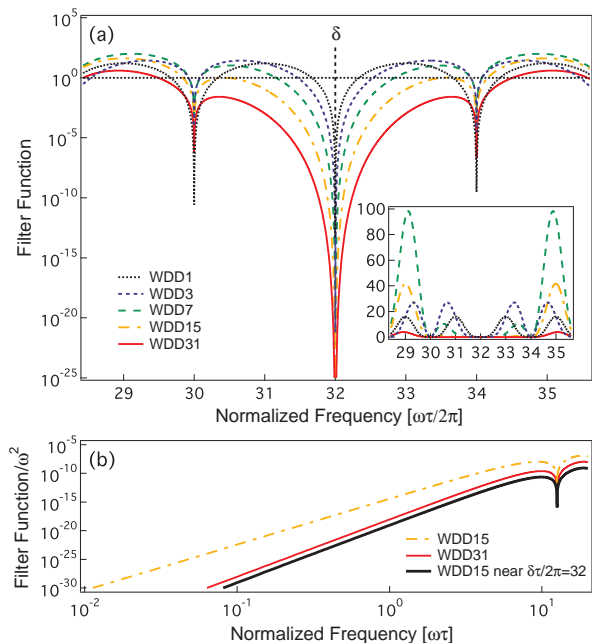

Figure 3: (color online) Noise suppression at fixed frequency. (a) Filter function about normalized frequency δ on a semilog scale for various WDD sequences, resulting in a notch where noise is suppressed. The notch bandwidth increases with sequency and the order of error suppression about δ increases with r. Inset: Same filter functions plotted on a linear scale. (b) Filter function normalized by ω^2 for WDD15 and WDD31. The solid thick line shows WDD15 at $\delta/2\pi = 2^5$, where the horizontal axis represents detuning from δ. Here the order of error suppression as a function of detuning has increased by one order and is comparable to that for WDD31 near zero frequency. See text for details.

sequence applied at the same times, except that the *argument of the filter function in the integrand is shifted by a value δ*, and instead of the power spectrum $S_\beta(\omega)$ the error probability density $P(\Delta)$ is used. To cancel the effects of Δ we require the filter function to have a zero of high multiplicity at $\omega = \delta$.

Such a zero may be realized in the filter function of the WDD sequence; from Sec. III B, we recall that the filter function $F(\omega\tau)$ has a natural period (translational invariance, see Fig. 3) at $\tau_{\min}/2\pi$, corresponding to $\tau/(2^{m+1}\pi)$. This produces the requisite zeroes in the filter function at fixed values of δ, a feature intrinsic to the formulation of WDD.

In order to gain a quantitative understanding, let us calculate the Fourier transform of the sequence propagator,

$$\tilde{y}(\omega\tau) = \int_0^\tau dt \mathsf{W}_n(t/\tau)e^{i(\delta+\Delta)t/\tau}. \qquad (19)$$

We extract the filter function's dependence on Δ near δ

by use of an equality that is similar to Eq. (10),

$$\int_0^1 dx \prod_{\{k\}} \mathsf{R}_k(x) e^{i2^{r+1}\pi x} \sum_{i=0}^r a_i x^i = 0, \qquad (20)$$

where the set $\{k\}$ has r elements in it, $k \geq 1$ and $\max(k) \leq r$. A sketch of the proof of Eq. (20) appears in the Appendix. If the frequency $\delta/2\pi = 2^k/\tau$, then Eq. (19) can be written as

$$\tilde{y}(\omega\tau) = \tau \int_0^1 dt' \prod_{\{k\}} \mathsf{R}(k,t') e^{i2^{r+1}\pi t'} \sum_{m=r+1}^\infty \frac{(i\Delta' t')^m}{m!},$$
(21)

where we have defined $t' = t/\tau$ and $\Delta' = \Delta\tau$. This expression indicates that $\tilde{y}(\omega\tau)$ scales like Δ^{r+1} in the vicinity of δ. Thus, the slope of the rolloff as a function of detuning is set by r, giving increasing robustness to fluctuations in detuning from δ. For fixed r, the bandwidth of the "notch" increases with sequency, as does the equivalent Ω_p. These properties are expressed in Fig 3.

While the filter function is translationally symmetric, we see that the form of filter performance around δ is not the same as it is around $\omega = 0$, due to the factor of ω^{-2} appearing in Eq. 2. An expansion as a function of Δ therefore yields improved performance relative to the zero-frequency rolloff when accounting for this factor.

It thus follows that for the specific case treated here by flipping the laser phase at times corresponding to the WDD_n, the error in the detuning δ can be suppressed up to order $r+1$, where as before r is the Hamming weight of n. This general approach was studied experimentally in Ref. [50]. Interestingly, this sequence can be interpreted as a high-order DCG where the primitive gates Q are defined as $\tau_{\min}$-long periods of evolution at detuning $\delta = 2\pi/\tau_{\min}$ and the X gates are implemented by phase flips. We can interpret all the resulting DCGs generated by WDD sequences as repetitions and concatenations of the basic sequence $XQXQ$, resulting in a corrected gate Q^{2^m} which is the desired entangling operation but has an error that scales with $\Delta^{2(r+1)}$.

V. BENEFITS OF WDD OVER OTHER OPTIMIZED APPROACHES

Interest in the DD community has recently focused on optimized pulse sequences producing high-order suppression of noise either through analytical approaches (e.g. UDD [10]), numerical optimization (e.g. Locally Optimized ODD (LODD) [20], Optimized noise Filtration DD (OFDD) [42], Bandwidth Adapted DD (BADD) [44]), or combinations of optimization and concatenation strategies (e.g. concatenated UDD (CUDD) [71], QDD [66], Nested UDD (NUDD) [51]). These sequences bring many benefits in terms of resource-efficient DD, as they generally optimize error-suppression against pulse number.

Studies have shown, however, that the extraordinary benefits provided by optimized sequences are largely suppressed in the presence of realistic constraints such as imperfect control pulses [21, 72, 73], digital clocking [16], and timing limitations [43]. Moreover, noise suppression benefits have been shown to be minimal for $S_\beta(\omega)$ dominated by low-frequency noise and exhibiting slowly decaying high-frequency tails [40, 42, 74].

The WDD sequences possess benefits over existing approaches along two primary metrics:
- Efficient hardware sequencing;
- Restricted search space for sequencing.

We will separately explore each of these benefits next.

A. Efficient hardware sequencing

Current experiments in DD employ a user-programmed microprocessor and a complex hardware chain (e.g., a signal generator controlled by a programmable logic device responsible for pulse timing and under PC control). This is appropriate for demonstration experiments, but fails to provide a scalable solution due to both sequencing challenges and the difficulty of input-output (I/O) in complex or quantum systems. We are therefore interested in finding solutions permitting all sequence generation to be performed at the *local* level.

Once we accept this consideration, our metrics of efficiency change relative to the majority of published literature. When considering optimized DD, sequencing is complex and will likely require either a local microprocessor to decode instructions and apply a DD sequence, or multiple high-bandwidth communication pathways to external controllers. Both situations pose challenges in terms of significant local power dissipation in control hardware and energy inflows associated with I/O pathways. The energy expended in the number of pulses applied may contribute only a small amount of the total local power dissipation given these considerations. Given a presumed need for local DD sequence generation we thus arrive at digital sequencing complexity as the relevant metric for efficiency.

WDD meets this challenge, its primary benefit being that the Walsh functions are easily produced using simple digital circuits [48]. We identify Rademacher function generators (square-wave generators) as basic hardware resources for the physical-layer implementation of dynamical error suppression, realizing the control propagator in real-time. Rademacher functions may be generated in hardware with relative ease from a distributed clock signal, and may be added/multiplied together via hardware logic to generate Walsh toggling frames (for instance Harmuth's array generator [48]). The number of Rademacher functions required to achieve a given error suppression may therefore be deemed a relevant quantitative means of establishing the sequencing complexity.

As we have shown, a WDD sequence derived from r Rademacher functions cancels the first $r - 1$ orders of dephasing noise, and the sequences derived from

$W_{2^r-1}(x) \equiv \text{CDD}_r$ do so for the smallest value of sequency (for fixed r). WDD therefore provides the highest-possible order of error suppression with respect to the number of Rademacher functions required to implement a given DD protocol, providing efficiency in the sequencing complexity. For the same sequencing complexity r, using sequences with different n permits modification of Ω_p, giving flexible control over noise filtering capabilities. By contrast, other optimized sequences discussed above require substantially greater sequencing resources, as the Walsh-transformation of a sequence such as UDD indicates the need for a large r in order to generate the requisite $\{\delta_j\}$.

Furthermore, the structure of WDD and Walsh function generation is compatible with a nested or concatenated structure of decoupling about multiple axes, which enables the generation of GWDD sequences for generic decoherence. In the special case of a large separation of timescales for different decoherence processes, GWDD also allows for efficient bit-stacking of Walsh functions derived from similar hardware, but executed with different time-bases. Finally, we note that certain sequences constructed by recursively repeating the same base sequence many times may find value in circumstances where long-time storage and low-latency memory interrupts are desired in addition to improved low-frequency error suppression. This topic will be addressed in a separate manuscript.

Once a given Walsh function (or set of Walsh functions) is generated, conversion to WDD in real time may be achieved via hardware differentiation or edge-triggering of separate circuitry producing pre-programmed control pulses. Overall complexity is reduced by separating *control sequencing* from *pulse generation*.

Using Walsh functions for generation of the control propagator and hardware techniques for the triggering of applied pulses therefore provides a means to efficiently realize dynamical error suppression at the local level. Through this approach the need for access to a user-controlled microprocessor in order to send pulse-sequence commands to hardware, or for a local processor to interpret externally generated commands may be obviated: all control and timing may be performed using relatively simple digital logic realized in an Application Specific Integrated Circuit. A system protected by WDD sequences may run semi-autonomously, implementing pre-selected WDD sequences, or it may be programmed externally, where the only communicated information required could be the value of n in binary (or as Gray code), a repetition number, and a trigger signal starting the sequence.

A significant challenge in such schemes is suppression of spurious rising/falling edge signals arising from, say, propagation delays in digital circuitry. However, latching WDD sequences to the clock signal should aid suppression of these signals. Accounting for nonzero-duration control pulses with τ_π an integer multiple of the clock period, may also be achieved through use of latching logic to introduce hardware delays while the control operations are applied. However, the effect of such delays on WDD performance would need to be characterized in detail. The creation of an optimized *hardware-based* generator of WDD sequences remains a problem for future study.

B. Restricted search space

While there already exists a plethora of DD procedures within the bang-bang limit, in practice not all these procedures will be suitable to the operating range of physical systems or the demands of quantum information processing protocols. This is true despite the fact that the objective in all such procedures is identical: maximize the fidelity of preserving an arbitrary quantum state for a desired time. For sufficiently simple experimental systems, an empirical optimization procedure (such as LODD) can sample over the space of all permissible DD pulse sequences and search for an optimal one. This idea has also been explored numerically in the above-mentioned OFDD and BADD protocols, as well in optimized DD for power-law spectra [74]. These approaches grow significantly in complexity as the pulse number increases, for instance in a large scale system requiring long-term storage. The central idea in WDD – that sequences using digital timing are made by attaching or repeating smaller sequences recursively – can be instrumental in significantly reducing the search space for finding optimal DD sequences empirically or numerically. We briefly outline potential implications of this idea here.

The canonical model of single-qubit dephasing with a given noise power spectral density used through most of this paper retains considerable structure in choosing when to apply the X pulses. This persists even when enforcing digital timing conditions as below. Given $N = 2^m$ time-bins, each of duration $\tau_{\min}$, one may elect to either apply or not apply a pulse, thus generating a space of $2^N = 2^{2^m} = 2^{2^{\tau/\tau_{\min}}}$ possible sequences. Performing a complete search over this space becomes extremely challenging as the total duration τ of the DD procedure grows, as observed in numerical approaches to generating randomized DD protocols [38, 75] and to optimized sequences [20, 42].

WDD provides a natural means to reduce the search space for the generation of dynamical error suppression sequences as there are only $N = 2^m = 2^{\tau/\tau_{\min}}$ WDD sequences within the operational constraints of the problem. It also provides an intuitive analytical framework allowing pre-selection of certain sequences within WDD, further reducing the search space. For instance, a search might exclude all WDD sequences with $r < r_{\min}$, as they will be known to provide insufficient low-frequency noise suppression. With such constraints, and the relative simplicity of implementing WDD in hardware, these sequences become especially attractive for an empirical search in which the performance of each sequence is tested in an actual experimental setting.

Perhaps even more interesting are the options offered when Walsh DD is extended to generic decoherence of a larger set of qubits in a quantum memory. Recently, an efficient perturbative DD procedure for this task has been explored (NUDD) that utilizes $(D_S^2)^r$ pulses for rth order decoupling of a generic $D_S = 2^{n_q}$-dimensional system of n_q qubits [51]. This procedure assumes the most general form of errors, including many-body errors. However, in reality error models are generally sparse. A fundamental problem of DD of many qubit systems (closely related to quantum simulation algorithms [70, 76, 77]) is to optimize the decoupling when the error model for the interaction of the qubits is locally sparse. Using the procedure described in Sec. IV B, we can progressively produce a search space for such an optimization procedure starting from elements of a basic DD sequence and recursively building longer sequences by the WDD construction procedure (repetition/concatenation).

VI. CONCLUSION

In this manuscript we have studied the problem of reducing sequencing complexity in a dynamical error suppression framework, through consideration of *digital* modulation schemes. Towards this end we have introduced the Walsh functions as a mathematical basis for the generation of dynamical error suppression sequences. We have revealed how Walsh dynamical decoupling naturally incorporates familiar sequences (e.g. PDD, CPMG, CDD), and examined the properties of all the possible recursively structured sequences corresponding to Walsh functions.

Our analysis has demonstrated that the order of error suppression achieved by a given WDD sequence is set by the number of elementary Rademacher functions required to generate that sequence. This is manifested as a scaling of the low-frequency rolloff in the filter function as $6(r+1)$ dB/octave, with r the number of Rademacher functions appearing in the sequence. Meanwhile, the high-frequency performance of a WDD sequence, and hence its noise-suppression bandwidth, is tunable via selection of n for a given r, providing significant flexibility in sequence construction. Further we have introduced a simple means to construct WDD sequences using concatenation and repetition, and a technique to suppress general errors via GWDD.

We have shown that the Walsh functions provide efficient performance against control complexity, quantified by the value of r, and the supporting sequencing hardware. We believe that sequencing complexity will serve as a useful new metric with significant weight in system-level analyses, going far beyond standard optimization over pulse number in a DD sequence.

These considerations are likely to make the Walsh functions an attractive framework for the development of a quantum memory incorporating hardware-efficient physical-layer error-suppression strategies. Interestingly,

the problem of searching for (digital) DD sequences with high order error suppression properties is also related to finding Littewood complex polynomials with high order zeroes [78, 79], which may point to further connections with both signal processing theory and polynomial analysis and approximation. From a practical standpoint, the variety of WDD sequences realizable through simple hardware sequencing provides a flexible solution for many future experimental systems.

Appendix A: Order of error suppression

In this section we prove that for a given Paley ordering r, as long as $i \leq r$:

$$\int_0^1 \mathsf{R}_1(x) \cdots \mathsf{R}_{j_{r+1}}(x) x^i dx = 0, \qquad \text{(A1)}$$

where j_k are non-zero positive integers in increasing order. This identity can be used to show the error suppression of the WDD sequences *directly in the time domain* and can be mathematically interpreted as the vanishing of the *moments* of the Walsh functions on the unit interval.

We proceed by induction on r starting with the base case, $r = 0$:

$$\int_0^1 \mathsf{R}_{j_1}(x) dx = 0. \qquad \text{(A2)}$$

Since $j_1 > 0$, the corresponding Rademacher function R_{j_1} will be periodic and balanced on $[0, 1]$, which validates the base case in Eq. (A2).

For the inductive step, let us assume that

$$\int_0^1 \mathsf{R}_{j_1}(x) ... \mathsf{R}_{j_{r+1}}(x) x^i dx = 0, \qquad \text{(A3)}$$

where $i \leq r$. We need to prove that

$$\int_0^1 \mathsf{R}_{j_1}(x) ... \mathsf{R}_{j_{r+1}}(x) \mathsf{R}_{j_{r+2}}(x) x^i dx = 0, \qquad \text{(A4)}$$

where $i \leq r + 1$.

The sign of Rademacher function $\mathsf{R}_j(x)$ is determined by the j-th binary digit of x which we refer to as $b_j(x)$. When $i = 0$, Eq. (A4) reduces to the average (zeroth moment) of the product of Rademacher functions over the interval $[0, 1]$. We give a probabilistic argument for the cancellation of this average which can also be proven using induction. The average of the product of the Rademacher functions can be written as

$$\int_0^1 (-1)^{b_{j_1}(x)} \cdots (-1)^{b_{j_{r+2}}(x)} dx. \qquad \text{(A5)}$$

We can interpret the above integral as an expectation value (denoted below by E) of a function of a real random variable x distributed uniformly over the interval $[0, 1]$.

The corresponding (Lebesgue) probability measure dx is equivalent to products of *independent* and uniform (discrete) measures for each binary digit of x [80]. We can thus rewrite Eq. (A5) in the form

$$E[(-1)^{b_{j_1}} \cdots (-1)^{b_{j_{r+2}}}] = E[(-1)^{b_{j_1}}] \cdots E[(-1)^{b_{j_{r+2}}}] = 0.$$

We can then use the $i = 0$ case as the base for another induction proof over i, where we assume that Eq. (A4) holds for all $i' \leq r$. We refer to this as the "inner induction" assumption and proceed to prove Eq. (A4) for $i = r + 1$, which we rewrite as

$$I = \int_0^1 \mathsf{W}_n(x)\mathsf{R}_{j_1}(x)\, x^{r+1}dx. \qquad (A6)$$

Note that we have singled out the lowest order Rademacher function in the product and written the product of the rest as a Walsh function with Paley ordering n. We can rewrite the integral as

$$I = \Big(\int_0^{2^{-j_1}} - \int_{2^{-j_1}}^{2\times 2^{-j_1}} + \int_{2\times 2^{-j_1}}^{3\times 2^{-j_1}} - \cdots \Big)\mathsf{W}_n(x)x^{r+1}dx,$$

where the integrals match the positive and negative values of the Rademacher function R_{j_1}.

Making the substitution $x' = x - 2^{-j_1}$ for the negative-signed terms, we can write

$$I = \Big(\int_0^{2^{-j_1}} + \int_{2\times 2^{-j_1}}^{3\times 2^{-j_1}} + \cdots \Big)[x^{r+1}\mathsf{W}_n(x) - (x + 2^{-j_1})^{r+1}\mathsf{W}_n(x + 2^{-j_1})]dx. \qquad (A7)$$

The Walsh function W_n is constructed entirely of Rademacher functions with indices larger than j_1 having periods that commensurate with 2^{-j_1}, which results in $\mathsf{W}_n(x + 2^{-j_1}) = \mathsf{W}_n(x)$. This allows us to write

$$I = \Big(\int_0^{2^{-j_1}} + \int_{2\times 2^{-j_1}}^{3\times 2^{-j_1}} + \cdots \Big)\mathsf{W}_n(x)[x^{r+1} - (x + 2^{-j_1})^{r+1}]dx. \qquad (A8)$$

We can rewrite the above as an integral over $[0, 1]$ using the kernel $(\mathsf{R}_{j_1}(x) + 1)/2$:

$$I = \int_0^1 \mathsf{W}_n(x)\frac{\mathsf{R}_{j_1}(x) + 1}{2}[x^{r+1} - (x + 2^{-j_1+1})^{r+1}]dx,$$

where the leading x^{r+1} powers cancel when we expand $(x + 2^{-j_1+1})^{r+1}$, leaving us with an integrand in which every power of x appears with an exponent i' less than or equal to r. This matches the assumptions of our inner induction, leading to the cancellation of the integral for $i = r + 1$ and thus establishing the main result in Eq. (A1).

We also note the following result, related to the WDD filter function at higher frequencies and used in Sec. IV C of the main text:

$$\int_0^1 \mathsf{R}_1(x) \cdots \mathsf{R}_{j_r}(x)e^{i2^{r+1}\pi x}x^i dx = 0, \qquad (A9)$$

where $i \leq r$ and we have used the highest Rademacher function in the product with a trigonometric function of the same period. This result can also be proven by induction in a similar manner.

It is interesting to note that the only properties of the Rademacher functions that are used in our proof are their "frequencies" and symmetry. Thus, any family of functions that mimics the periods and the sign changes of the Rademacher function family could in principle be used in our proof.

Acknowledgments

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

[2] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**,

14

2417 (1999).

[3] P. Zanardi, Phys. Lett. **258**, 77 (1999).

[4] D. Vitali and P. Tombesi, Phys. Rev. A **59**, 4178 (1999).

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

[6] M. S. Byrd and D. A. Lidar, Phys. Rev. A **67**, 012324 (2003).

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

[8] K. Khodjasteh and D. A. Lidar, Phys. Rev. Lett. **95**, 180501 (2005).

[9] W. Yao, R. B. Liu, and L. J. Sham, Phys. Rev. Lett. **98**, 077602 (2007).

[10] G. S. Uhrig, Phys. Rev. Lett. **98**, 100504 (2007).

[11] G. Gordon, G. Kurizki, and D. A. Lidar, Phys. Rev. Lett. **101**, 010403 (2008).

[12] K. Khodjasteh and L. Viola, Phys. Rev. Lett. **102**, 80501 (2009).

[13] K. Khodjasteh and L. Viola, Phys. Rev. A **80**, 032314 (2009).

[14] K. Khodjasteh, D. A. Lidar, and L. Viola, Phys. Rev. Lett. **104**, 090501 (2010).

[15] W. Yang, Z.-Y. Wang, and R.-B. Liu, Frontiers Phys. **6**, 1 (2010).

[16] M. J. Biercuk, A. C. Doherty, and H. Uys, J. Phys. B **44**, 154002 (2011).

[17] T. D. Ladd, D. Maryenko, and Y. Yamamoto, Phys. Rev. B. **71**, 014401 (2005).

[18] E. Fraval, M. J. Sellars, and J. J. Longdell, Phys. Rev. Lett. **95**, 030506 (2005).

[19] H. G. Krojanski and D. Suter, Phys. Rev. Lett. **97**, 150503 (2006).

[20] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature **458**, 996 (2009).

[21] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Phys. Rev. A **79**, 062324 (2009).

[22] J. Du, X. Rong, N. Zhao, Y. Wang, J. Yang, and R. B. Liu, Nature **461**, 1265 (2009).

[23] J. R. West, D. A. Lidar, B. H. Fong, M. F. Gyure, X. Peng, and D. Suter, arXiv:0911.2398 (2009), 0911.2398.

[24] S. Damodarakurup, M. Lucamarini, G. Di Giuseppe, D. Vitali, and P. Tombesi, Phys. Rev. Lett. **103**, 040502 (2009).

[25] D. J. Szwer, S. C. Webster, A. M. Steane, and D. M. Lucas, J. Phys. B: At. Mol. Opt. Phys. **44**, 025501 (2011).

[26] G. A. Álvarez, A. Ajoy, X. Peng, and D. Suter, Phys. Rev. A **82**, 042306 (2010).

[27] G. de Lange, Z. H. Wang, D. Riste, V. V. Dobrovitski, and R. Hanson, Science **330**, 60 (2010).

[28] Y. Sagi, I. Almog, and N. Davidson, Phys. Rev. Lett. **105**, 053201 (2010).

[29] H. Bluhm, S. Foletti, I. Neder, M. Rudner, D. Mahalu, V. Umansky, and A. Yacoby, Nature Physics **7** (2011).

[30] C. Barthel, J. Medford, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **105**, 266808 (2010).

[31] C. A. Ryan, J. S. Hodges, and D. G. Cory, Phys. Rev. Lett. **105**, 200402 (2010).

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

[33] J. Medford, L. Cywinski, C. Barthel, C. M. Marcus, M. P. Hanson, and A. C. Gossard, arXiv:1108.3682 (2011).

[34] E. L. Hahn, Phys. Rev. **80**, 580 (1950).

[35] U. Haeberlen, *High Resolution NMR in Solids, Advances in Magnetic Resonance Series* (Academic Press, New York, 1976).

[36] L. M. K. Vandersypen and I. Chuang, Rev. Mod. Phys. **76**, 1037 (2004).

[37] L. Viola and E. Knill, Phys. Rev. Lett. **94**, 060502 (2005).

[38] L. F. Santos and L. Viola, Phys. Rev. Lett. **97**, 150501 (2006).

[39] L. Faoro and L. Viola, Phys. Rev. Lett. **92**, 117905 (2004).

[40] G. S. Uhrig, New J. Phys. **10**, 083024 (2008).

[41] P. Kuopanporttii, M. Mottonen, V. Bergholm, O.-P. Saira, J. Zhang, and K. B. Whaley, Phys. Rev. A. **77**, 032334 (2008).

[42] H. Uys, M. J. Biercuk, and J. J. Bollinger, Phys. Rev. Lett. **103**, 040501 (2009).

[43] T. E. Hodgson, L. Viola, and I. D'Amico, Phys. Rev. A **81**, 062321 (2010).

[44] K. Khodjasteh, T. Erdelyi, and L. Viola, Phys. Rev. A. **83**, 020305R (2011).

[45] J. M. Martinis, S. Nam, J. A. Aumentado, K. M. Lang, and C. Urbina, Phys. Rev. B. **67**, 094510 (2003).

[46] L. Cywinski, R. Lutchyn, C. Nave, and S. Das Sarma, Phys. Rev. B. **77**, 174509 (2008).

[47] A. Ajoy, G. A. Alvarez, and D. Suter, Phys. Rev. A. **83**, 032303 (2011).

[48] K. G. Beauchamp, *Walsh Functions and their Applications* (Academic Press, London, 1975).

[49] S. G. Tzafestas, ed., *Walsh functions in signal and systems analysis and design*, vol. 31 of *Benchmark papers in electrical engineering and computer science* (Van Nostrand Reinhold Company, New York, 1985).

[50] D. Hayes, S. M. Clark, S. Debnath, D. Hucul, Q. Quraishi, and C. M. Monroe, arXiv:1104.1347 (2011), 1104.1347.

[51] Z.-Y. Wang and R.-B. Liu, Phys. Rev. A **83**, 022306 (2011).

[52] Note that when s is odd, a final pulse must be added just before the readout to reset the qubit state.

[53] J. Walsh, Am. J. Math. **45**, 5 (1923).

[54] H. Rademacher, Math. Ann. **87**, 112 (1922).

[55] R. Paley, Proceedings of the London Mathematical Society **2**, 241 (1932).

[56] If s and n are stored in digital registers, we have `s=n^(n>>1)` where `^` denotes bitwise addition mod 2 (XOR) and `>>` denotes a shift right operation.

[57] B. Lee, W. Witzel, and S. Das Sarma, Phys. Rev. Lett. **100**, 160505 (2008).

[58] W. Yang and R.-B. Liu, Phys. Rev. Lett. **101**, 180403 (2008).

[59] M. J. Biercuk and H. Bluhm, Phys. Rev. B **83**, 235316 (2011).

[60] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[61] K. Khodjasteh and D. A. Lidar, Phys. Rev. A **75**, 062310 (2007).

[62] *The On-Line Encyclopedia of Integer Sequences*, available at http://oeis.org/A000975.

[63] C. Uchiyama and M. Aihara, Phys. Rev. A **66**, 032313 (2002).

[64] S. Kitajima, M. Ban, and F. Shibata, Journal of Physics B: Atomic, Molecular and Optical Physics **43**, 135504

(2010).

[65] D. A. Lidar, P. Zanardi, and K. Khodjasteh, Phys. Rev. A **78**, 012308 (2008).

[66] J. R. West, B. H. Fong, and D. A. Lidar, Phys. Rev. Lett. **104**, 130501 (2010).

[67] L. Jiang and A. Imambekov, arXiv:1104.5021 (2011), 1104.5021.

[68] W. Zhang, V. V. Dobrovitski, L. F. Santos, L. Viola, and B. N. Harmon, Phys. Rev. B **75**, 201302(R) (2007).

[69] N. P. Konstantinidis, W. Zhang, V. V. Dobrovitski, L. F. Santos, L. Viola, and B. N. Harmon, Phys. Rev. B **77**, 125336 (2008).

[70] P. Wocjan, M. Rötteler, D. Janzing, and T. Beth, Quantum Inf. Comput. **2**, 133 (2002).

[71] G. S. Uhrig, Phys. Rev. Lett. **102**, 120502 (2009).

[72] Z. Xiao, L. He, and W.-G. Wang, Phys. Rev. A **83**, 032322 (2011).

[73] Z.-H. Wang and V. V. Dobrovitski, arXiv:1101.0292 (2010), 1101.0292.

[74] S. Pasini and G. S. Uhrig, Phys. Rev. A. **81**, 012309 (2010).

[75] L. F. Santos and L. Viola, New J. Phys. **10**, 083009 (2008).

[76] D. Leung, J. Mod. Opt. **49**, 1199 (2002).

[77] M. Rötteler, J. Math. Phys. **49**, 042106 (2008).

[78] J. S. Byrnes and D. J. Newman, IEEE Trans. Antennas and Propagation **36**, 301 (1988).

[79] D. Berend and S. Golan, Math. Comput. **75**, 1541 (2006).

[80] M. Adams and V. Guillemin, *Measure Theory and Probability* (Birkhäuser, Boston, 1996).