Noise-reducing encoding strategies for spin chains

Catherine Keele and Alastair Kay

Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, England, United Kingdom

We present an encoding technique that reduces the effects of noise on quantum spin systems whose operation is driven by Hamiltonian evolution. This technique is widely applicable, being most relevant to the scenarios where there are insufficient qubits to permit full-scale error correction. Instead, our technique can be implemented over small numbers of qubits and still leads to noticeable improvements in the fidelity of operations. The encoding scheme is easy to implement, flexible with respect to choice of Hamiltonian, and close to optimal.

I. INTRODUCTION

Quantum computers hold the promise of being able to solve problems that are too complex for classical computers [1]; however, fault-tolerant quantum computation remains firmly in the future. To harness the power of quantum computation on near-term devices, there is currently a focus on noisy intermediate-scale quantum (NISQ) devices, which operate in the presence of noise but without access to full error correction due to a limited number of qubits [2]. Even in this setting, quantum supremacy was recently demonstrated [3]. Given the noise, but a lack of error correction, we must discover alternative strategies for error mitigation. Here we introduce a simple technique that can mitigate the effects of noise, improving the quality of operation in devices whose dynamics is driven by Hamiltonian evolution, using a small number of qubits relative to the number available.

The study of state transfer [4–7] is a key test bed for the development of ideas based upon Hamiltonian evolution. It provides a concrete task, that of transferring an unknown quantum state from one qubit to another, for us to study and demonstrates some of the key criteria for use in a NISQ device in that it is a useful elementary building block within more complex tasks and, in the Hamiltonian formulation, permits a factor of 2 speed enhancement compared to the circuit model [8] which, while irrelevant to a computational scaling perspective, could be absolutely critical to achieving the maximum number of computational steps in a finite time before decoherence overwhelms the system.

The primary focus of early studies of state transfer was on perfect transfer [5]. However, in an imperfect world, perfect transfer can never be achieved and it is preferable to consider near-perfect transfer if, in trade, the system might be more tolerant of noise by, for example, achieving the transfer faster. Numerous methods have been considered for creating high-quality transfer chains, from modifying some of the couplings in a chain [9] to encoding inputs and outputs [10–12].

Nevertheless, these studies have generally focused on unitary evolution. Few studies have even quantified the effects of noise [7,13], let alone attempted to directly improve the tolerance to noise. Those that have require unrealistic assumptions about the nature of the environment [14,15] or knowledge of when and where errors might happen [16]. Ultimately, these effects can be addressed by error correction [17–20], but we are interested in the regime for which error correction is not a realistic prospect because it requires too many qubits and too many operations.

Our approach is an encoding strategy that is a lighter touch than error correction. It generalizes a method introduced by Haselgrove [11], in which an optimal encoding could be found for the task of quantum state transfer under unitary evolution. We further developed the technique in [12] for application to unitary evolution, but we now include the effects of noise. The method is broadly applicable to a wide range of Hamiltonians and noise types and could readily be applied to tasks beyond that of state transfer. We show how two common types of noise, i.e., amplitude damping and dephasing, may be treated within our formalism, on an otherwise perfect system, and how fidelity in these types of system can be improved by use of encoding strategies.

In Sec. II we introduce the setting of our technique and the previous work by Haselgrove [11] that we build upon here. Section III introduces our system and a measure of success. Section IV introduces our encoding scheme. In Sec. V we incorporate the treatment of noise. We demonstrate the application of our encoding scheme to two example Hamiltonians: one that implements perfect state transfer [5,7] and another specified in [9] which is the best chain that we know of in terms of the speed-fidelity trade-off. These are just expository and neither of these has been optimized for the scenario. We concentrate primarily on the single-excitation subspace, but show how the results can be extended into the regime of multiple-excitation subspaces in Sec. VII. Throughout the paper, we consider dimensionless quantities by taking $\hbar = 1$.

II. SETTING

We consider a set of N qubits, all prepared in the state $|0\rangle$. The task of state transfer requires the introduction of an unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ in a given site a and for the evolution of the system Hamiltonian to cause this state

to move to the output site b. We assume that the Hamiltonian H is excitation preserving, meaning $[H, \sum_{n=1}^{N} Z_n] = 0$. As such, $|\mathbf{0}\rangle := |0\rangle^{\otimes N}$ is an eigenstate. (We will typically use $|\mathbf{0}\rangle$ to denote the all-0 state of any number of qubits, where the number should be clear from context.) Thus, the only evolution that we have to focus on is that of the single-excitation subspace, which is spanned by a basis $|\mathbf{n}\rangle = |0\rangle^{\otimes(n-1)}|1\rangle|0\rangle^{\otimes(N-n)}$, within which the Hamiltonian is H_1. Without loss of generality, we take the input site to be labeled 1 and the output site N. Thus, perfect state transfer occurs in a system at time t if $|e_{N,1}^{-iH_1 t}| = 1$.

We will mostly focus on the single-excitation subspace. This makes the task computationally tractable, but is also motivated by [12], in which it was shown that, for unitary evolution, the single-excitation subspace is the optimal for encoding in across a broad parameter range. These encodings were capable of outperforming error-correcting codes of the same size.

A. Encoding

Instead of controlling a single site at the input and output, we assume control of two (small) sets of sites Λ_{in} and Λ_{out}. We are able to prepare a single-excitation state $|\psi\rangle$ on the sites Λ_{in},

$$|\Psi_{\text{in}}\rangle = \alpha|\mathbf{0}\rangle + \beta|\psi\rangle,$$

and receive it on the sites Λ_{out}. In the context of unitary evolution, the challenge of finding the optimal encoding and decoding strategy for a fixed H_1 and time t was solved by Haselgrove [11]. If

$$P_{\text{in}} = \sum_{i \in \Lambda_{\text{in}}} |\mathbf{i}\rangle\langle\mathbf{i}|, \quad P_{\text{out}} = \sum_{i \in \Lambda_{\text{out}}} |\mathbf{i}\rangle\langle\mathbf{i}|$$

are the projectors onto the input and output regions, then $P_{\text{in}}|\psi\rangle = |\psi\rangle$. We proceed by evaluating the $|\Lambda_{\text{out}}| \times |\Lambda_{\text{in}}|$ matrix

$$\tilde{U}_t = P_{\text{out}} e^{-iH_1 t} P_{\text{in}}.$$

If we select an input state $|\psi\rangle$ to be a right singular vector of $\tilde{U}_t$ and $|\phi\rangle$ to be the corresponding left-singular vector, then by encoding in $|\psi\rangle$ and waiting a time t, the arriving state on the decoding region is $|\phi\rangle$ with a probability amplitude corresponding to the singular value. The maximum success probability is therefore just the square of the largest singular value of $\tilde{U}_t$. We aim to generalize this encoding strategy to the context of noisy systems.

III. VECTORIzATION

In order to treat noise in a straightforward manner, we make use of a vectorization procedure for the density matrix [21] so that we have a density vector given as

$$|\rho\rangle = \sum_{ij} \langle i|\rho|j\rangle |ij\rangle.$$

The advantage of this is that it converts noise superoperators into linear operators. As an example, for $|\Psi_{\text{in}}\rangle = \alpha|\mathbf{0}\rangle + \beta|\mathbf{1}\rangle$, the density matrix is $\rho = |\Psi_{\text{in}}\rangle\langle\Psi_{\text{in}}|$ and the density vector is

$$|\rho\rangle = |\alpha|^2|\mathbf{00}\rangle + \alpha^*\beta|\mathbf{10}\rangle + \alpha\beta^*|\mathbf{01}\rangle + |\beta|^2|\mathbf{11}\rangle$$
$$= |\Psi_{\text{in}}\Psi_{\text{in}}^*\rangle.$$

We need to know how operators on the density matrix ρ manifest within this formalism.

Lemma 1. Let A and B be two linear operators that act on a density matrix ρ. The vectorized form of $A\rho B$ is

$$|A\rho B\rangle = A \otimes B^T |\rho\rangle.$$

Proof. Applying the definition for vectorization, we have

$$|A\rho B\rangle = \sum_{kl} \langle k|A\rho B|l\rangle |k, l\rangle.$$

Using the completeness relation yields

$$|A\rho B\rangle = \sum_{ijkl} \langle k|A|i\rangle\langle i|\rho|j\rangle\langle j|B|l\rangle |k, l\rangle$$
$$= \sum_{ijkl} |k, l\rangle\langle k, l|A|i\rangle\langle l|B^T|j\rangle\langle i|\rho|j\rangle$$
$$= \left(\sum_{kl} |k, l\rangle\langle k, l|A \otimes B^T\right) \sum_{ij} \langle i|\rho|j\rangle |i, j\rangle$$
$$= A \otimes B^T |\rho\rangle. \qquad \blacksquare$$

As our system can be divided into subspaces, we denote the components of $|\rho\rangle$ on a given subspace by $|\rho_{00}\rangle, |\rho_{01}\rangle, |\rho_{10}\rangle$ and $|\rho_{11}\rangle$. The 00 subspace is a single element $|\mathbf{00}\rangle$, while the 11 subspace is spanned by basis states $|\mathbf{ij}\rangle$. The 01 and 10 subspaces are the coherences between these two and are spanned by $|\mathbf{0i}\rangle$ and $|\mathbf{i0}\rangle$, respectively.

A. Noisy evolution

We describe noise using the Lindblad master equation

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_{n=1}^{N} \left(L_n \rho L_n^\dagger - \frac{1}{2}L_n^\dagger L_n \rho - \frac{1}{2}\rho L_n^\dagger L_n\right),$$

where the $\{L_n\}$ specify the noise. Under our vectorization technique, we write

$$\mathcal{Q} = \mathcal{H} + \sum_n L_n \otimes L_n^* - \frac{1}{2}L_n^\dagger L_n \otimes \mathbb{1} - \frac{1}{2}\mathbb{1} \otimes (L_n^\dagger L_n)^*,$$

where $\mathcal{H} = -iH \otimes \mathbb{1} + i\mathbb{1} \otimes H^T$. The Lindblad equation then becomes

$$\frac{d|\rho\rangle}{dt} = \mathcal{Q}|\rho\rangle$$

such that the evolution is given by

$$|\rho(t)\rangle = e^{\mathcal{Q}t}|\rho(0)\rangle.$$

B. Trace

We will also need to take the partial trace over a set of sites $\bar{\Lambda}$, leaving just the set of qubits Λ remaining. We define

$$T_\Lambda|\rho\rangle = |\text{Tr}_{\bar{\Lambda}}(\rho)\rangle.$$

This is also a linear operator; if $\{|u_i\rangle\}$ is an orthonormal basis over the qubits $\bar{\Lambda}$,

$$\text{Tr}_{\bar{\Lambda}}(\rho) = \sum_i (\mathbb{1}_\Lambda \otimes \langle u_i|)\rho(\mathbb{1}_\Lambda \otimes |u_i\rangle),$$

such that, by Lemma 1,

$$T_\Lambda = \sum_i (\mathbb{1}_\Lambda \otimes \langle u_i|) \otimes (\mathbb{1}_\Lambda \otimes \langle u_i^*|).$$

Typically, one picks the standard basis for performing the trace, in which case $|u_i^*\rangle = |u_i\rangle$.

C. Quality of transfer

If our aim is to successfully transfer a state from one location to another, we must introduce a measure of success. For Hamiltonian evolution, the measure of success is the transfer fidelity [5], whose derivation we reproduce here, using the vectorized notation, before later (Secs. IV and V) expanding it to include the noisy evolution and encoding and decoding.

The density matrix after evolution is given as

$$|\rho'\rangle = U_t \otimes U_t^* |\rho\rangle,$$

where $U_t = e^{-iHt}$. The fidelity of state transfer is given by

$$F = (\alpha^*\langle 0| + \beta^*\langle 1|) \otimes (\alpha\langle 0| + \beta\langle 1|) T_N |\rho'\rangle$$

for a specific input state $\alpha|0\rangle + \beta|1\rangle$. However, to truly judge the efficacy of the protocol, one should average over all possible input states, by identifying $\alpha = \cos\frac{\theta}{2}$ and $\beta = \sin\frac{\theta}{2}$ such that the average fidelity of state transfer is

$$\overline{F} = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi F \sin\theta \, d\theta \, d\phi$$
$$= \frac{1}{6}(3 + 2\sqrt{F_{\text{ex}}} + F_{\text{ex}}),$$

where, for an excitation preserving Hamiltonian,

$$F_{\text{ex}} = \langle 11| T_N e^{\mathcal{H}_1 t} |11\rangle = |\langle \mathbf{N}| e^{-iH_1 t} |\mathbf{1}\rangle|^2$$

and $\mathcal{H}_1 = -iH_1 \otimes \mathbb{1} + i\mathbb{1} \otimes H_1^T$. Although we want our transfer fidelity to be as high as possible, we note that there is a natural threshold of $\frac{2}{3}$ [4], which is the fidelity achieved by classically transferring a quantum state, that we need to beat.

IV. ENCODING STRATEGY

We now introduce our encoding scheme, where the initial state is encoded over the zero- and one-excitation subspaces of a set of qubits Λ_{in}. Our aim is to find the ideal choice of state $|\psi\rangle$ for the initial encoding of $|\Psi_{\text{in}}\rangle = \alpha|0\rangle + \beta|\psi\rangle$ and a decoding unitary U acting on the decoding region Λ_{out} such that the state $\alpha|0\rangle + \beta|1\rangle$ is reproduced on a single site with the maximum fidelity. We are assuming here that U is performing the decoding onto a single qubit in the decoding region. In [12] we took a different approach of encoding onto a separate ancilla qubit. However, our motivation here is that we will be encoding over as many qubits as we can, so that precludes the possibility of making additional qubits interact.[1]

---

[1]Ultimately this choice makes no difference until we look at higher-excitation subspaces but is included to demonstrate a different assumption.

The initial state evolves through time according to

$$|\rho'\rangle = e^{\mathcal{Q}t}(\alpha|\mathbf{0}\rangle + \beta|\psi\rangle)(\alpha^*|\mathbf{0}\rangle + \beta^*|\psi^*\rangle),$$

which we will then decode. Upon decoding, we trace out all other qubits because they are irrelevant. Hence, the transfer fidelity is

$$F = (\alpha^*\langle 0| + \beta^*\langle 1^*|)(\alpha\langle 0| + \beta\langle 1|) T_N U \otimes U^* T_{\Lambda_{\text{out}}} |\rho'\rangle.$$

Let $\mathcal{R} = T_N U \otimes U^* T_{\Lambda_{\text{out}}} e^{\mathcal{Q}t}$. As in Sec. III C, we average over all possible input states (parameters α and β) to give

$$\overline{F} = \frac{1}{6}(2\langle 00|\mathcal{R}|00\rangle + 2\langle 11|\mathcal{R}|\psi\psi^*\rangle + \langle 00|\mathcal{R}|\psi\psi^*\rangle$$
$$+ \langle 11|\mathcal{R}|00\rangle + \langle 01|\mathcal{R}|0\psi^*\rangle + \langle 10|\mathcal{R}|\psi 0\rangle).$$

Since $\mathcal{R}|\psi\psi^*\rangle$ describes a one-qubit density matrix, which has trace 1, we have that $\langle 00|\mathcal{R}|\psi\psi^*\rangle + \langle 11|\mathcal{R}|\psi\psi^*\rangle = 1$. We make one further assumption about the noise model, that it is excitation nonincreasing. This means that $e^{\mathcal{Q}t}|\mathbf{00}\rangle = |\mathbf{00}\rangle$. From this we infer that the decoding unitary should map

$$U|\mathbf{0}\rangle_{\Lambda_{\text{out}}} = |\mathbf{0}\rangle_{\Lambda_{\text{out}}}.$$

Consequently, $\langle 00|\mathcal{R}|00\rangle = 1$ and $\langle 11|\mathcal{R}|00\rangle = 0$. Hence, under this assumption,

$$\overline{F}(t) = \frac{1}{6}(3 + \langle 11|\mathcal{R}|\psi\psi^*\rangle$$
$$+ \langle 01|\mathcal{R}|0\psi^*\rangle + \langle 10|\mathcal{R}|\psi 0\rangle). \quad (1)$$

We remain free to choose $|\psi\rangle$ and the action of U on the single-excitation subspace, to maximize $\overline{F}$. We start by considering the components $\langle 10|\mathcal{R}|\psi 0\rangle$ and $\langle 11|\mathcal{R}|\psi\psi^*\rangle$ separately.

We start with $\langle 11|\mathcal{R}|\psi\psi^*\rangle$. How are we to pick U? Note that if there is an excitation on the decoding region, that can only result from an excitation in the input (as the noise cannot introduce excitations). As such, we definitely want to provide a $|1\rangle$ state on the output if possible. Consequently, we impose that

$$U|\mathbf{n}\rangle = |N, \phi_n\rangle \, \forall \, n \in \Lambda_{\text{out}},$$

i.e., a single excitation on the output spin, and some arbitrary state over the other qubits of the decoding region, subject to the $\{|\phi_n\rangle\}$ forming an orthonormal basis over $\Lambda_{\text{out}\backslash N}$.

If we explicitly write out the state after evolution, the only terms remaining in the (1,1) subspace are

$$e^{\mathcal{Q}t}|\psi\psi^*\rangle \to \sum_{n,m=1}^N \gamma_{nm}|\mathbf{nm}\rangle.$$

Thus,

$$\langle 11|\mathcal{R}|\psi\psi^*\rangle = \langle 11| T_N U \otimes U^* \sum_{n,m\in\Lambda_{\text{out}}} \gamma_{nm}|\mathbf{nm}\rangle$$

$$= \langle 11| T_N \sum_{n,m\in\Lambda_{\text{out}}} \gamma_{nm}|N, \phi_n\rangle|N, \phi_m^*\rangle.$$

Using the orthonormal basis $\{|\phi_n\rangle\}$ for the trace T_N leaves

$$\langle 11|\mathcal{R}|\psi\psi^*\rangle = \sum_{n\in\Lambda_{\text{out}}} \gamma_{nn}.$$

This allows us to see that our choice of U is irrelevant (beyond our earlier very natural assumptions) for the $\langle 11|\mathcal{R}|\psi\psi^*\rangle$

component. We need now only to find the input state $|\psi\rangle$ to maximize $\sum \gamma_{nn}$.

Let

$$R(t) = \sum_{i,j \in \Lambda_{\text{in}}} \sum_{m \in \Lambda_{\text{out}}} |j\rangle\langle i|\langle \mathbf{mm}|e^{\mathcal{Q}t}|\mathbf{ij}\rangle$$

such that $\langle 11|\mathcal{R}|\psi\psi^*\rangle = \sum \gamma_{nn} = \langle \psi|R|\psi\rangle$. It follows that we are able to maximize the $\langle 11|\mathcal{R}|\psi\psi^*\rangle$ component by selecting $|\psi\rangle$ to be the eigenvector of R with the maximum eigenvalue.

We now continue, to understand how to independently maximize the other component in Eq. (1). It is sufficient to maximize only $\langle 10|\mathcal{R}|\psi 0\rangle$, as this can always be made real by incorporating a phase on U such that $\langle 10|\mathcal{R}|\psi 0\rangle = \langle 01|\mathcal{R}|0\psi^*\rangle$ and both are real.

For an excitation nonincreasing $\mathcal{Q}$, after evolution we can parametrize the term

$$e^{\mathcal{Q}t}|\psi 0\rangle = \gamma_{00}|\mathbf{00}\rangle + \sum_{n=1}^{N} \gamma_{n0}|\mathbf{n0}\rangle.$$

One can readily calculate

$$U \otimes U^* T_{\Lambda_{\text{out}}} e^{\mathcal{Q}t}|\psi 0\rangle$$
$$= \gamma_{00}|\mathbf{00}\rangle + \sum_{n \in \Lambda_{\text{out}}} \gamma_{n0}|N, \phi_n\rangle|\mathbf{0}\rangle. \quad (2)$$

The application of the final trace yields

$$\mathcal{R}|\psi 0\rangle = \gamma_{00}|00\rangle + \sum_{n \in \Lambda_{\text{out}}} \gamma_{n0}\langle \mathbf{0}|\phi_n\rangle|10\rangle. \quad (3)$$

The required overlap

$$\langle 10|\mathcal{R}|\psi 0\rangle = \sum_{n \in \Lambda_{\text{out}}} \gamma_{n0}\langle \mathbf{0}|\phi_n\rangle = \sum_{n \in \Lambda_{\text{out}}} \gamma_{n0}\langle \mathbf{N}|U|\mathbf{n}\rangle$$

is maximized by setting $U^\dagger|\mathbf{N}\rangle$ parallel to $\sum_n \gamma_{n0}|\mathbf{n}\rangle$. One component of each of these states is fixed,

$$\langle \mathbf{0}|\phi_n\rangle = \frac{\gamma_{n0}^*}{\sqrt{\sum_{m \in \Lambda_{\text{out}}} |\gamma_{m0}|^2}},$$

and leaves them otherwise free. This yields an optimal value of the component

$$\langle 10|\mathcal{R}|\psi 0\rangle = \sqrt{\sum_{m \in \Lambda_{\text{out}}} |\gamma_{m0}|^2},$$

which can alternatively be expressed as

$$\|(P_{\text{out}} \otimes \langle \mathbf{0}|)e^{\mathcal{Q}t}(|\psi\rangle \otimes |\mathbf{0}\rangle)\|.$$

In the case of unitary evolution, this recovers the result of Haselgrove [11]. Therefore, our encoding $|\psi\rangle$ will be the eigenvector corresponding to the largest right singular vector of

$$S(t) = (P_{\text{out}} \otimes \langle \mathbf{0}|)e^{\mathcal{Q}t}(P_{\text{in}} \otimes |\mathbf{0}\rangle). \quad (4)$$

Instead of finding the maximum right singular vector of S, we can alternatively find the maximum eigenvector of $S^\dagger S$.

Note that, in optimizing the two components separately, the conditions on the choice of U are mutually compatible; it is

only the choice of encoding $|\psi\rangle$ that could potentially differ. The overall expression for the fidelity is thus

$$\bar{F}(t) = \frac{1}{2} + \frac{1}{3}\sqrt{\langle \psi|S^\dagger S|\psi\rangle} + \frac{1}{6}\langle \psi|R|\psi\rangle. \quad (5)$$

An exactly optimal choice of $|\psi\rangle$ in all circumstances is nontrivial, but we will see in Sec. V that there are instances where this can be solved exactly. Furthermore, an extremely good approximation can be made in many reasonable cases. Equally, our intended operating regime for these encodings is with small sizes of the encoding and decoding region, for which exact optimization is possible.

V. SPECIAL CASES

We will now study some special cases in which we can find the optimal $|\psi\rangle$ and hence evaluate $\bar{F}$.

A. Unitary evolution

We now derive the optimal encoding strategy in the noise-free case, which coincides exactly with Haselgrove's strategy [11]. We consider first the R term

$$R = \sum |j\rangle\langle i|\langle \mathbf{mm}|e^{(-iH_1 \otimes \mathbb{1} + i\mathbb{1} \otimes H_1^T)t}|\mathbf{i}, \mathbf{j}\rangle. \quad (6)$$

We can rearrange this to get

$$R = \sum_{i,j \in \Lambda_{\text{in}}} \sum_{m \in \Lambda_{\text{out}}} |j\rangle\langle i|\langle \mathbf{j}|e^{iH_1 t}|\mathbf{m}\rangle\langle \mathbf{m}|e^{-iH_1 t}|\mathbf{i}\rangle$$
$$= \sum_{i,j} |j\rangle\langle i|\langle \mathbf{j}|e^{iH_1 t}P_{\text{out}}e^{-iH_1 t}|\mathbf{i}\rangle$$
$$= S^\dagger S.$$

In this case, the R term does not further constrain our choice of $|\psi\rangle$ and we are free to pick it to optimize the S term via Eq. (4), thereby reproducing the strategy of Haselgrove.

B. Large decoding region

Consider the case where Λ_{out} comprises every qubit in the system. This is clearly not a realistic scenario, but is nevertheless interesting. We restrict the noise model to being excitation preserving, with the maximally mixed state (of each excitation subspace) being the fixed points of the map. One such example is dephasing noise. The state $\sum_{m=1}^{N}|\mathbf{mm}\rangle$ is an eigenstate of $\mathcal{Q}$ because it is the maximally mixed state and thus the fixed point of the map. Hence, $R = \mathbb{1}$. Again, the choice of $|\psi\rangle$ is irrelevant to this term and we need only consider the term arising from S.

C. Amplitude damping noise

One important type of noise that we can treat within this formalism is amplitude damping noise. This describes the relaxation of a system as it loses energy to the environment. The Lindblad operators are $L_i = \sqrt{\frac{\Gamma_x}{2}}(X_i + iY_i)$, where $\Gamma_x = \frac{1}{T_1}$ is a measure of the strength of the noise, T_1 being the longitudinal paramagnetic relaxation time [22].

Due to the structure of the amplitude damping noise $\mathcal{Q}_A$, we can find analytical solutions for evolution constrained
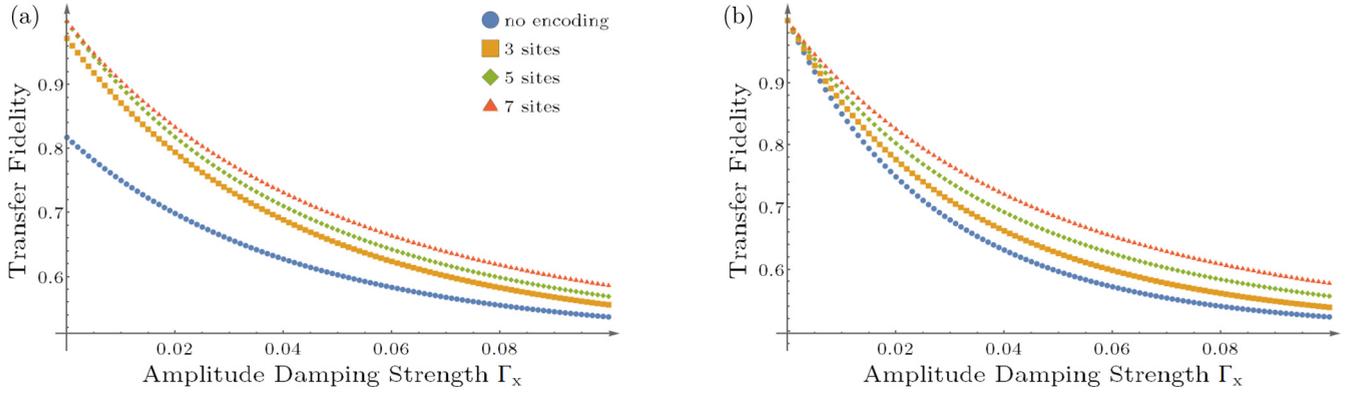
FIG. 1. Performance of different Hamiltonians in the presence of amplitude damping noise, chain length 35, combated with encoding of differing sizes: (a) uniformly coupled Hamiltonian and (b) perfect state transfer Hamiltonian [5].

within the zero- and single-excitation subspaces by considering the evolution of each subspace of the density vector separately. We first look at evolution of the $|11\rangle$ subspace, which is given by

$$\frac{d|\rho_{11}\rangle}{dt} = (-2\Gamma_x \mathbb{1} + \mathcal{H}_1)|\rho_{11}\rangle.$$

This has a direct solution

$$|\rho_{11}(t)\rangle = e^{-2\Gamma_x t} e^{\mathcal{H}_1 t} |\rho_{11}(0)\rangle.$$

Applying the same logic to the other subspaces of $|\rho\rangle$ tells us that the components evolve as

$$|\rho_{00}\rangle \rightarrow |\rho_{00}\rangle, \tag{7}$$

$$|\rho_{01}\rangle \rightarrow e^{-\Gamma_x t} e^{iH_1 t} |\rho_{01}\rangle, \tag{8}$$

$$|\rho_{10}\rangle \rightarrow e^{-\Gamma_x t} e^{-iH_1 t} |\rho_{10}\rangle, \tag{9}$$

$$|\rho_{11}\rangle \rightarrow e^{-2\Gamma_x t} e^{\mathcal{H}_1 t} |\rho_{11}\rangle + (1 - e^{-2\Gamma_x t})|\rho_{00}\rangle. \tag{10}$$

This result gives us an exact solution for the evolution of the density vector when subject to amplitude damping noise and shows that the effect of this noise is just to add the deterioration term Γ_x. We can see instantly that solutions can be taken directly from the noise-free case and our only opportunity to minimize noise is to make the transfer as fast as possible.

In this instance, we have that $S = e^{-\Gamma_x t} S_0$, where S_0 is the S matrix in the absence of noise. Similarly,

$$R = e^{-2\Gamma_x t} \sum |j\rangle\langle i|\langle \mathbf{mm}|e^{\mathcal{H}_1 t}|\mathbf{ij}\rangle$$
$$= e^{-2\Gamma_x t} \sum_{i,j} |j\rangle\langle i|\langle \mathbf{j}|e^{iH_1 t} P_{\text{out}} e^{-iH_1 t}|\mathbf{i}\rangle = S^\dagger S.$$

Again, R does not affect the choice of $|\psi\rangle$. Hence, the strength of the noise does not affect the choice of $|\psi\rangle$, so one can find the optimal encoding in the noiseless case (i.e., utilizing Haselgrove's technique), and this is the optimal encoding for all noise strengths, just with reduced fidelity

$$\bar{F}_{\Gamma_x} = \frac{1}{3} + \frac{1}{6}[1 + e^{-\Gamma_x t}(\sqrt{6\bar{F}_0 - 2} - 1)]^2.$$

Even so, when considering optimizing over time as well, adding noise will tend to make the optimal time (marginally) earlier.

In Fig. 1 we plot the effects of amplitude damping noise (having optimized for time) for two different Hamiltonians. It is noteworthy that by the time we use an encoding and decoding region of size 7, the transfer fidelity has been significantly enhanced and there is essentially no difference between the performance of the two Hamiltonians.

D. Optimizing over components

We have now seen a number of cases in which R does not influence the optimal choice of $|\psi\rangle$, which is just selected to be the maximum eigenvector of $S^\dagger S$ as $R \propto \mathbb{1}$ of $S^\dagger S$. In the case of general noise, we do not expect this to always hold, but we anticipate that R will hold less relevance and the S term will dominate. This motivates our simplifying assumption that $|\psi\rangle$ will be close to being an eigenvector of $S^\dagger S$, in which case we can approximate

$$\bar{F}(t) \approx \frac{1}{2} + \frac{1}{6}\langle\psi|2\sqrt{S^\dagger S} + R|\psi\rangle. \tag{11}$$

Thus, $|\psi\rangle$ is just the maximum eigenvector of $2\sqrt{S^\dagger S} + R$. Using the $|\psi\rangle$ in this way must represent a lower bound on the achievable fidelity. In contrast, the independent optimization of the R and S terms yields an upper bound to this value. In Fig. 2 we take the case of dephasing noise ($L_i = \sqrt{\frac{\Gamma_z}{2}}Z_i$), which is not expected to have $R \propto \mathbb{1}$. For the maximum possible opportunity to see a discrepancy between the upper and lower bounds, we push the dephasing so strong as to render the transfer fidelities unusable. Even at this extreme, we see that the upper and lower bounds coincide and thus anticipate that they will do so at all intermediate regimes. We thus expect this method to be essentially optimal across all relevant parameters.

VI. EXAMPLES

For the sake of concrete examples, we choose to perform transfer along the length of a chain coupled via a
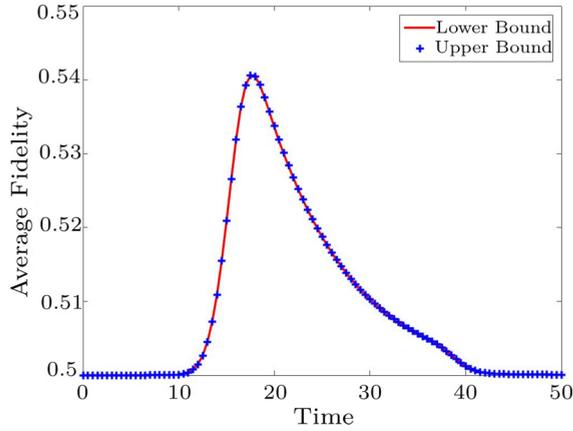
FIG. 2. Comparison of upper and lower bounds for average fidelity for a uniformly coupled chain of $N = 35$ and dephasing strength of $\Gamma_z = 0.3$.

Hamiltonian

$$H = \frac{1}{2} \sum_{n=1}^{N-1} J_n (X_n X_{n+1} + Y_n Y_{n+1}) - \frac{1}{2} \sum_{n=1}^{N} B_n Z_n.$$

We impose that $J_i \leqslant 1$ for all i in order to facilitate a fair comparison. There is a wide variety of different coupling strengths that one could consider. We have not made a detailed study of performance of the many different options. However, the $e^{-\Gamma_x t}$ factor that arises in amplitude damping noise strongly suggests that we must find solutions that are as fast as possible. We therefore consider three cases: the perfect state transfer chain [5], since it is the fastest such case [8,23]; the uniformly coupled chain [4]; and a tweaked version that is optimized for end-to-end transfer with high fidelity [9].

We consider two noise models: amplitude damping, as introduced in Sec. V C, and dephasing noise. For dephasing noise, the Lindblad operators are $L_i = \sqrt{\Gamma_z/2} Z_i$, where Γ_z is a measure of the strength of dephasing noise and $\Gamma_z = \frac{1}{T_2}$, where T_2 is the transverse paramagnetic relaxation time.

The results for the uniformly coupled chain are plotted in Fig. 3, comparing the effects of dephasing and amplitude damping noise. These plots, and indeed all others for different

Hamiltonians, consist of contours of constant fidelity that are straight lines. This is because the dominant noise term is of the form $e^{-(\Gamma_x + \Gamma_z)t}$ on a single excitation; we already saw in Sec. V C that this is the only contribution from amplitude damping. For dephasing noise, note that the noise terms are of the form

$$\mathcal{Q} - \mathcal{H} = \frac{\Gamma_z}{2} \sum_n Z_n \otimes Z_n - \frac{\Gamma_z}{2} \mathbb{1} \otimes \mathbb{1}.$$

This is diagonal. Of the N^2 diagonal elements in the single-excitation subspace, $N^2 - N$ of them are $-2\Gamma_z$, the exceptions being terms of the form $|\mathbf{nn}\rangle$, which are 0. Informally, then the dominant behavior will be similar to $-2\Gamma_z \mathbb{1}$, albeit with some correction due to the $\{|\mathbf{nn}\rangle\}$, the most relevant component of which is that the exponential decay tends towards the maximally mixed state of the subspace, instead of leaking out of the subspace (as was the case for amplitude damping). It is this to which we ascribe the reason for a given fidelity contour being realized by a marginally higher value of Γ_z than Γ_x.

Given that the dominant noise term is for the form $e^{-(\Gamma_x + \Gamma_z)t}$, it is clear that for a fixed transfer fidelity, we follow a contour of approximately constant $\Gamma_x + \Gamma_z$, which is what we observe. As such, it is not necessary to reproduce versions of Fig. 3 for multiple different encoding region sizes and Hamiltonian models. Instead, it is sufficient to refer back to Fig 1, which provides the equivalent plots with $\Gamma_z = 0$ fixed.

The state transfer time is not a significant parameter in these models; encoding is very good at producing wave packets that move at the group velocity of the chain. The leading effect on the transfer time as we increase the size of the encoding and decoding regions is simply a reduction in the transfer distance (transfer distance $N - 1 - 2M$) at the same transfer velocity (while also benefiting from lower dispersion due to the tighter wave packet and reduced noise impact due to the shorter transfer time).

VII. TOWARDS A MULTIPLE-EXCITATION ENCODING

So far, we have concentrated on encoding into the single-excitation subspace. This was partially motivated by the observation in [12] that the optimal encoding choice in the case of unitary evolution is in the single-excitation subspace
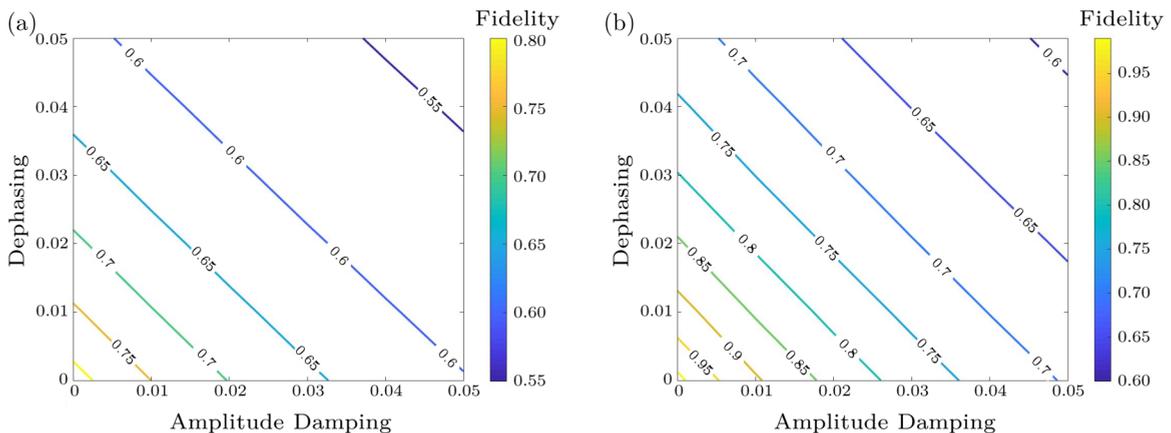
FIG. 3. Comparison of average fidelity achieved for the uniformly coupled Hamiltonian with varying noise parameters for $N = 35$: (a) no encoding and (b) seven-site encoding.

and that this outperforms any usage of error correction. Nevertheless, it is certainly possible that using the higher-excitation subspaces and error-correcting codes could enhance the protection against noise: Loss of excitations from the decoding region is our source of error, so if we had several excitations and could tolerate the loss of all but one, this is going to achieve a higher fidelity (although, exactly the same expectation can be applied, erroneously, to the unitary evolution case).

In fact, the formalism developed so far is readily generalized, being aware of the discrepancy indicated in [12] for the case of decoding when multiple excitations are present, compared to what Haselgrove originally claimed to be optimal [11]. Consider a decoding region of size $|\Lambda_{\text{out}}| = M$ and let us encode using the excitation subspaces up to and including the kth. We will restrict consideration to $k < M/2$. The purpose behind this assumption on the excitation number is that it lets us define U much as before: $U|\mathbf{0}\rangle = |\mathbf{0}\rangle$ and

$$U|x\rangle = |N, \phi_x\rangle$$

for any $x \in \{0, 1\}^M$ with weight $w_x \geqslant 1$ and $w_x \leqslant k$, where the $|\phi_x\rangle$ are orthonormal. Here we use x to describe the basis state of the M qubits in the decoding region. There are $\sum_{n=1}^{k} \binom{M}{n}$ of these and the states $|\phi_x\rangle$ only have support on $M - 1$ qubits, so the maximum number must be 2^{M-1}. Since $\sum_{n=1}^{M} \binom{M}{n} = 2^M - 1$, we need to pick k to restrict to less than half the possible total sum, i.e., $k < M/2$. For larger weights, we cannot achieve the orthonormal condition[2] and a further approximation would be necessary. We should also note that the exact solution for amplitude damping is no longer applicable.

Repeating the previous calculations, the $\langle 11|\mathcal{R}|\psi\psi^*\rangle$ term becomes otherwise independent of the choice of U and we recover a matrix similar to that used before,

$$R = \sum_{\substack{x,y\in\{0,1\}^{|\Lambda_{\text{in}}|}}} \sum_{\substack{z\in\{0,1\}^N \\ w_z \leqslant k \\ w_z^{\text{out}} \geqslant 1}} |y\rangle\langle x|\langle z, z|e^{\mathcal{Q}t}|x\mathbf{0}, y\mathbf{0}\rangle. \quad (12)$$

Here w_z^{out} indicates the weight of the bit string restricted only to the components in Λ_{out}. Similarly, optimization of the $\langle 10|\mathcal{R}|\psi\mathbf{0}\rangle$ term requires selection of the unitary such that

$$\langle \mathbf{0}|\phi_x\rangle = \frac{\gamma^*_{x\mathbf{0},\mathbf{00}}}{\sqrt{\sum_{z\in\{0,1\}^M}|\gamma_{z\mathbf{0},\mathbf{00}}|^2}},$$

where

$$\gamma_{x,y} = \langle x, y|e^{\mathcal{Q}t}|\psi\mathbf{0}\rangle.$$

This leads to a matrix

$$S = \sum_{\substack{x\in\{0,1\}^M \\ w_x \geqslant 1}} |x\rangle\langle x\mathbf{0}, \mathbf{00}|e^{\mathcal{Q}t} \sum_{\substack{z\in\{0,1\}^{|\Lambda_{\text{in}}|} \\ w_z \geqslant 1}} |z\mathbf{0}, \mathbf{00}\rangle\langle z|$$

$$= (P_{\text{out}} \otimes \langle \mathbf{0}|)e^{\mathcal{Q}t}(P_{\text{in}} \otimes |\mathbf{0}\rangle).$$

---

[2]In [12] we decoded onto a separate qubit, which removes this constraint.
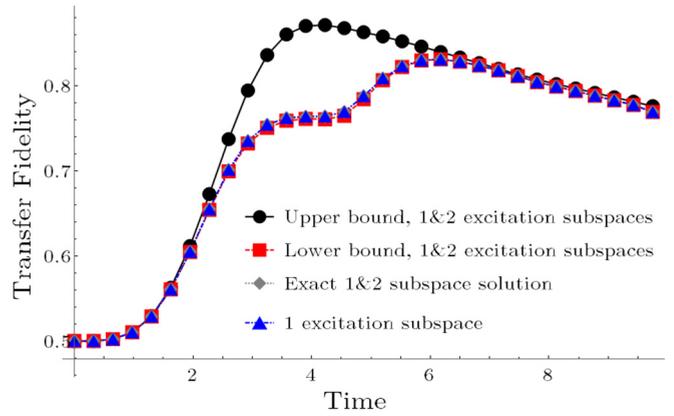
FIG. 4. Average fidelity over time for a uniformly coupled chain of $N = 13$ in the encoding and decoding region of five qubits. We compare the upper bound for fidelity in the two-excitation case with a lower bound and an exact solution. The exact solution case is hidden under the single-excitation case. There is no advantage to using higher-excitation subspaces. Here $\gamma_x = 0.02$ and $\gamma_z = 0.04$.

This time P_{in} and P_{out} are projectors onto the 1- to k-excitation subspaces on the input and output regions, respectively, and onto all other qubits being in the $|0\rangle$ state.

For all the cases we have been considering, S divides into a block-diagonal structure based on excitation number. For excitation preserving Q, this is trivial; $|z_{\text{in}}\mathbf{0}\rangle|\mathbf{0}\rangle \mapsto |x_{\text{out}}\mathbf{0}\rangle|\mathbf{0}\rangle$ only if x and z have the same weight. For amplitude damping noise, the excitation-decreasing terms are due to terms of the form $(X + iY) \otimes (X + iY)$. This requires that we are able to remove an excitation from both halves of $|\psi\rangle|\mathbf{0}\rangle$, which is clearly not possible; they have no effect and we revert to the excitation-preserving case.

The R has an identical subspace structure. To see this, consider the term $\langle zz|e^{\mathcal{Q}t}|x\mathbf{0}, y\mathbf{0}\rangle$ from Eq. (12). From our previous discussion we know that the Hamiltonian and noise will either preserve the number of excitations of x and y or decrease them by an equal number. However, since they must both end up having the same excitation number w_z, they must have started with the same excitation number. Note, however, that although both S and R are block diagonal, it does not necessarily mean that $|\psi\rangle$ is always supported on just one-excitation subspace; this is an effect of the square root in Eq. (5).

With both S and R in place, Eq. (5) still holds. Does the approximation in Eq. (11) remain useful? Does encoding in higher-excitation subspaces yield an improved fidelity?

In the noise-free case, so long as the maximum transfer amplitude in the single-excitation subspace is at least $\sqrt{2} - 1$, the optimal encoding is in the single-excitation subspace. By continuity, we expect this to remain true for weak noise as well.

We have not performed extensive studies of the multiple-excitation subspaces, which are computationally far more demanding. However, we have performed simple tests on more modestly sized systems and have never found an improvement by going to higher-excitation subspaces. In Fig. 4 we depict a typical case: a single noisy instance of a uniformly coupled chain. For reference, we display (blue triangles) the

behavior in the single-excitation subspace. With black circles we show an upper bound on the achievable fidelity based on separate optimization of $|\psi\rangle$ for the matrices R and S. With red squares we show the fidelity achieved due to the approximation of Eq. (11), a lower bound of what can be achieved. In higher-excitation subspaces, the coincidence of these two lines is clearly not as tight as it was in the first-excitation subspace. For such small cases, we can exactly find the optimal solution (gray diamonds). In all such cases we have tried, it has always been close to the lower bound that is the best achievable value.

VIII. CONCLUSION

We have a presented a simple to implement a scheme that can be applied to a wide range of Hamiltonians to improve fidelity in the presence of noise, by giving near-optimal encoding strategies. This is applicable to any Hamiltonian that is excitation preserving. We have demonstrated that larger encoding and decoding regions lead to better transfer in the presence of noise although reducing the overall transfer distance. Even modest sizes of encoding region can convert scenarios that are impossible for transfer into reasonable propositions. This is most compelling in the case of the uniform chain, which suggests there is little value in considering other state transfer systems with more complex, harder to implement, coupling schemes.

We have also shown that it is of benefit to choose a Hamiltonian that allows faster transfer over one that (in the absence of noise) produces higher-fidelity transfer. Our scheme leads to further improvement in these faster transfer chains. This technique can be applied to NISQ devices to allow some improvement in state transfer fidelity with respect to noise without implementing a full error correction scheme.

We have explicitly considered a specific form of Hamiltonian based on a spin chain. However, our derivation only assumed an excitation-preserving Hamiltonian and does not depend on any underlying coupling geometry. Similarly, a broad class of noise models can be handled. We have primarily focused on the single-excitation subspace, but have provided a route via which the formalism can continue to higher-excitation encodings, although numerics have failed to find any gain from doing so.

Here we focused on state transfer for the sake of having a concrete task to talk about. However, the formalism could easily be adapted to other Hamiltonian-based tasks [24,25], particularly those whose success is measured by fidelity. One might also be able to extend this work to cover other Hamiltonian models such as those that have different subspace structures [26,27] including those from the Jordan-Wigner transformation.

[1] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. **26**, 1484 (1997).

[2] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[3] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[4] S. Bose, Quantum Communication through an Unmodulated Spin Chain, Phys. Rev. Lett. **91**, 207901 (2003).

[5] M. Christandl, N. Datta, A. Ekert, and A. J. Landahl, Perfect State Transfer in Quantum Spin Networks, Phys. Rev. Lett. **92**, 187902 (2004).

[6] S. Bose, Quantum communication through spin chain dynamics: An introductory overview, Contemp. Phys. **48**, 13 (2007).

[7] A. Kay, A review of perfect state transfer and its application as a constructive tool, Int. J. Quantum Inf. **08**, 641 (2010).

[8] M.-H. Yung, Quantum speed limit for perfect state transfer in one dimension, Phys. Rev. A **74**, 030303(R) (2006).

[9] T. J. G. Apollaro, L. Banchi, A. Cuccoli, R. Vaia, and P. Verrucchi, 99%-fidelity ballistic quantum-state transfer through long uniform channels, Phys. Rev. A **85**, 052319 (2012).

[10] T. J. Osborne and N. Linden, Propagation of quantum information through a spin system, Phys. Rev. A **69**, 052315 (2004).

[11] H. L. Haselgrove, Optimal state encoding for quantum walks and quantum communication over spin systems, Phys. Rev. A **72**, 062326 (2005).

[12] C. Keele and A. Kay, preceding paper, Combating the effects of disorder in quantum state transfer, Phys. Rev. A **105**, 032612 (2022).

[13] A. Marais, I. Sinayskiy, A. Kay, F. Petruccione, and A. Ekert, Decoherence-assisted transport in quantum networks, New J. Phys. **15**, 013038 (2013).

[14] N. Behzadi, A. Ektesabi, and B. Ahansaz, Well-protected quantum state transfer in a dissipative spin chain, Sci. Rep. **8**, 7906 (2018).

[15] D. Burgarth and S. Bose, Universal destabilization and slowing of spin-transfer functions by a bath of spins, Phys. Rev. A **73**, 062321 (2006).

[16] C. Marletto, A. Kay, and A. Ekert, How to counteract systematic errors in quantum state transfer, Quantum Inf. Comput. **12**, 648 (2012).

[17] D. Burgarth and S. Bose, Perfect quantum state transfer with randomly coupled quantum chains, New J. Phys. **7**, 135 (2005).

[18] D. Burgarth, V. Giovannetti, and S. Bose, Efficient and perfect state transfer in quantum chains, J. Phys. A: Math. Gen. **38**, 6793 (2005).

[19] A. Kay, Perfect coding for dephased quantum state transfer, Phys. Rev. A **97**, 032317 (2018).

[20] A. Kay, Quantum error correction for state transfer in noisy spin chains, Phys. Rev. A **93**, 042320 (2016).

[21] A. Gilchrist, D. R. Terno, and C. J. Wood, Vectorization of quantum operations and its use, arXiv:0911.2539.

[22] B. Rost, B. Jones, M. Vyushkova, A. Ali, C. Cullip, A. Vyushkov, and J. Nabrzyski, Simulation of thermal relaxation in spin chemistry systems on a quantum computer using inherent qubit decoherence, arXiv:2001.00794.

[23] A. Kay, A note on the speed of perfect state transfer, arXiv:1609.01854.

[24] A. Kay, Coprocessors for quantum devices, Phys. Rev. A **97**, 032316 (2018).

[25] A. Kay, Tailoring spin chain dynamics for fractional revivals, Quantum **1**, 24 (2017).

[26] A. Kay, Unifying Quantum State Transfer and State Amplification, Phys. Rev. Lett. **98**, 010501 (2007).

[27] C. Di Franco, M. Paternostro, and M. S. Kim, Perfect State Transfer on a Spin Chain without State Initialization, Phys. Rev. Lett. **101**, 230502 (2008).