

Compressing Measurements in Quantum Dynamic Parameter Estimation

Easwar Magesan, Alexandre Cooper, and Paola Cappellaro
*Nuclear Science and Engineering Department and Research Laboratory of Electronics,
Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

We present methods that can provide an exponential savings in the resources required to perform dynamic parameter estimation using quantum systems. The key idea is to merge classical compressive sensing techniques with quantum control methods to efficiently estimate time-dependent parameters in the system Hamiltonian. We show that incoherent measurement bases and, more generally, suitable random measurement matrices can be created by performing simple control sequences on the quantum system. Since random measurement matrices satisfying the restricted isometry property can be used to reconstruct *any* sparse signal in an efficient manner, and many physical processes are approximately sparse in some basis, these methods can potentially be useful in a variety of applications such as quantum sensing and magnetometry. We illustrate the theoretical results throughout the presentation with various practically relevant numerical examples.

I. INTRODUCTION

Quantum sensors have emerged as promising devices to beat the shot-noise limit in metrology and, more broadly, to perform measurements at the nanoscale. In particular, quantum systems can be used to perform parameter estimation, where the goal is to estimate a set of unknown parameters by manipulating the system dynamics and measuring the resulting state. A typical scheme for parameter estimation can be cast in the form of *Hamiltonian identification*, whereby one couples the quantum system to external degrees of freedom so that the parameter of interest is embedded in the Hamiltonian governing the system evolution. Estimates of the desired parameter can then be obtained by estimating the relevant quantities in the Hamiltonian. There can be significant advantages in performing sensing using quantum systems, for instance, gains in terms of sensitivity [1] and both spatial [2] and field amplitude resolution [3]. In general however, parameter estimation can be a difficult problem, especially when the parameters and Hamiltonian are time-dependent. In addition, quantum measurements can be time-intensive and thus a costly resource for practical parameter estimation schemes. The goal of this paper is to provide methods for performing dynamic parameter estimation in a robust manner, while significantly reducing the number of required measurements for high fidelity estimates.

The general scenario we are interested in is when the Hamiltonian of the system can be written as

$$\hat{H}(t) = \left[\frac{\omega_0}{2} + \gamma b(t) \right] \sigma_z, \quad (1.1)$$

where we have set $\hbar = 1$ and we are interested in reconstructing over some time interval $I = [0, T]$ a field $b(t)$, which is coupled to the qubit by an interaction of strength γ . To make the presentation more concrete, throughout we will consider a specific application of dynamic parameter estimation; quantum magnetometry. Applications of magnetic field sensing with quantum probes can be found in a wide variety of emerging research areas such as biomedical imaging [4], cognitive neuroscience [5, 6],

geomagnetism [7], and detecting gravitational waves [8]. We emphasize that, although we will use the language of magnetometry throughout the rest of the presentation, the methods we propose can be applied in generality as long as Eq. (1.1) describes the evolution of the system, up to redefinition of the constants. In the setting of magnetometry, we assume our quantum sensor is a single spin- $\frac{1}{2}$ qubit operating under the Zeeman effect, an example of which is the NV center in diamond [9]. We expect that many of the methods discussed in the context of a spin qubit could be adapted for other types of quantum sensors, such as superconducting quantum interference devices (SQUID's) [3], nuclear magnetic resonance (NMR) and imaging (MRI) techniques [10, 11], and quantum optical and atomic magnetic field detectors [12, 13].

Let $b(t)$ represent the magnetic field of interest; setting the gyromagnetic ratio, γ , equal to 1 and moving into the rotating frame gives the unitary evolution

$$e^{-i \int_0^t H(t') dt'} = e^{-i \int_0^t b(t') dt'} \sigma_z. \quad (1.2)$$

If $b(t)$ is constant, $b(t) = b$, one can use a simple Ramsey experiment [14] to determine b . One realization of the Ramsey protocol consists of implementing a $\frac{\pi}{2}$ pulse about X , waiting for time T , implementing a $-\frac{\pi}{2}$ pulse about Y , and performing a measurement in the computational basis $\{|0\rangle, |1\rangle\}$. Repeating these steps many times allows one to gather measurement statistics and estimate b since the probability p_0 of obtaining outcome 0 is

$$p_0 = \frac{1 + \sin\left(\int_{t=0}^T b dt\right)}{2} = \frac{1 + \sin(bT)}{2}. \quad (1.3)$$

Under the assumption that $bT \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, we can unambiguously determine b from the above equation. When $b(t)$ is not constant, one must design a new protocol for reconstructing the profile of $b(t)$ since the usual destructive nature of quantum measurements implies that continuous monitoring is not possible. For example, one could partition $I = [0, T]$ into n uniformly spaced intervals and measure $b(t)$ in each interval. However, this is often impractical since, in order for the discretized reconstruction of $b(t)$ to be accurate, n must be large. This

entails large overhead associated with readout and re-initialization of the sensor. Recently, Ref. [15] proposed an alternative method to estimating time-dependent parameters with quantum sensors. In this paper, we discuss how one can build on that result by merging compressive sensing (CS) techniques [16, 17] and quantum control methods to reproduce the profile of $b(t)$ with the potential for an exponential savings in the number of required measurements. From a more general perspective, compressive sensing techniques can be an ideal potential solution to the problem of costly measurement resources in quantum systems.

Compressive sensing (CS) is a relatively new sub-field of signal processing that can outperform traditional methods of *transform coding*, where the goal is to acquire, transform, and store signals as efficiently as possible. Suppose the signal F of interest is either naturally discrete or discretized into an element of \mathbb{R}^n . When the signal is sparse in some basis Ψ of \mathbb{R}^n , most traditional methods rely on measuring the signal with respect to a complete basis Φ , transforming the acquired data into the natural (sparse) basis Ψ , and finally compressing the signal in this basis by searching for and keeping only the largest coefficients. The end result is an encoded and compressed version of the acquired signal in its sparse basis that can be used for future communication protocols. There are various undesirable properties of this procedure. First, the number of measurements in the basis Φ is typically on the order of n , which can be extremely large. Second, transforming the signal from the acquisition basis Φ to its natural basis Ψ can be computationally intensive. Lastly, discarding the small coefficients involves locating the largest ones, which is a difficult search problem. Thus, it is clearly desirable to have methods that minimize the number of required measurements, maximize the information contained in the final representation of F in Ψ , and circumvent performing large transformations to move between representations.

CS theory shows that one can design both a “suitable” measurement matrix Φ and an efficient convex optimization algorithm so that only a small number of measurements $m \ll n$ are required for *exact* reconstruction of the signal F . Hence, the compression is performed at the measurement stage itself and all three of the problems listed above are solved; there is a significant reduction in the number of measurements and, since the convex optimization algorithm directly finds the sparse representation in an efficient manner, the reconstruction is exact and no large basis transformation is required. Finding suitable measurement matrices and reconstruction algorithms are active areas of research in compressive sensing and signal processing theory. CS techniques have already been applied in a wide array of fields that include, but is certainly not limited to, medical imaging [18], channel and communication theory [19], computational biology [20], geophysics [21], radar techniques [22], tomography [23], audio and acoustic processing [24], and computer graphics [25]. The wide applicability of CS tech-

niques is an indication of both its power and generality, and here we show many of these techniques are also amenable to quantum sensing. In the realm of quantum metrology, CS has been used for Hamiltonian identification in the case of static interactions [26], and more generally for quantum process tomography [23]. In contrast, here we introduce methods for dynamic parameter estimation

The paper is organized as follows. In Sec. II we provide a three-part review of the relevant CS results we will need throughout the presentation. The first part discusses compressive sensing from the viewpoint of the sparse and measurement bases being fixed and incoherent, which provides the foundation for understanding CS. The second part discusses how one can use randomness to create measurement bases that allow the reconstruction of *any* sparse signal. In the final part, we discuss the extent to which CS is robust against both approximate sparseness and noise.

We then move on to the main results of the paper. We first show in Sec. III how one can utilize both the discrete nature of quantum control sequences and the existence of the Walsh basis to reconstruct signals that are sparse in the time-domain. As an example application, we consider reconstructing magnetic fields produced by spike trains in neurons. We then generalize the discussion to *arbitrary* sparse bases in Sec. IV and show the true power of using CS in quantum parameter estimation. We show that for any deterministic dynamic parameter that is sparse in a known basis, one can implement randomized control sequences and utilize CS to reconstruct the magnetic field. Hence, since this procedure works for signals that are sparse in any basis, it can be thought of as a “universal” method for performing dynamic parameter estimation. We also show these protocols are robust in the cases of approximate sparsity and noisy signals. We provide various numerical results to quantify these results and make concluding remarks in Sec. V.

II. REVIEW OF COMPRESSIVE SENSING THEORY

Compressive sensing techniques [16, 17] allow for the reconstruction of signals using a much smaller number of non-adaptively chosen measurements than required by traditional methods. The success of CS is based on the fact that signals that are naturally sparse in some basis can be efficiently reconstructed using a small number of measurements *if* the measurement (sensing) basis is “*incoherent*” [27] with the sparse basis. Loosely speaking, incoherence generalizes the relationship between time and frequency, and makes precise the idea that the sensing basis is spread out in the sparse basis (and vice versa). As discussed in the introduction, many situations of interest, such as medical imaging [18] and communication theory [19] deal with sensing sparse signals. The main advantages of CS over traditional techniques derive from

the compression of the signal at the sensing stage and the efficient and exact convex reconstruction procedures. Let us now describe the general ideas and concepts of CS theory, highlighting those that will be important for the purpose of dynamic parameter estimation with quantum probes.

Suppose $F \in \mathbb{R}^n$ is the deterministic signal we want to reconstruct and F is S -sparse when written in the basis $\Psi := \{\psi_j\}_{j=1}^n$,

$$F = \sum_{j=1}^n \langle F, \psi_j \rangle \psi_j = \sum_{j=1}^n F_j^\Psi \psi_j, \quad (2.1)$$

that is, only $S \leq n$ of the coefficients F_j^Ψ are non-zero. For simplicity, let us assume that Ψ is ordered so that the magnitude of the coefficients F_j^Ψ monotonically decrease as j increases. In most physically realistic scenarios, which include quantum parameter estimation, measurements of F are modeled as linear functionals on \mathbb{R}^n . By the Riesz representation theorem [28], each measurement can be associated to a unique element of \mathbb{R}^n . Suppose we have access to a set of n orthonormal measurements represented by the orthonormal basis of \mathbb{R}^n , $\Phi := \{\phi_k\}_{k=1}^n$. Since we can represent F as

$$F = \sum_{j=1}^n \langle F, \phi_j \rangle \phi_j = \sum_{j=1}^n F_j^\Phi \phi_j, \quad (2.2)$$

the output of a measurement ϕ_k is the k 'th coefficient, $\langle F, \phi_k \rangle$, of F with respect to this basis. One of the central questions compressive sensing attempts to answer is, under the assumption of S -sparsity, do there exist

1. conditions on the pair (Φ, Ψ) and
2. efficient reconstruction algorithms

that allow one to reconstruct f using a small number of measurements from Φ ? Ref. [16] has shown that if (Φ, Ψ) is an incoherent measurement basis pair then one can use an l_1 -minimization algorithm to efficiently reconstruct F with a small number of measurements. Let us explicitly define these concepts and present a set of well-known compressive sensing theorems that we have ordered in terms of increasing complexity.

A. Compressive Sensing from Incoherent Bases

The initial formulation of compressive sensing [16, 17] required the natural (Ψ) and measurement (Φ) bases to be incoherent, which loosely speaking means that these bases are as far off-axis from each other as possible. Rigorously, coherence is defined as follows [27].

Definition 1. Coherence

If (Φ, Ψ) is a basis pair then the coherence between Φ and Ψ is defined to be

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{1 \leq i, j \leq n} |\langle \phi_i, \psi_j \rangle|. \quad (2.3)$$

Thus, the coherence is a measure of the largest possible overlap between elements of Φ and Ψ . If Φ and Ψ are orthonormal bases then

$$\mu(\Phi, \Psi) \in [1, \sqrt{n}]. \quad (2.4)$$

When $\mu(\Phi, \Psi)$ is close to 1, Φ and Ψ are said to be incoherent and when the coherence is equal to 1, the pair is called maximally incoherent. CS techniques provide significant advantages when Φ and Ψ are incoherent. Suppose we select m measurements uniformly at random from Φ and we denote the set of chosen indices by $M \subset \{1, \dots, n\}$ so that $|M| = m$. The first CS theorem [29] gives a bound on the probability that the solution to the convex optimization problem,

Convex Optimization Problem (COP 1)

argmin $\{\|x\|_1 : x \in \mathbb{R}^n\}$ subject to : $\forall j \in M,$

$$F_j^\Phi = \left\langle \phi_j, \sum_{k=1}^n x_k \psi_k \right\rangle, \quad (2.5)$$

is equal to the vector F^Ψ , where ‘‘argmin’’ refers to finding the vector $x \in \mathbb{R}^n$ that minimizes the 1-norm $\|x\|_1$.

Theorem 1. Let Φ be the sensing basis and $F \in \mathbb{R}^n$ be S -sparse in its natural basis Ψ . Then, if $\delta > 0$, and

$$m \geq C \mu(\Phi, \Psi)^2 S \log\left(\frac{n}{\delta}\right) \quad (2.6)$$

for a fixed constant C , we have that the solution to COP 1 is equal to F^Ψ with probability no less than $1 - \delta$.

Note that the constant C in Eq. (2.6) is independent of n and is typically not very large. A general rule of thumb is the ‘‘factor of 4’’ rule which says that approximately $m = 4S$ measurements usually suffice when the bases are maximally incoherent. To summarize, under the conditions of Theorem 1, the vector of coefficients in the Ψ basis that minimizes the 1-norm *and* is consistent with the m measurement results will be equal to $(F_1^\Psi, \dots, F_n^\Psi)$ with probability no less than $1 - \delta$.

Clearly, there are two important factors in Eq. (2.6), the sparsity S and the level of coherence μ between the bases Ψ and Φ . When the bases are maximally coherent, there is no improvement over estimating all n coefficients. However, when the bases are maximally incoherent, one needs to only perform $O(S \log(\frac{n}{\delta}))$ measurements, which is a significant reduction in measurements (especially for large n). There are various examples of incoherent pairs, for instance, it is straightforward to verify that the pairs

1. standard basis/Fourier basis,
2. standard basis/Walsh basis,
3. noiselets/wavelets [29],

are incoherent, with the first two being maximally incoherent. The second pair will be especially useful for using CS to perform magnetometry using quantum systems. We now analyze how to relax the incoherence condition by using random matrices.

B. Random Measurement Matrices

As we have seen, CS techniques can provide significant advantages when the measurement and sparse bases are incoherent. However, for a given sparse basis, the requirement of incoherence places restrictions on the types of measurements one can perform. To overcome this, a large body of theory has been developed regarding how to construct measurement matrices that still afford the advantages of CS. To generalize the discussion, we can see that the convex optimization problem given in Theorem 1 can be put into the following generic form

Convex Optimization Problem (COP 2)

$\operatorname{argmin}\{\|\tilde{x}\|_1 : \tilde{x} \in \mathbb{R}^n\}$ subject to $y = A\tilde{x}$.

This can be seen by taking $A = R\Phi^\dagger\Psi$, where R is an $m \times n$ matrix that picks out m rows from $\Phi^\dagger\Psi$. Focusing on this form of the convex optimization problem, we first discuss conditions on A which ensure exact recovery of sparse signals before describing how to construct such A .

1. Restricted Isometry Property

The restricted isometry property and constant are defined as follows [30].

Definition 2. Restricted Isometry Property (RIP)

We say that a matrix A satisfies the restricted isometry property (RIP) of order (S, δ) if $\delta \in (0, 1)$ is such that for all S -sparse vectors $x \in \mathbb{R}^n$,

$$(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2. \quad (2.7)$$

Note that this is equivalent to:

1. The spectral radius of $A^T A$, denoted $\sigma(A^T A)$, lying in the range $(1 - \delta, 1 + \delta)$,

$$1 - \delta \leq \sigma(A^T A) \leq 1 + \delta, \quad (2.8)$$

2. and also

$$\sqrt{1 - \delta} \leq \|A\|_{2,S} \leq \sqrt{1 + \delta}, \quad (2.9)$$

where we have defined the matrix 2-norm of sparsity level S for A ,

$$\|A\|_{2,S} = \max_{S\text{-sparse } x : \|x\|_2=1} \|Ax\|_2. \quad (2.10)$$

Definition 3. Restricted Isometry Constant (RIC)

The infimum over all δ , denoted δ_S , for which A satisfies the RIP at sparsity level S is called the restricted isometry constant (RIC) of A at sparsity level S . We also say that A satisfies the RIP of order S if there is some $\delta \in (0, 1)$ for which A satisfies the RIP of order (S, δ) .

The RIP is fundamental in compressive sensing. If A satisfies the RIP with $\delta_S \ll 1$ then A acts like an isometry on S -sparse vectors, that is, it preserves the Euclidean 2-norm of S -sparse vectors. Hence, S -sparse vectors are *guaranteed* to not be in the kernel of A and, if A constitutes a measurement matrix, one might hope x can be reconstructed via sampling from A . As shown in Theorem 2 below, which was first proved in [31], this is indeed the case.

2. Reconstruction for Sparse Signals

Theorem 2. Suppose $x \in \mathbb{R}^n$ satisfies the following conditions

1. x is S -sparse,
2. the underdetermined $m \times n$ matrix A satisfies the RIP of order $2S$ with $\delta_{2S} < 0.4652$,
3. $y = Ax$.

Then there is a unique solution x^* of COP 2, $x^* = x$.

The constant 0.4652 is not known to be optimal. It is important to note that Theorem 2 is not probabilistic. If A satisfies the RIP of order $2S$, and the associated RIC δ_{2S} is small enough, then exact recovery will always occur. Thus, recalling that Theorem 1 was probabilistic, it is clear that even if the basis pair (Φ, Ψ) is incoherent, the matrix $R\Phi^\dagger\Psi$ is not guaranteed to satisfy the RIP property for m given by Eq. (2.6). Equivalently, choosing this many rows at random from Φ is not guaranteed to produce a matrix $R\Phi^\dagger\Psi$ which satisfies the RIP of order $2S$ with $\delta_{2S} < 0.4652$. However, if we allow for $m \sim O(S(\log(n))^4)$, then $R\Phi^\dagger\Psi$ does satisfy the RIP with probability 1.

With Theorem 2 in hand, the remaining question is how to create $m \times n$ matrices A that satisfy the RIP with small δ_{2S} . Deterministically constructing such A is a hard problem, however various random matrix models satisfy this condition with high probability if m is chosen large enough. For a detailed discussion, see Ref. [32]. The key result is the following theorem.

Theorem 3. Suppose

1. S , n , and $\delta \in (0, 1)$ are fixed and
2. the mn entries of A are chosen uniformly at random from a probability distribution to form a matrix-valued random variable $A(\omega)$ which, for all $\epsilon \in (0, 1)$, satisfies the concentration inequality

$$\mathbb{P} \left[\left| \|A(\omega)x\|_2^2 - \|x\|_2^2 \right| \geq \epsilon \|x\|_2^2 \right] \leq 2e^{-nC_0(\epsilon)}. \quad (2.11)$$

Note that Eq. (2.11) means that the probability of a randomly chosen vector x satisfying $\left| \|A(\omega)x\|_2^2 - \|x\|_2^2 \right| \geq \epsilon \|x\|_2^2$ is less than or equal to $2e^{-nC_0(\epsilon)}$, where $C_0(\epsilon) > 0$ is a constant that depends only on ϵ .

Then there exist constants $C_1, C_2 > 0$ that depend only on δ such that, if

$$m \geq C_1 S \log\left(\frac{n}{S}\right), \quad (2.12)$$

then, with probability no less than $1 - 2e^{-C_2 n}$, A satisfies the RIP of order (S, δ) .

We note some important points of Theorem 3. First, in the spirit of Sec. II A, let us fix some *arbitrary* basis Ψ and choose a random $m \times n$ measurement matrix $R\Phi^\dagger$ according to a probability distribution that satisfies Eq. (2.11). Then Theorem 3 *also holds* for the matrix $R\Phi^\dagger\Psi$ even though only $R\Phi^\dagger$ was chosen randomly. In this sense, the random matrix models above form “universal encoders” because, if the RIP property holds for $R\Phi^\dagger$, then it also holds for $R\Phi^\dagger\Psi$ *independently* of Ψ . So, as long as the signal is sparse in some basis that is known a priori then, with high probability, we can recover the signal using a random matrix model and convex optimization techniques. Second, this theorem is equivalent to Theorem 5.2 of Ref. [32], where they fixed m, n , and δ and deduced upper bounds on the sparsity (therefore the constant C_1 above is the inverse of C_1 in Ref. [32]).

Some common examples of probability distributions that lead to the concentration inequality in Eq. (2.11) are

1. Sampling the n columns uniformly at random from \mathbb{S}^{m-1} ,
2. Sampling each entry from a normal distribution with mean 0 and variance $\frac{1}{m}$,
3. Sampling the m rows by random m -dimensional projections P in \mathbb{R}^n (and normalizing by $\sqrt{\frac{n}{m}}$),
4. Sampling each entry from the symmetric Bernoulli distribution $\mathbb{P}\left(A_{i,j} = \pm \frac{1}{\sqrt{m}}\right) = \frac{1}{2}$,
5. Sampling each entry from from the set $\left\{-\sqrt{\frac{3}{m}}, 0, \sqrt{\frac{3}{m}}\right\}$ according to the probability distribution $\left\{\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right\}$.

Example 4 will be of particular importance for our main results. We now analyze how to relax the condition of sparsity to compressibility and also how to take noise effects into account.

C. Compressive Sensing in the Presence of Noise

The last part of this brief outline of CS techniques shows that they are robust to both small deviations from sparsity as well as noise in the signal. First, for any vector x , let x_S represent the vector resulting from only keeping the S entries of largest magnitude. Thus, x_S is the best S -sparse approximation to x and x is called “ S -compressible” if $\|x - x_S\|_1$ is small. In many applications,

the signal of interest x only satisfies the property of compressibility, since many of its coefficients can be small in magnitude but are not exactly equal to 0. In addition, real signals typically are prone to noise effects. Suppose the signal and measurement processes we are interested in are affected by noise that introduces itself as an error ϵ in the measurement vector y . We have the following convex optimization problem which includes the noise term of strength ϵ :

Convex Optimization Problem (COP 3)

$\text{argmin}\{\|\tilde{x}\|_1 : \tilde{x} \in \mathbb{R}^n\}$ subject to $\|y - A\tilde{x}\|_2 \leq \epsilon$.

We now have the following CS theorem [31], which is the most general form that we will be considering.

Theorem 4. If the matrix A satisfies the RIP of order $2S$, and $\delta_{2S} < 0.4652$, then the solution x^* of the COP 3 satisfies

$$\|x^* - x\|_2 \leq \frac{C_3 \|x - x_S\|_1}{\sqrt{S}} + \epsilon, \quad (2.13)$$

where x_S is the S -compressed version of x .

Theorem 4 is deterministic, holds for *any* x , and says that the recovery is robust to noise and is just as good as if one were to only keep the S largest entries of x (the compressed version of x). If the signal is exactly S -sparse then $x = x_S$ and the recovery is exact, up to the noise term ϵ . We now discuss how to apply CS theory to dynamic parameter estimation and quantum magnetometry.

III. QUANTUM DYNAMIC PARAMETER ESTIMATION WITH COMPRESSIVE SENSING

We now present the main results of the paper, combining ideas from CS presented above with coherent control of quantum sensors. For concreteness, we adopt notation suitable for quantum magnetometry. Thus, we assume the deterministic function of interest is a magnetic field $b(t)$ that we want to reconstruct on the time interval $[0, T]$. We partition $I = [0, T]$ into n uniformly spaced intervals with endpoints $t_j = \frac{jT}{n}$ for $j \in \{0, \dots, n\}$ and discretize $[0, T]$ into the n mid-points $\{s_j\}_{j=0}^{n-1}$ of these intervals,

$$s_j = \frac{(2j+1)T}{2n}. \quad (3.1)$$

The discretization of $b(t)$ to a vector $B \in \mathbb{R}^n$ is defined by $B_j = b(s_j)$ for each $j \in \{0, \dots, n-1\}$. In principle, each B_j can be approximately estimated by performing a Ramsey protocol in each time interval and assuming the magnetic field is constant over the partition width $\delta = \frac{T}{n}$. The result of the Ramsey protocol is to acquire $\frac{1}{\delta} \int_{t_j}^{t_{j+1}} b(t) dt$. The key idea is that, instead of this

naive method that requires n measurements, we can apply control sequences during the total time T to modulate the evolution of $b(t)$ before making a measurement at T . While we still need to repeat the measurement for different control sequences, this method is amenable to CS techniques, with the potential for exponential savings in the number of measurements needed for an accurate reconstruction of B .

Using coherent control to reconstruct simple sinusoidal fields was performed in Ref. [33] by using the Hahn spin-echo [34] and its extensions, such as the periodic dynamical decoupling (PDD) and Carr-Purcell-Meiboom-Gill (CPMG) sequences [35, 36]. Recently, control sequences based on the Walsh basis [37] have been proposed to reconstruct fields of a completely general form in an accurate manner [15]. The main point in all of these methods is to control the system during the time $[0, T]$ by performing π rotations at pre-determined times $t_j \in [0, T]$. Let us briefly describe the control and measurement processes.

At each t_j , a π -pulse is applied according to some pre-defined algorithm encoded as a length- n bit string u . The occurrence of a 1 (0) in u indicates that a π -pulse should (should not) be applied. The evolution of the system is then given by

$$U(T) = \prod_{j=n-1}^0 \left[\left(e^{-i \int_{t_j}^{t_{j+1}} b(t) dt} \sigma_z \right) \pi^{u(j)} \right] \quad (3.2)$$

$$= e^{-i \int_0^T \kappa_u(t) b(t) dt} \sigma_z = e^{-iT \langle \kappa_u(t), b(t) \rangle \sigma_z},$$

where $\kappa_u(t)$ is the piecewise constant function taking values ± 1 on each (t_j, t_{j+1}) and a switch $1 \leftrightarrow -1$ occurs at t_j if and only if a π -pulse is implemented at t_j (we assume without loss of generality that $\kappa_u(t) = 1$ for $t < 0$). The value of κ_u on an interval $[t_j, t_{j+1})$, $j \in \{0, \dots, n-1\}$, is determined by the parity of the rank of the truncated sequence $u^j = (u(0), \dots, u(j))$,

$$\kappa_u [(t_j, t_{j+1})] = (-1)^{\text{parity}(\sum_{i=0}^j u(t_i))}. \quad (3.3)$$

Hence, performing a π -modulated experiment produces a phase evolution given by

$$\phi_u(T) = T \langle \kappa_u(t), b(t) \rangle = \int_0^T \kappa_u(t) b(t) dt. \quad (3.4)$$

Performing a measurement in the computational basis $\{|0\rangle, |1\rangle\}$ gives the following probability of obtaining outcome “0”

$$p_0 = \frac{1 + \sin(T \langle \kappa_u(t), b(t) \rangle)}{2}. \quad (3.5)$$

Hence, for each u , one can solve for $\langle \kappa_u(t), b(t) \rangle$ by estimating p_0 and solving for $\langle \kappa_u(t), b(t) \rangle$.

If the set of all κ_u form an orthonormal basis for the set of square-integrable functions on $[0, T]$, denoted $L^2[0, T]$,

then we can write

$$b(t) = \sum_u \langle \kappa_u(t), b(t) \rangle \kappa_u(t). \quad (3.6)$$

We know that the κ_u are piecewise continuous functions and take the values ± 1 . An example of a piecewise continuous orthonormal basis of $L^2[0, T]$ is the Walsh basis [37], which we denote $\{w_m\}_{m=0}^\infty$ (see Appendix A for details). Each m corresponds to a different control sequence and one can reconstruct $b(t)$ according to its Walsh decomposition

$$b(t) = \sum_{m=0}^\infty \langle w_m(t), b(t) \rangle w_m(t). \quad (3.7)$$

The Walsh basis will be useful for our first set of results in Sec. III where we use incoherent measurement bases to reconstruct $b(t)$.

On the other hand, the set of all κ_u need not be a basis. They can be chosen to be random functions, which are also useful from the point of view of random matrices and the RIP. We use randomness and the RIP for our second set of results in Sec. IV, where $b(t)$ can be sparse in *any* basis.

A. Reconstructing Temporally Sparse Magnetic Fields Using Incoherent Measurement Bases

We first focus on sparse time domain signals, which are important since they can model real physical phenomena such as action potential pulse trains in neural networks [38]. In this case, the parameter $b(t)$ has an S -sparse discretization B when written in the standard spike basis of \mathbb{R}^n . The spike basis is just the standard basis that consists of vectors with a “1” in exactly one entry and “0” elsewhere. To keep notation consistent, we denote the spike basis by Ψ . From Theorem 1, if we want to reconstruct B using a measurement basis Φ , then we need Ψ and Φ to be incoherent. It is straightforward to show that the discrete orthonormal Walsh basis $\{W_j\}_{j=0}^\infty$ (see Appendix A) is maximally incoherent with the spike basis. Thus, let us suppose that the measurement basis Φ is the discrete Walsh basis in sequency ordering [37]. The Walsh basis is particularly useful because it can be easily implemented experimentally [15] and it has the added advantage of refocusing dephasing noise effects [39].

In order to estimate the k 'th coefficient B_k^Φ one needs to apply control sequences that match the k 'th Walsh function. More precisely, for $j \in \{1, \dots, n-1\}$, if a switch $+1 \leftrightarrow -1$ occurs in the k 'th Walsh function at time t_j then a π -pulse is applied at t_j . We further assume that available resources limit us to implementing Walsh sequences of order N so that $n = 2^N$. Practically, the resources that determine the largest possible n one could use depends on the situation. We are therefore constrained to the information contained in the discretization of $b(t)$ to the vector $B \in \mathbb{R}^{2^N}$. The discretized

magnetic field B is given in the Ψ and Φ bases by

$$\begin{aligned} B &= \sum_{k=0}^{2^N-1} B_k^\Psi \psi_k = \sum_{k=0}^{2^N-1} b(s_k) \psi_k = \sum_{k=0}^{2^N-1} B_k^\Phi \Phi_k \quad (3.8) \\ &= \sum_{k=0}^{2^N-1} \langle B, \Phi_k \rangle \Phi_k = \sum_{k=0}^{2^N-1} \left[\sum_{j=0}^{n-1} b(s_j) \frac{w_k(s_j)}{\sqrt{n}} \right] \Phi_k. \end{aligned}$$

From Theorem 1 we expect that, since B is assumed to be sparse in the standard basis, very few measurements of Walsh coefficients (much smaller than 2^N) are sufficient to reconstruct b with high probability. Let us rephrase Theorem 1 in the notation introduced here.

Theorem 5. Let $B = (b(s_0), \dots, b(s_{n-1})) \in \mathbb{R}^n$ be the N 'th order discretization $b(t) \in [0, T]$ ($n = 2^N$) and suppose B is S -sparse in the spike basis Ψ . Let us select m measurements (discrete Walsh vectors) uniformly at random from Φ and denote the set of chosen indices by M ($|M| = m$). Then, if $\delta > 0$ and

$$m \geq CS \log \left(\frac{n}{\delta} \right), \quad (3.9)$$

the solution to the following convex optimization problem is equal to B^Ψ with probability no less than $1 - \delta$,

Convex Optimization Problem

argmin $\{\|x\|_1 : x \in \mathbb{R}^n\}$ subject to : $\forall j \in M$,

$$B_j^\Phi = \left\langle \phi_j, \sum_{j=0}^{n-1} x_j \psi_j \right\rangle. \quad (3.10)$$

Note that Theorem 5 also holds in the case where there is noise in the measurement results B_j^Φ .

An important point that needs to be clarified is that real measurements performed using a spin system such as the NV center produce the coefficients of $b(t)$ with respect to the *continuous* Walsh basis $\{w_k(t)\}_{k=0}^\infty$. These coefficients, which we denote by \hat{b}_k , are not equal to the coefficients B_k^Φ . Thus, to be completely rigorous, we need to understand how to compare \hat{b}_k and B_k^Φ . It will be more useful to multiply \hat{b}_k by \sqrt{n} and compare the distance between $\sqrt{n}\hat{b}_k$ and B_k^Φ . If we measure m coefficients using the spin system and obtain a vector $y = (\hat{b}_{\alpha_1}, \dots, \hat{b}_{\alpha_m})$, we can then bound the distance between y and $z = Ax$ where

$$z_{\alpha_k} = B_{\alpha_k}^\Phi. \quad (3.11)$$

Then, if we obtain

$$\|y - z\|_2 \leq \epsilon, \quad (3.12)$$

for some $\epsilon > 0$, we can treat the vector y obtained from the real sensing protocol as a *noisy version* of z . Since CS is robust to noise, the reconstruction will still be successful if ϵ is small. So, let us bound $\|y - z\|_2$ by first

bounding the distance between $\sqrt{n}\hat{b}_k$ and B_k^Φ for some k . We have

$$\hat{b}_k = \frac{1}{T} \int_0^T b(t) w_k(t) dt = \frac{1}{T} \int_0^T g_k(t) dt, \quad (3.13)$$

where we let w_k denote the scaled version of the k 'th Walsh function to $[0, T]$ and we define the function g_k by $g_k(t) = b(t) w_k(t)$. Now

$$B_k^\Phi = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} w_k(s_j) b(s_j) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} g_k(s_j), \quad (3.14)$$

and so

$$\begin{aligned} \left| \sqrt{n}\hat{b}_k - B_k^\Phi \right| &= \left| \frac{\sqrt{n}}{T} \int_0^T g_k(t) dt - \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} g_k(s_j) \right| \\ &= \frac{\sqrt{n}}{T} \left| \int_0^T g_k(t) dt - \frac{T}{n} \sum_{j=0}^{n-1} g_k(s_j) \right|. \end{aligned} \quad (3.15)$$

By the midpoint error formula for approximating integrals by Riemann sums [40] we have

$$\begin{aligned} \left| \sqrt{n}\hat{b}_k - B_k^\Phi \right| &\leq \frac{\sqrt{n}}{T} \max_{t \in [0, T]} |b''(t)| \frac{T}{24} \frac{T^2}{n^2} \\ &= \frac{T^2}{24n^{\frac{3}{2}}} \max_{t \in [0, T]} |b''(t)|, \end{aligned} \quad (3.16)$$

and so, for y and z defined above,

$$\|y - z\|_2 \leq \frac{\sqrt{m}T^2}{24n^{\frac{3}{2}}} \max_{t \in [0, T]} |b''(t)|. \quad (3.17)$$

Thus, we can set

$$\epsilon = \frac{\sqrt{m}T^2}{24n^{\frac{3}{2}}} \max_{t \in [0, T]} |b''(t)|, \quad (3.18)$$

which is small since CS requires $m \sim O(S \log(n))$ measurements, where S is the sparsity level. Thus, we have

$$\epsilon \sim \frac{T^2}{24} \max_{t \in [0, T]} |b''(t)| \sqrt{\frac{S \log(n)}{n^3}}, \quad (3.19)$$

which converges to 0 quickly in n . Hence, using the coefficients \hat{b}_k obtained from the physical sensing protocol will still provide robust reconstructions of the magnetic field. We now use these results to discuss applications to neural magnetometry, with the goal being to reconstruct the magnetic field profile of firing action potentials in neurons.

B. Numerical Simulations

Here, we present a numerical analysis of using CS to reconstruct time-sparse magnetic fields. Since our only

constraint is that the signal is sparse in the standard basis, there is clearly a wide range of possible models we can choose from. To place the model within the context of a physically relevant scenario, we assume the magnetic field is a series of two-sided spikes as is the case when a series of action potentials is produced by a firing neuron [41–44]. There is large variation in the physical parameters describing such neuronal activity. We chose a set of parameters that both aid in highlighting the potential advantages of CS and are close to current specifications for a sensing system such as the NV center [5, 6, 15]. We assumed a total acquisition time of $T = 1$ ms and defined an “event” to be a single action potential, which we assumed to last $10\mu\text{s}$. As well, we assumed that five events occur in the 1 ms time-frame and that control pulse times last approximately 10 ns. We chose these parameters, which lie at the current extremes of each system in order to have many different events occurring in $[0, T]$ (see figures below). Parameters closer to current experimental capabilities (e.g. pulse times of 40 ns and action potentials of $100\mu\text{s}$) would have resulted in a smaller number of events and thus less meaningful numerical comparisons.

Each of the two spikes in an action potential was assumed to have maximum magnitude of 1 nT and last $\Delta = 5\mu\text{s}$. If τ_P denotes the pulse time then, in practice, one only has to choose a reconstruction order N such that the resolution $\frac{1}{n} = 2^{-N}$ of the partition of size 2^N on the time interval $[0, T]$ satisfies the following two-sided inequality condition

$$\tau_P < 2^{-N}T < \Delta. \quad (3.20)$$

Hence, we can take $N = 10$ (so $n = 2^{10}$) as a suitable reconstruction order. More precisely, $2^{-10}\text{ms} \sim 1\mu\text{s}$ is a fine enough resolution to capture events of length Δ , yet coarse enough so that 10ns pulses can be approximated as δ -pulses. For $n = 1024$, the average and maximum number of pulses one has to apply in a sequence are 512 and 1024 respectively. Since CS techniques are robust to noise, we expect the reconstructions to be relatively robust against pulse errors and imperfections.

We implemented a CS reconstruction using random, non-adaptive, measurements in the Walsh basis. We emphasize that the Walsh and spike bases are maximally incoherent and that the Walsh measurement basis is easily implemented in practice. Again, the number of events in the spike train was chosen to be 5 so that there are 10 total magnetic field pulses (5 of each polarity). For simplicity, the times of each active potential event were chosen uniformly at random in $[0, T]$. We therefore have a sparsity level of $S = 50$ and so the number of measurements m we should require is

$$\begin{aligned} m &\sim O(S \log(n)) \\ &\sim O(500). \end{aligned} \quad (3.21)$$

As mentioned in Sec. II there is an empirical rule that says in many cases around $4S$ or $5S$ measurements typically suffice to reconstruct the signal. Thus, we might

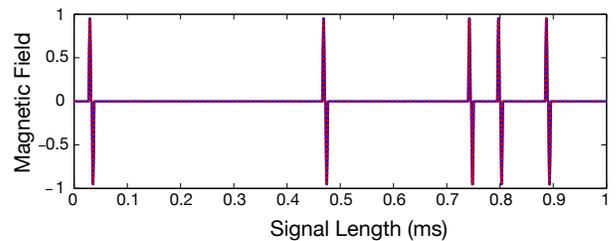


FIG. 1: Simulated (blue solid) and Successful CS Reconstructed (red dotted) Magnetic Fields (5 events with $m=200$ and $\text{MSQE}=7.6109\text{e-}12$)

expect around 200 or 250 measurements are adequate to reconstruct the signal with high probability.

The protocol was numerically implemented for various values of m up to 500. Once m reached 200, as expected by the above empirical rule, the probability of successful exact reconstruction, p_{suc} , began to quickly converge to 1 which verifies that the empirical rule is satisfied here. At $m = 250$, the probability was essentially equal to 1. For each $m \in \{200, 210, 220, \dots, 290, 300\}$ we recorded p_{suc} from a sample of 1000 trials. Since the reconstruction is either close to exact or has observable error, and the majority of the actual signal in the time-domain is equal to 0 (which implies the total variation of the signal is small), we set a stringent threshold for determining successful reconstruction; if the mean-squared error (MSQE) between the simulated (B_{sim}) and reconstructed (B_{rec}) magnetic fields was less than 10^{-9} (nT)²s then the reconstruction was counted as successful. The results are contained in Table I.

The main message is that, if the relevant experimental parameters are taken into account, one obtains a range of possible values for n which defines the required resolution of the grid on $[0, T]$. This provides an upper bound on the number of operations in a Walsh sequence that have to be performed. If S is the expected sparsity of the signal with respect to the chosen value for n then taking $m \sim O(S \log(n))$ measurements implies the probability of successful reconstruction, p_{suc} , will be high and converges to 1 quickly as m grows larger. We plotted examples of successful and unsuccessful reconstructions in Fig. 1 and 2 respectively. Typically, the CS reconstruction either works well ($\text{MSQE} < 10^{-9}$) or clearly “fails” ($\text{MSQE} \sim 0.01$).

1. Accounting for Decoherence

From Theorem 4, we know that CS techniques are robust in the presence of noise. We model the noise and evolution of the system as follows. We suppose that the Hamiltonian of the system is given by

$$\hat{H}(t) = [b(t) + \beta(t)] \sigma_z, \quad (3.22)$$

m	200	210	220	230	240	250	260	270	280	290	300
p_{suc}	0.870	0.950	0.974	0.991	0.996	0.999	1.000	1.000	1.000	1.000	1.000

TABLE I: Probability of successful CS reconstruction, p_{suc} , for different values of $m \ll n = 1024$.

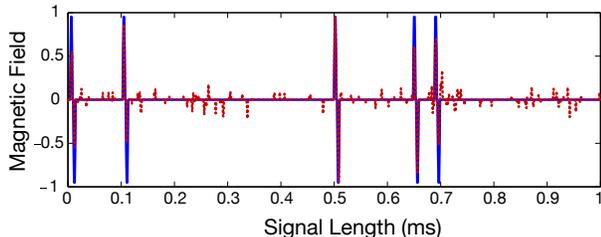


FIG. 2: Simulated (blue solid) and Unsuccessful CS Reconstructed (red dotted) Magnetic Fields (5 events with $m=200$ and $\text{MSQE}=0.0051747$).

where $\beta(t)$ is a zero-mean stationary stochastic process. By the Wiener-Khinchine theorem, the spectral density function of $\beta(t)$, $S_\beta(\omega)$, is equal to the Fourier transform of the auto-correlation function of $\beta(t)$ and thus the decay in coherence of the quantum system is given by $v = e^{-\chi(t)}$, where

$$\chi(T) = \int_0^\infty \frac{S_\beta(\omega)}{\omega^2} F(\omega T), \quad (3.23)$$

and $F(\omega T)$ is the filter function for the process [45]. It is important to note that one can experimentally reconstruct $S_\beta(\omega)$ by applying pulse sequences of varying length and inter-pulse spacings that match particular frequencies [46–48].

Applying control sequences that consist of π -pulses during $[0, T]$ modulates $\chi(T)$ by modifying the form of $F(\omega T)$ in Eq. (3.23) [45]. In most experimentally relevant scenarios, low frequency noise gives the most significant contribution to $S_\beta(\omega)$. Hence, typically, the goal is to design control sequences that are good high-pass filters. When the control sequence is derived from the j 'th Walsh function, we have

$$\chi_j(T) = \int_0^\infty \frac{S_\beta(\omega)}{\omega^2} F_j(\omega T), \quad (3.24)$$

where $F_j(\omega T)$ is the filter function associated with the j 'th Walsh function. The low frequency behavior of each F_j has been analyzed in detail [39, 49]; in general, if the noise spectrum is known, one can obtain an estimate of each $\chi_j(T)$, and thus each v_j .

The signal S_j acquired from a sensing protocol with the j 'th Walsh control sequence is given by

$$S_j(T) = \frac{1}{2} [1 + v_j(T) \sin(z_j(T))], \quad (3.25)$$

where

$$z_j(T) = \frac{1}{T} \int_0^T w_j(t) b(t) dt. \quad (3.26)$$

We note that for zero-mean noise we have

$$\left\langle \int_0^T w_j(t) \beta(t) dt \right\rangle = 0. \quad (3.27)$$

Now, for N_j measurement repetitions, we have that the sensitivity in estimating z_j , denoted Δz_j , is given by [49]

$$\Delta z_j = \frac{1}{\sqrt{N_j T v_j}}. \quad (3.28)$$

Thus, fluctuations in obtaining the measurement results z_j are on the order of $\frac{1}{\sqrt{N_j T v_j}}$ and so the 2-norm distance between the ideal measurement outcome vector z and actual measurement vector y is

$$\|z - y\|_2 \sim \sqrt{\sum_{j=1}^m \left[\frac{1}{\sqrt{N_j T v_j}} \right]^2} = \sqrt{\sum_{j=1}^m \frac{1}{N_j T^2 v_j^2}}. \quad (3.29)$$

Since CS reconstructions are robust up to $\|z - y\|_2$, we have that a realistic CS reconstruction in a system such as the NV center is robust up to $\sqrt{\sum_{j=1}^m \frac{1}{N_j T^2 v_j^2}}$.

IV. RECONSTRUCTING GENERAL SPARSE MAGNETIC FIELDS

While time-sparse signals are important in various applications, they are far from being common. Ideally, we would like to broaden the class of signals that we can reconstruct with CS techniques and quantum sensors. In this section, we present a method for reconstructing signals that are sparse in *any* known basis. This generality of the signal is significant since most real signals can be approximated as sparse in some basis. We use the results of Sec. II B to show that performing *random* control sequences creates random measurement matrices that satisfy the RIP (see Def. 2) with high probability. Therefore, by Theorem 2, a small number of measurements suffice for exact reconstruction of the signal. Again, we emphasize the key point that the signal can be sparse in *any* basis of \mathbb{R}^n .

A. Applying Random Control Sequences

Suppose $b(t)$ is a deterministic magnetic field on $I = [0, T]$ and we partition I into n uniformly spaced intervals with grid-points $t_j = \frac{jT}{n}$ for $j \in \{0, \dots, n\}$. Previously, we used the discrete Walsh basis $\{w_k\}_{k=0}^{2^N-1}$ as our measurement basis Φ . π -pulses were applied at each t_j according to whether a switch between $+1$ and -1 occurred at t_j . Now, for each $j \in \{0, \dots, n-1\}$, we choose to either apply or not apply a π -pulse at t_j according to the symmetric Bernoulli distribution

$$\mathbb{P}[\pi \text{ applied at } t_j] = \frac{1}{2}. \quad (4.1)$$

The result of this sampling defines the n -bit string u , where the occurrence of a 0 indicates a π -pulse is not applied, and a 1 indicates a π -pulse is applied. Following Eq. (3.2)-(3.4), the evolution of the system is

$$U(T) = e^{-iT \langle \kappa_u(t), b(t) \rangle \sigma_z}. \quad (4.2)$$

As before, let us discretize $[0, T]$ into n points $\{s_j\}_{j=0}^{n-1}$ where

$$s_j = \frac{(2j+1)T}{2n}. \quad (4.3)$$

Define $\tilde{\kappa}_u$ to be the random element of $\{-1, 1\}^n$ with entries given by $\kappa_u(s_j)$ for each $j \in \{0, \dots, n-1\}$. As well, define $B \in \mathbb{R}^n$ to be the discretization of $b(t)$ at each s_j , $B_j = b(s_j)$.

Suppose there is an orthonormal basis Ψ of \mathbb{R}^n for which B has an approximate S -sparse representation. To maintain notational consistency, we use Ψ to represent the $n \times n$ matrix whose columns are the basis elements $\psi_j \in \mathbb{R}^n$. Let x denote the coordinate representation of B in Ψ ,

$$B = \sum_{j=1}^n x_j \psi_j. \quad (4.4)$$

We emphasize that Ψ is an *arbitrary* but known basis. Assume we have chosen m random measurement vectors κ_{u_i} , $i = 1, \dots, m$. and let us define the matrix \mathcal{G} whose entries are given by

$$\mathcal{G}_{i,j} = \frac{\kappa_{u_i}(s_j)}{\sqrt{m}} = \frac{\tilde{\kappa}_{u_i}(j)}{\sqrt{m}}, \quad (4.5)$$

where $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Also, let us define the $m \times n$ matrix

$$A = \mathcal{G}\Psi. \quad (4.6)$$

Since the entries of \mathcal{G} were chosen according to a probability distribution that satisfies Eq. (2.11), from Theorem 3, there exist constants $C_1, C_2 > 0$ which depend only on δ such that, with probability no less than

$$1 - 2e^{-C_2 n}, \quad (4.7)$$

\mathcal{G} , and thus A , satisfies the RIP of order $(2S, \delta)$ when

$$m \geq 2C_1 S \log\left(\frac{n}{2S}\right). \quad (4.8)$$

By Theorem 4 this implies that, if we let y be the noisy version of the ideal measurement Ax and $\|y - Ax\|_2 \leq \epsilon$, the solution x^* of COP 3 satisfies,

$$\|x^* - x\|_2 \leq \frac{C_3 \|x - x_S\|_1}{\sqrt{S}} + \epsilon \quad (4.9)$$

where x_S is the best S -sparse approximation to x and C_3 is a constant.

We note that the real sensing procedure in a spin system calculates the continuous integrals of the form $\int_0^T \kappa_{u_i}(t)b(t)dt$ rather than the discrete inner products $\langle \frac{1}{\sqrt{m}} \tilde{\kappa}_{u_i}, B \rangle$. To take this into account, let us bound the distance between $\frac{n}{T\sqrt{m}} \int_0^T \kappa_u(t)b(t)dt$ and $\frac{1}{\sqrt{m}} \sum_{j=0}^{n-1} \kappa_u(s_j)b(s_j)$. For ease of notation, let $g_u : [0, T] \rightarrow \mathbb{R}$ be defined by $g_u(t) = \kappa_u(t)b(t)$. We have by the the midpoint Riemann sum approximation [40],

$$\begin{aligned} & \left| \frac{n}{T\sqrt{m}} \int_0^T g_u(t)dt - \frac{1}{\sqrt{m}} \sum_{j=0}^{n-1} g_u(s_j) \right| \\ &= \frac{n}{T\sqrt{m}} \left| \int_0^T g_u(t)dt - \frac{T}{n} \sum_{j=0}^{n-1} g_u(s_j) \right| \\ &\leq \frac{n}{T\sqrt{m}} \max_{t \in [0, T]} |b''(t)| \frac{T}{24} \frac{T^2}{n^2} \\ &= \frac{T^2}{24n\sqrt{m}} \max_{t \in [0, T]} |b''(t)|. \end{aligned} \quad (4.10)$$

Hence, if we make m measurements and obtain a vector $y = (y_{u_1}, \dots, y_{u_m})$ where

$$y_{u_j} = \frac{n}{T\sqrt{m}} \int_0^T \kappa_{u_j}(t)b(t)dt, \quad (4.11)$$

then

$$\|y - Ax\|_2 \leq \frac{T^2\sqrt{m}}{24n} \max_{t \in [0, T]} |b''(t)|. \quad (4.12)$$

Setting

$$\epsilon = \frac{T^2\sqrt{m}}{24n} \max_{t \in [0, T]} |b''(t)|, \quad (4.13)$$

implies we can treat y as a noisy version of Ax ,

$$\|y - Ax\|_2 \leq \epsilon. \quad (4.14)$$

Therefore, by Theorem 4 and the fact that $m \sim O(S \log(\frac{n}{S}))$, we have that the CS reconstruction will still be successful up to ϵ (and becomes exact as $n \rightarrow \infty$). Hence, we can treat the difference between the actual continuous integrals we obtain in \tilde{y} and the ideal discrete

measurements y as a fictitious error term in the acquisition process and use Theorem 4. Finally, if there is actual error ϵ_1 in the magnetic field then taking $\epsilon = \epsilon_1 + \tilde{\epsilon}$ and using Theorem 4 again gives that the solution x^* of COP 2 satisfies,

$$\|x^* - x\|_2 \leq \frac{C_3 \|x - x_S\|_1}{\sqrt{S}} + \epsilon. \quad (4.15)$$

We reiterate that Ψ is arbitrary but known *a priori*. More precisely, in order to implement the convex optimization procedures, one needs to know the basis Ψ . However there are no restrictions on what this basis *could* be. If the signal B has an approximately S -sparse representation in the basis Ψ , then the above result implies we can recover B using a small number of samples.

It is also important to note for the case of when Ψ is the Fourier basis, this result is distinct, and in many cases stronger, than Nyquist's theorem. Nyquist's theorem states that if B has finite bandwidth with upper limit f_B , one can reconstruct \tilde{B} by sampling at a rate of $2f_B$. Compressive sensing tells us that even when there is no finite bandwidth of the signal, exact reconstruction is possible using a small number of measurements that depends only on the sparsity level.

Before analyzing numerical examples, we give a theoretical analysis of how to account for decoherence effects. The general idea is similar to the discussion in Sec. III B 1. The main point is that applying a random control sequence according to the random binary string u gives

$$\chi_u(T) = \int_0^\infty \frac{S_\beta(\omega)}{\omega^2} F_u(\omega T), \quad (4.16)$$

where $F_u(\omega T)$ is the filter function associated to this control sequence. In principle, both the nose spectrum $S_\beta(\omega)$ and low frequency behavior of each $F_u(\omega T)$ can be determined [45] so one can estimate $\chi_u(T)$.

The signal S_u one measures from the binary string u is

$$S_u = \frac{1}{2} [v_u(T) \sin(z_u(T))], \quad (4.17)$$

where

$$z_u(T) = \frac{1}{T} \int_0^T \kappa_u(t) b(t) dt. \quad (4.18)$$

Again, noting that an ensemble of measurements is required to acquire the signal above, we have for a zero-average noise $\beta(t)$

$$\left\langle \int_0^T \kappa_u(t) \beta(t) dt \right\rangle = 0, \quad (4.19)$$

Now, for N_u measurement repetitions we have

$$\Delta z_u = \frac{1}{\sqrt{N_u T v_u}}. \quad (4.20)$$

As before, fluctuations in obtaining the measurement results z_u are on the order of $\frac{1}{\sqrt{N_u T v_u}}$ and so the 2-norm distance between the ideal measurement outcome vector z and actual measurement vector y is

$$\|z - \bar{y}\|_2 \sim \sqrt{\sum_u \left[\frac{1}{\sqrt{N_u T v_u}} \right]^2} = \sqrt{\sum_u \frac{1}{N_u T^2 v_u^2}}. \quad (4.21)$$

Since CS reconstructions are robust up to $\|z - y\|_2$, we have that a realistic CS reconstruction in a system such as the NV center is robust up to $\sqrt{\sum_u \frac{1}{N_u T^2 v_u^2}}$.

1. Reconstruction Examples

We provide a set of numerical examples that show how random measurement matrices allow for reconstructing sparse signals in any known basis Ψ . We first analyze the case of $b(t)$ being sparse in the frequency domain and then revisit the case of $b(t)$ being sparse in the time domain (the latter of which can provide a comparison with the results of Sec. III B). To keep the discussion both simple and consistent with the example in Sec. III B, we choose a total acquisition time of $T = 1$ ms, assume pulse widths of $P = 10$ ns, and discretize $[0, T]$ into $n = 2^{10}$ intervals.

Frequency-Sparse Signals

Since we have $n = 2^{10}$ intervals in $[0, 1]$ our high-frequency cut-off is equal to 1.024 MHz. More precisely, a partition of $[0, 1]$ into $\frac{1}{n}$ subintervals implies a resolution of n kHz. We choose a sparsity level $S = 8$ with frequencies (in MHz);

$$f_j \in \{1.024 x_j\}_{j=1}^S, \quad (4.22)$$

where the x_j are chosen uniformly at random

$$x_j \in_r \left\{ \frac{1}{k} \right\}_{k=1}^{1024}. \quad (4.23)$$

For each randomly chosen frequency, we also choose uniformly random phases $\phi_j \in_r [0, 2\pi]$ and amplitudes $A_j \in_r [0, 1]$. The resulting field $b(t)$ is given by

$$b(t) = \sum_{j=1}^S A_j \cos(2\pi f_j t + \phi_j). \quad (4.24)$$

The threshold for the MSQE in this case naturally needs to be larger than for the neural magnetic field reconstruction in Sec. III B. The reason for this is that the majority of the signal b in the neuron case is equal to 0 and thus the total variation of that signal is very small (which leads to an extremely small MSQE when successful CS reconstruction occurs). In the case considered here, choosing

random frequencies leads to signals with a larger total variation.

Fig. 3 shows reconstruction for randomly chosen x_j of $\{\frac{1}{2}, \frac{1}{61}, \frac{1}{78}, \frac{1}{328}, \frac{1}{551}, \frac{1}{788}, \frac{1}{881}, \frac{1}{1022}\}$. Clearly the recon-

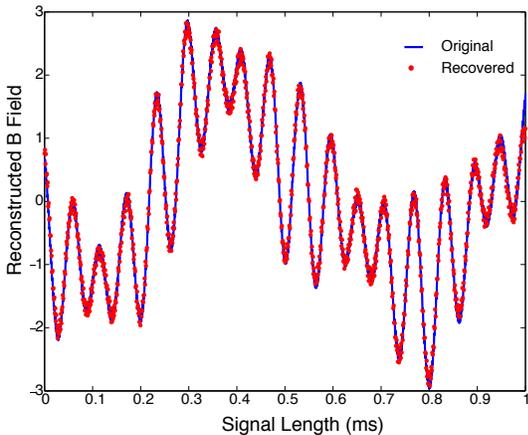


FIG. 3: CS Reconstruction of (Frequency) Sparse Magnetic Field With 8 Random Frequencies (reconstruction with $m = 250$ and $\text{MSQE}=0.0048006$).

struction is fairly accurate. We analyzed the probability of successful CS reconstruction using a MSQE threshold of 0.005 for $m \in \{250, 260, \dots, 350\}$. The results are contained in Table II. As expected, the probability of successful reconstruction quickly converges to 1 for $m \ll 1024$.

Time-Sparse Signals

For this example we used the same parameters as the example in Sec. III B however the measurement matrix is now a random Bernoulli matrix of size $m \times n$. Fig 4 shows the original and reconstructed signal for $m = 250$ where a MSQE of $\sim 10^{-18}$ was obtained. The probability of successful reconstruction followed a similar form to that in Sec. III B in that the probability converged to 1 very quickly as m became larger than 200.

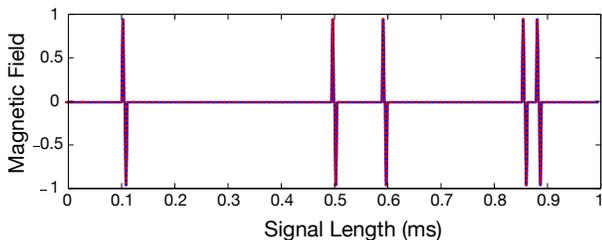


FIG. 4: CS Reconstruction of Time Sparse Magnetic Field Produced By Firing Neuron: Simulated (blue solid) and CS Reconstructed (red dotted) Magnetic Fields (5 events with $m=250$ random Walsh measurements and $\text{MSQE}=3.2184 \times 10^{-4}$)

V. CONCLUSION

We have shown that performing quantum magnetometry using the Zeeman effect in spin systems allows one to utilize the power of compressive sensing (CS) theory. The number of measurements required to reconstruct many physically relevant examples of magnetic fields can be greatly reduced without the degradation of the reconstructed signal. This can have significant impact when measurement resources are extremely valuable, as is the case when using quantum systems for reconstructing magnetic fields. In addition, CS is robust to noise in the signal which makes it an extremely attractive tool to use in physically relevant scenarios, especially when phenomena such as phase jitter are present. We have provided a basic introduction to CS by first discussing the concept of incoherent measurement bases and then moving to the random measurement basis picture. The main idea behind CS is to reconstruct sparse signals using both a small number of non-adaptive measurements in some basis and efficient reconstruction algorithms. We have used l_1 -minimization algorithms however this is only one example of many different methods.

Our first main example utilized the fact that the Walsh and standard measurement basis are maximally incoherent to model the reconstruction of neural spike trains. The signals are sparse in the time-domain and, since we can perform control pulse sequences that represent Walsh sequences, the signal can be represented using a Walsh measurement matrix. We looked at how the probability of successful reconstruction increased as the number of measurements m increased. The probability saturated at 1 when $m \ll n$, where $\log(n)$ is the chosen Walsh reconstruction order. This order will typically be chosen by the physical parameters of the system used to reconstruct the field. The parameters we chose were relevant for a real system such as the Nitrogen-Vacancy (NV) center in diamond. We also verified that the reconstruction is robust to noise in the signal.

The second main section of our results pertains to random measurements, in particular, random symmetric Bernoulli measurement matrices. These measurement matrices satisfy the Restricted Isometry Property (RIP) with high probability when $m \sim O(2S \log(\frac{n}{2S}))$. We first analyzed signals that are sparse in the frequency domain, and showed that efficient reconstruction occurs with probability 1 for $m \ll n$. In addition, we added a Gaussian noise component and showed successful reconstruction still occurs. We again reiterate that in the case of frequency-sparse signals, CS is completely distinct from Nyquist's Theorem and should be viewed as an independent result. This is easily seen by the fact that we were able to sample at a rate much lower than the highest frequency present in the sample and still obtain an exact reconstruction. Next, we revisited the neuron signals and showed that random measurement matrices are just as effective as the Walsh basis at reconstructing the time-sparse signals.

m	250	260	270	280	290	300	310	320	330	340	350
p_{suc}	0.914	0.958	0.983	0.993	0.998	0.999	0.999	1.000	1.000	1.000	1.000

TABLE II: Probability of successful CS reconstruction, p_{suc} , for different values of $m \ll n = 1024$.

There are various directions of future research. First, it will be interesting to investigate whether it is possible to incorporate the framework of CS into other methods for performing magnetometry. In this work, we were able to map control sequences directly onto the evolution of the phase which provides the relevant measurement coefficients for efficient reconstruction of the waveform. A different application of CS may be needed for other sensing mechanisms with quantum systems. While we numerically analyzed an example inspired by neuronal magnetometry, there is a wide array of physical phenomena that produce magnetic fields which display sparsity, or more generally compressibility, in some basis. A more detailed analysis of using CS in these cases will be worthwhile to consider.

Appendix A: The Walsh basis

Let $f : [0, T] \rightarrow \mathbb{R}$ be a square-integrable function, that is, $f \in L^2[0, T]$. Then, f has a Walsh representation [37] given by

$$f(t) = \sum_{j=0}^{\infty} \hat{f}_j w_j(t), \quad (\text{A1})$$

where

$$\hat{f}_j = \frac{1}{T} \int_0^T f(t) w_j \left(\frac{t}{T} \right) dt,$$

and the $w_j : [0, 1] \rightarrow \mathbb{R}$, $j \in \mathbb{N}$, denote the Walsh functions. The Walsh functions are useful from a digital viewpoint in that they are a binary system (they only take the values ± 1) and also form an orthonormal basis of $L^2[0, 1]$.

One can explicitly define the Walsh functions via the set of Rademacher functions $\{R_k\}_{k=1}^{\infty}$. For each $k = 1, 2, \dots$, the k 'th Rademacher function $R_k : [0, 1] \rightarrow \mathbb{R}$ is defined by

$$R_k(t) = (-1)^{t_k}, \quad (\text{A2})$$

where t_k is the k 'th digit of the binary expansion of $t \in [0, 1]$,

$$t = \sum_{k=1}^{\infty} \frac{t_k}{2^k} = \frac{t_1}{2} + \frac{t_2}{2^2} + \frac{t_3}{2^3} + \dots = 0.t_1 t_2 t_3 \dots \quad (\text{A3})$$

Equivalently, one can define R_k as the k 'th square-wave function

$$R_k(t) = \text{sgn}(\sin(2^k \pi t)). \quad (\text{A4})$$

The Walsh basis is just equal to the group (under multiplication) generated by the set of all Rademacher functions. Different orderings of the Walsh basis can be obtained by considering the ordering in which one multiplies Rademacher functions together. For instance, two particularly common orderings of the Walsh basis are the ‘‘sequency’’ [37] and ‘‘Paley’’ [50] orderings. Sequency ordering arises from writing the set of binary sequences according to ‘‘Gray code’’ [51], which corresponds to having each successive sequence differ in exactly one digit from the previous sequence, under the assumption that right-most digits vary fastest. If one assigns each Rademacher function to its corresponding digit in the binary expansion (for instance R_1 is associated to the right-most digit, R_2 is associated to the next digit and so on) then, when each integer i is written in Gray code, $i = \dots i_k \dots i_2 i_1$, we have that the i 'th Walsh function in sequency ordering is

$$w_i(t) = \prod_{k=1}^{\infty} [R_k(t)]^{i_k} = \prod_{k=1}^{\infty} (-1)^{i_k t_k}. \quad (\text{A5})$$

Paley ordering of the Walsh basis is obtained in the same manner as just described for sequency ordering, the only difference being that binary sequences are ordered according to standard binary code (rather than Gray code).

There are a couple of important points to remember about Walsh functions. First, since sequency and Paley orderings differ only in terms of how each integer i is represented in terms of binary sequences, switching between these orderings reduces to switching between Gray and standard binary code ordering. Second, it is straightforward to verify that the set of the first 2^k sequency-ordered Walsh functions is equal to the first 2^k Paley ordered Walsh functions. Thus, rearrangements of orderings differ only within the sets of functions whose size is a power of 2.

For the remainder of this paper, whenever the Walsh basis is used, we will assume the functions are sequency-ordered. We define the n 'th partial sum of $f \in L^2[0, T]$, denoted f_n , to be the sum of the first n terms of its Walsh representation

$$f_n(t) = \sum_{j=0}^{n-1} \hat{f}_j w_j \left(\frac{t}{T} \right). \quad (\text{A6})$$

The N 'th order reconstruction of f corresponds to the 2^N 'th partial sum.

Lastly, we note that for any $n = 2^N$ with $N \geq 0$, one can also define a discrete Walsh basis for \mathbb{R}^n . To see this, first note that the first 2^N Walsh functions $\{w_j\}_{j=0}^{2^N-1}$ are piecewise constant on the 2^N uniform-length subintervals of $[0, 1]$. Now, just associate the values of each w_j on these subintervals to a vector in \mathbb{R}^n . The resulting set of vectors is an orthogonal basis, and dividing each vector

by \sqrt{n} gives the discrete orthonormal Walsh basis, which we denote by $\{W_j\}_{j=0}^{\infty}$. As an example, let $N = 2$ so $n = 4$. Then the four unnormalized Walsh vectors are $\{(1, 1, 1, 1), (1, 1, -1, -1), (1, -1, -1, 1), (1, -1, 1, -1)\}$. Normalizing each vector by 2 gives the orthonormal Walsh basis $\{W_j\}_{j=0}^3$ of \mathbb{R}^4 .

-
- [1] V. Giovannetti, S. Lloyd, and L. Maccone, *Science* **306**, 1330 (2004).
- [2] A. Gruber, A. Drbenstedt, C. Tietz, L. Fleury, J. Wrachtrup, and C. v. Borczyskowski, *Science* **276**, 2012 (1997).
- [3] R. C. Jaklevic, J. Lambe, A. H. Silver, and J. E. Mercereau, *Phys. Rev. Lett.* **12**, 159 (1964).
- [4] S. Xu, V. V. Yashchuk, M. H. Donaldson, S. M. Rochester, D. Budker, and A. Pines, *Proc Natl Acad Sci* **103**, 12668 (2006).
- [5] L. T. Hall, G. C. G. Beart, E. A. Thomas, D. A. Simpson, L. P. McGuinness, J. H. Cole, J. H. Manton, R. E. Scholten, F. Jelezko, J. Wrachtrup, S. Petrou, and L. C. L. Hollenberg, *Nature Sci. Rep.* **2** (2012), 10.1038/srep00401.
- [6] L. M. Pham, D. Le Sage, P. L. Stanwix, T. K. Yeung, D. Glenn, A. Trifonov, P. Cappellaro, P. R. Hemmer, M. D. Lukin, H. Park, A. Yacoby, and R. L. Walsworth, *Phys. Rev. Lett.* **106**, 045021 (2011).
- [7] E. B. Aleksandrov, *Physics-Uspekhi* **53**, 487 (2010).
- [8] D. F. Blair, *The Detection of Gravitational Waves* (Cambridge University Press, Great Britain, 1991).
- [9] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. G. Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin, *Nature* **455**, 644 (2008).
- [10] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch, *Phys. Rev.* **53**, 318 (1938).
- [11] R. Damadian, *Science* **171**, 1151 (1971).
- [12] A. L. Bloom, *Appl. Opt.* **1**, 61 (1962).
- [13] J. Dupont-Roc, S. Haroche, and C. Cohen-Tannoudji, *Phys. Lett. A* **28**, 638 (1969).
- [14] N. F. Ramsey, *Phys. Rev.* **78**, 695 (1950).
- [15] A. Cooper, E. Magesan, H. N. Yum, and P. Cappellaro, (2013), arXiv:1305.6082.
- [16] E. J. Candes, J. K. Romberg, and T. Tao, *Communications on Pure and Applied Mathematics* **59**, 1207 (2006).
- [17] D. Donoho, *Information Theory, IEEE Transactions on* **52**, 1289 (2006).
- [18] M. Lustig, D. Donoho, and J. M. Pauly, *Magn. Reson. Med.* **58**, 1182 (2007).
- [19] G. Taubock and F. Hlawatsch, in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on* (2008) pp. 2885–2888.
- [20] W. Dei, M. Sheikh, O. Milenkovic, and R. Baraniuk, *EURASIP J. Bioinform. Syst. Biol.* **1**, 162824 (2009).
- [21] T. T. Lin and F. J. Herrmann, *Geophysics* **72**, SM77 (2007).
- [22] R. Baraniuk and P. Steeghs, in *Radar Conference, 2007 IEEE* (2007) pp. 128–133.
- [23] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Phys. Rev. Lett.* **105**, 150401 (2010).
- [24] A. Griffin, T. Hirvonen, C. Tzagkarakis, A. Mouchtaris, and P. Tsakalides, *Audio, Speech, and Language Processing, IEEE Transactions on* **19**, 1382 (2011).
- [25] P. Sen and S. Darabi, *IEEE Transactions on Visualization and Computer Graphics* **17**, 487 (2011).
- [26] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, *Phys. Rev. Lett.* **106**, 100401 (2011).
- [27] D. Donoho and X. Huo, *Information Theory, IEEE Transactions on* **47**, 2845 (2001).
- [28] W. Rudin, *Functional Analysis* (McGraw-Hill Science/Engineering/Math, 1991).
- [29] E. Candes and J. Romberg, *Inverse Problems* **23**, 969 (2007).
- [30] E. Candes and T. Tao, *Information Theory, IEEE Transactions on* **51**, 4203 (2005).
- [31] S. Foucart, *Appl. and Comp. Harmonic Analysis* **29**, 97 (2010).
- [32] R. Baraniuk, M. Davenport, R. Devore, and M. Wakin, *Constr. Approx.* **2008** (2007).
- [33] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin, *Nature Physics* **4**, 2417 (2008).
- [34] E. L. Hahn, *Phys. Rev.* **80**, 580 (1950).
- [35] H. Y. Carr and E. M. Purcell, *Phys. Rev.* **94**, 630 (1954).
- [36] S. Meiboom and D. Gill, *Rev. Sci. Instrum.* **29**, 688 (1958).
- [37] J. L. Walsh, *Amer. J. Math.* **45**, 5 (1923).
- [38] P. Dayan and L. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, U.S.A., 2005).
- [39] D. Hayes, K. Khodjasteh, L. Viola, and M. J. Biercuk, *Phys. Rev. A* **84**, 062323 (2011).
- [40] M. Spivak, *Calculus* (Cambridge University Press, Cambridge, UK, 1967).
- [41] J. K. Woosley, B. J. Roth, and J. P. W. Jr., *Math. Biosciences* **76**, 1 (1985).
- [42] L. S. Chow, A. Dagens, Y. Fu, G. G. Cook, and M. N. Paley, *Magnetic Resonance in Medicine* **60**, 1147 (2008).
- [43] S. M. Anwar, G. G. Cook, L. S. Chow, and M. N. Paley, in *Proc. Intl. Soc. Mag. Reson. Med.* **17** (2009).
- [44] S. Murakami and Y. Okada, *The Journal of Physiology* **575**, 925 (2006).
- [45] G. Uhrig, *New J. Phys.* **10**, 083024 (2008).
- [46] L. Cywiński, R. M. Lutchyn, C. P. Nave, and S. Das Sarma, *Phys. Rev. B* **77**, 174509 (2008).
- [47] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. Cory, N. Yasunobu, J.-S. Shen, and W. Oliver, *Nature Physics* **7**, 565 (2011).
- [48] N. Bar-Gill, L. Pham, C. Belthangady, D. Le Sage, P. Cappellaro, J. Maze, M. Lukin, A. Yacoby, and R. Walsworth, *Nat. Commun.* **2**, 858 (2012).

- [49] E. Magesan, A. Cooper, H. N. Yum, and P. Cappellaro, (2013), arXiv:1305.6604.
- [50] R. E. A. C. Paley, Proc. London Math. Soc. **34**, 241 (1932).
- [51] F. Gray, U.S. patent no. 2632058 (1953).