Time-dependent Hamiltonian simulation with L^1-norm scaling

Dominic W. Berry[1], Andrew M. Childs[2,3], Yuan Su[2,3], Xin Wang[3,4], and Nathan Wiebe[5,6,7]

[1]Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia
[2]Department of Computer Science, University of Maryland, College Park, MD 20742, USA
[3]Institute for Advanced Computer Studies and Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD 20742, USA
[4]Institute for Quantum Computing, Baidu Research, Beijing 100193, China
[5]Department of Physics, University of Washington, Seattle, WA 98195, USA
[6]Pacific Northwest National Laboratory, Richland, WA 99354, USA
[7]Google Inc., Venice, CA 90291, USA

The difficulty of simulating quantum dynamics depends on the norm of the Hamiltonian. When the Hamiltonian varies with time, the simulation complexity should only depend on this quantity instantaneously. We develop quantum simulation algorithms that exploit this intuition. For sparse Hamiltonian simulation, the gate complexity scales with the L^1 norm $\int_0^t \mathrm{d}\tau \, \|H(\tau)\|_{\max}$, whereas the best previous results scale with $t \max_{\tau\in[0,t]} \|H(\tau)\|_{\max}$. We also show analogous results for Hamiltonians that are linear combinations of unitaries. Our approaches thus provide an improvement over previous simulation algorithms that can be substantial when the Hamiltonian varies significantly. We introduce two new techniques: a classical sampler of time-dependent Hamiltonians and a rescaling principle for the Schrödinger equation. The rescaled Dyson-series algorithm is nearly optimal with respect to all parameters of interest, whereas the sampling-based approach is easier to realize for near-term simulation. These algorithms could potentially be applied to semi-classical simulations of scattering processes in quantum chemistry.

Contents

1 Introduction **3**

2 Preliminaries **5**
2.1 Time-dependent Hamiltonian evolution 5
2.2 Notation for norms 6
2.3 Hamiltonian input models 7
2.4 Simulation algorithms with L^1-norm scaling 9
2.5 Query complexity with L^1-norm scaling 10

3 Continuous qDRIFT **13**
3.1 A classical sampler of time-dependent Hamiltonians 14
3.2 Universality 19
3.3 Complexity of the continuous qDRIFT algorithm 23

4 Rescaled Dyson-series algorithm **24**
4.1 A rescaling principle for the Schrödinger equation 24
4.2 Complexity of the rescaled Dyson-series algorithm 25

5 Applications to chemistry and scattering theory **29**

6 Discussion **31**

A Continuous qDRIFT and Hamiltonian averaging **37**
A.1 Hamiltonian averaging 38
A.2 Implementing Hamiltonian averaging by continuous qDRIFT 39

1 Introduction

Simulating the Hamiltonian dynamics of a quantum system is one of the most promising applications of a quantum computer. The apparent classical intractability of simulating quantum dynamics led Feynman [25] and others to propose the idea of quantum computation. Quantum computers can simulate various physical systems, including condensed matter physics [3], quantum field theory [28], and quantum chemistry [2, 14, 37, 48]. The study of quantum simulation has also led to the discovery of new quantum algorithms, such as algorithms for linear systems [27], differential equations [10], semidefinite optimization [11], formula evaluation [23], quantum walk [17], and ground-state and thermal-state preparation [20, 41].

Let $H(\tau)$ be a Hamiltonian defined for $0 \leq \tau \leq t$. The problem of Hamiltonian simulation is to approximate the evolution $\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\ H(\tau)\right)$ using a quantum circuit comprised of elementary quantum gates, where $\exp_{\mathcal{T}}$ denotes the time-ordered matrix exponential. If the Hamiltonian $H(\tau) = H$ does not depend on time, the evolution operator can be represented in closed form as e^{-itH}. Then the problem can be greatly simplified and it has been thoroughly studied by previous works on quantum simulation [1, 4, 5, 7–9, 13, 16, 18, 19, 29, 31–33, 35, 36].

Simulating a general time-dependent Hamiltonian $H(\tau)$ naturally subsumes the time-independent case, and can be applied to devising quantum control schemes [38, 40], describing quantum chemical reactions [12], and implementing adiabatic quantum algorithms [22]. However, the problem becomes considerably harder and there are fewer quantum algorithms available. Wiebe, Berry, Høyer, and Sanders designed a time-dependent Hamiltonian simulation algorithm based on higher-order product formulas [50]. They assume that $H(\tau)$ is smooth up to a certain order and they give an example in which a desired approximation cannot be achieved due to the non-differentiability of the Hamiltonian. The smoothness assumption is relaxed in subsequent work by Poulin, Qarry, Somma, and Verstraete [42] based on techniques of Hamiltonian averaging and Monte Carlo estimation. The fractional-query algorithm of Berry, Childs, Cleve, Kothari, and Somma can also simulate time-dependent Hamiltonians [6], with an exponentially improved dependence on precision and only logarithmic dependence on the derivative of the Hamiltonian. A related quantum algorithm for time-dependent Hamiltonian simulation was suggested by Berry, Childs, Cleve, Kothari, and Somma based on the truncated Dyson series [8], which is analyzed explicitly in [29, 36].

In this paper, we study time-dependent Hamiltonian simulation based on a simple intuition: the difficulty of simulating a quantum system should depend on the integrated norm of the Hamiltonian. To elaborate, first consider the special case of simulating a time-independent Hamiltonian. The complexity of such a simulation depends on $t\,\|H\|$ [15], where $\|\cdot\|$ is a matrix norm that quantifies the size of the Hamiltonian. It is common to express the complexity in terms of the spectral norm $\|H\|_\infty$ (i.e., the Schatten ∞-norm), which quantifies the maximum energy of H.

In the general case where the Hamiltonian $H(\tau)$ is time dependent, we expect a quantum simulation algorithm to depend on the Hamiltonian locally in time, and therefore to have complexity that scales with the integrated norm $\int_0^t \mathrm{d}\tau\,\|H(\tau)\|$. This is the L^1 norm of $\|H(\tau)\|$ when viewed as a function of τ, so we say such an algorithm has L^1-norm scaling. Surprisingly, existing analyses of quantum simulation algorithms fail to achieve this complexity; rather, their gate complexity scales with the worst-case cost $t\max_{\tau\in[0,t]}\|H(\tau)\|$. It is therefore reasonable to question whether our intuition is correct, or if there exist faster

time-dependent Hamiltonian simulation algorithms that can exploit this intuition.[1]

Our work answers this question by providing multiple faster quantum algorithms for time-dependent Hamiltonian simulation. These algorithms have gate complexity that scales with $\int_0^t \mathrm{d}\tau \, \|H(\tau)\|$, in contrast to the best previous scaling of $t \max_{\tau\in[0,t]} \|H(\tau)\|$. As the norm inequality $\int_0^t \mathrm{d}\tau \, \|H(\tau)\| \leq t \max_{\tau\in[0,t]} \|H(\tau)\|$ always holds but is not saturated in general, these algorithms provide a strict speedup over existing algorithms. We further analyze an application to simulating scattering processes in quantum chemistry, showing that our improvement can be favorable in practice.

We introduce notation and terminology and state our assumptions in Section 2. Following standard assumptions about quantum simulation, we consider two different input models of Hamiltonians. The first is the sparse matrix (SM) model common for analyzing Hamiltonian simulation in general, in which the Hamiltonian is assumed to be sparse and access to the locations and values of nonzero matrix elements are provided by oracles. We quantify the complexity of a simulation algorithm by the number of queries and additional gates it uses. The second model, favorable for practical applications such as condensed matter physics and quantum chemistry simulation, assumes that the Hamiltonian can be explicitly decomposed as a linear combination of unitaries (LCU), where the coefficients are efficiently computable on a classical computer and the summands can be exponentiated and controlled on a quantum computer. We ignore the cost of implementing the coefficient oracle and focus mainly on the gate complexity. Quantum simulation algorithms can sometimes work more efficiently in other input models, but we study these two models since they are versatile and provide a fair comparison of the gate complexity.

Reference [6] claims that the fractional-query algorithm can simulate time-dependent Hamiltonians with L^∞-norm scaling. However, it is not hard to see that its query complexity in fact scales with the L^1 norm. While we do not show how to achieve this scaling for the gate complexity, our analysis is simple and suggests that such a result might be possible. We analyze the query complexity of the fractional-query algorithm in Section 2.5.

We develop two new techniques to simulate time-dependent Hamiltonians with L^1-norm scaling. Our first technique is a classical sampling protocol for time-dependent Hamiltonians. In this protocol, we randomly sample a time $\tau \in [0, t]$ and evolve under the time-independent Hamiltonian $H(\tau)$, where the probability distribution is designed to favor those τ with large $\|H(\tau)\|$. Campbell introduced a discrete sampling scheme for time-independent Hamiltonian simulation [13] and our protocol can be viewed as its continuous analog, which we call "continuous qDRIFT". We show that continuous qDRIFT is universal, in the sense that any Hamiltonian simulable by [13] can be simulated by continuous qDRIFT with the same complexity. In addition, we shave off a multiplicative factor in the analysis of [13] by explicitly evaluating the rate of change of the evolution with respect to scaling the Hamiltonian. Continuous qDRIFT and its analysis are detailed in Section 3. Our algorithm is also similar in spirit to the approach of Poulin et al. [43] based on Hamiltonian averaging and Monte Carlo estimation, although their algorithm does not have L^1-norm scaling. We discuss the relationship between these two approaches in Appendix A.

We also present a general principle for rescaling the Schrödinger equation in Section 4. In the rescaled Schrödinger equation, the time-dependent Hamiltonian $H(\tau)$ has the same norm for all $\tau \in [0, t]$, so the norm inequality $\int_0^t \mathrm{d}\tau \, \|H(\tau)\| \leq t \max_{\tau\in[0,t]} \|H(\tau)\|$ holds

[1] For the Dyson-series approach, Low and Wiebe claimed that the worst-case scaling may be avoided by a proper segmentation of the time interval [36]. However, it is unclear how their analysis can be formalized to give an algorithm with complexity that scales with the L^1 norm. In Section 4, we propose a rescaling principle for the Schrödinger equation and develop a rescaled Dyson-series algorithm with L^1-norm scaling.

with equality. Using this principle, we show that the simulation algorithm based on the truncated Dyson series [8, 29, 36] can also be improved to have L^1-norm scaling.

To illustrate how our results might be applied, we identify a specific problem in quantum chemistry for which our L^1-norm improvement is advantageous: semi-classical scattering of molecules in a chemical reaction. For such a simulation, $\|H(\tau)\|$ changes dramatically throughout the evolution, so its L^1 norm can be significantly smaller than its L^∞ norm. We discuss this application further in Section 5.

Finally, we conclude in Section 6 with a brief discussion of the results and some open questions.

2 Preliminaries

2.1 Time-dependent Hamiltonian evolution

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$. By default, we assume that $H(\tau)$ is continuously differentiable and $H(\tau) \neq 0$ everywhere, and we defer the discussion of pathological cases to Section 6. If the Hamiltonian $H(\tau) = H$ is time independent, the evolution is given in closed form by the matrix exponential e^{-itH}. However, there exists no such closed-form expression for a general $H(\tau)$ and we instead represent the evolution by $\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right)$, where $\exp_{\mathcal{T}}$ denotes the time-ordered matrix exponential. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right) = -iH(t)\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right). \tag{1}$$

If $G(\tau)$ is another time-dependent Hamiltonian, the evolutions generated by $H(\tau)$ and $G(\tau)$ have distance bounded by the following lemma.

Lemma 1 (L^1-norm distance bound of time-ordered evolutions [45, Appendix B])**.** *Let $H(\tau)$ and $G(\tau)$ be time-dependent Hamiltonians defined on the interval $0 \leq \tau \leq t$. Then,*

$$\left\|\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right) - \exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, G(\tau)\right)\right\|_\infty \leq \int_0^t \mathrm{d}\tau\, \|H(\tau) - G(\tau)\|_\infty . \tag{2}$$

Here, $\|\cdot\|_\infty$ denotes the spectral norm.

We will abbreviate the evolution operator as $\mathrm{E}(t,s) := \exp_{\mathcal{T}}\left(-i\int_s^t \mathrm{d}\tau\, H(\tau)\right)$ when there is no ambiguity. In the special case where $H(\tau) = H$ is time independent, the evolution e^{-itH} only depends on the time duration so we denote $\mathrm{E}(t) := \mathrm{E}(t,0)$. Therefore, we have the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(t,0) = -iH(t)\mathrm{E}(t,0), \qquad \mathrm{E}(0,0) = I. \tag{3}$$

We may further obtain an integral representation of $\mathrm{E}(t,0)$. To this end, we apply the fundamental theorem of calculus to the Schrödinger equation and obtain

$$\mathrm{E}(t,0) - I = \mathrm{E}(t,0) - \mathrm{E}(0,0) = \int_0^t \mathrm{d}\tau\, \frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{E}(\tau,0) = -i\int_0^t \mathrm{d}\tau\, H(\tau)\mathrm{E}(\tau,0). \tag{4}$$

Equivalently, $\mathrm{E}(t,0)$ satisfies the integral equation

$$\mathrm{E}(t,0) = I - i\int_0^t \mathrm{d}\tau\, H(\tau)\mathrm{E}(\tau,0). \tag{5}$$

For any $0 \leq s \leq t$, the evolution operator satisfies the multiplicative property

$$\mathrm{E}(t,0) = \mathrm{E}(t,s)\mathrm{E}(s,0). \tag{6}$$

The operator $\mathrm{E}(0,t)$ with $t \geq 0$ is understood as the inverse evolution operator

$$\mathrm{E}(0,t) := \mathrm{E}^{-1}(t,0) = \mathrm{E}^{\dagger}(t,0). \tag{7}$$

For a thorough mathematical treatment of time-dependent Hamiltonian evolution, we refer the reader to [21]. Finally, the quantum channel corresponding to the unitary evolution $\mathrm{E}(t,0)$ is denoted as $\mathcal{E}(t,0)$ and is defined by

$$\mathcal{E}(t,0)(\rho) := \mathrm{E}(t,0)\rho\mathrm{E}^{\dagger}(t,0) = \mathrm{E}(t,0)\rho\mathrm{E}(0,t). \tag{8}$$

For time-independent Hamiltonians, we denote $\mathcal{E}(t) := \mathcal{E}(t,0)$.

2.2 Notation for norms

We introduce norm notation for vectors, matrices, operator-valued functions, and linear maps on the space of matrices.

Let $\alpha = \begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_L\end{bmatrix} \in \mathbb{C}^L$ be an L-dimensional vector. We use $\|\alpha\|_p$ to represent the vector ℓ_p norm of α. Thus,

$$\|\alpha\|_1 := \sum_{j=1}^{L} |\alpha_j|, \qquad \|\alpha\|_2 := \sqrt{\sum_{j=1}^{L} |\alpha_j|^2}, \qquad \|\alpha\|_\infty := \max_{j\in\{1,2,\ldots,L\}} |\alpha_j|. \tag{9}$$

For a matrix A, we define $\|A\|_p$ to be the Schatten p-norm of A [47, 51]. We have

$$\|A\|_1 := \mathrm{Tr}(\sqrt{A^\dagger A}), \qquad \|A\|_2 := \sqrt{\mathrm{Tr}(A^\dagger A)}, \qquad \|A\|_\infty := \max_{|\psi\rangle} \|A|\psi\rangle\|_2 . \tag{10}$$

Finally, if $f\colon [0,t] \to \mathbb{C}$ is a continuous function, we use $\|f\|_p$ to mean the L^p norm of the function f. Thus,

$$\|f\|_1 := \int_0^t \mathrm{d}\tau\, |f(\tau)|, \qquad \|f\|_2 := \sqrt{\int_0^t \mathrm{d}\tau\, |f(\tau)|^2}, \qquad \|f\|_\infty := \max_{\tau\in[0,t]} |f(\tau)|. \tag{11}$$

We combine these norms to obtain norms for vector-valued and operator-valued functions. Let $\alpha\colon [0,t] \to \mathbb{C}^L$ be a continuous vector-valued function, with the jth coordinate at time τ denoted $\alpha_j(\tau)$. We use $\|\alpha\|_{p,q}$ to mean that we take the ℓ_p norm $\|\alpha(\tau)\|_p$ for every τ and compute the L^q norm of the resulting scalar function. For example,

$$\|\alpha\|_{1,1} := \int_0^t \mathrm{d}\tau \sum_{j=1}^{L} |\alpha_j(\tau)|, \qquad \|\alpha\|_{1,\infty} := \max_{\tau\in[0,t]} \sum_{j=1}^{L} |\alpha_j(\tau)|. \tag{12}$$

Note that $\|\alpha(\tau)\|_p$ is continuous as a function of τ, so $\|\alpha\|_{p,q}$ is well defined and is indeed a norm for vector-valued functions. Similarly, we also define $\|A\|_{p,q}$ for a continuous operator-valued function by taking the Schatten p-norm $\|A(\tau)\|_p$ for every τ and computing the L^q norm of the resulting scalar function. In rare cases, we will also encounter time-dependent linear combinations of operators of the form $A(\tau) = \sum_{l=1}^{L} A_l(\tau)$, and we write $\|A\|_{p,q,r}$ to

mean that we take the Schatten p-norm $\|A_l(\tau)\|_p$ of each summand, and apply the ℓ_q norm and L^r norm to the resulting vector-valued functions. For example,

$$\|A\|_{2,1,\infty} := \max_{\tau\in[0,t]} \sum_{l=1}^{L} \|A_l(\tau)\|_2 \,. \tag{13}$$

We also define $\|A\|_{\max}$ as the largest matrix element of A in absolute value,

$$\|A\|_{\max} := \max_{j,k} |A_{j,k}|. \tag{14}$$

The norm $\|A\|_{\max}$ is a vector norm of A but does not satisfy the submultiplicative property of a matrix norm. It relates to the spectral norm by the inequality [15, Lemma 1]

$$\|A\|_{\max} \leq \|A\|_{\infty} \,. \tag{15}$$

If A is a continuous operator-valued function, we interpret $\|A\|_{\max,q}$ in a similar way as above. Therefore,

$$\|A\|_{\max,1} := \int_0^t \mathrm{d}\tau \, \|A(\tau)\|_{\max} \,, \qquad \|A\|_{\max,\infty} := \max_{\tau\in[0,t]} \|A(\tau)\|_{\max} \,. \tag{16}$$

Finally, we define a norm for linear maps on the space of matrices. Let $\mathcal{E}\colon A \mapsto \mathcal{E}(A)$ be a linear map on the space of matrices on $\mathcal{H}$. The diamond norm of $\mathcal{E}$ is

$$\|\mathcal{E}\|_{\diamond} := \max\{\|(\mathcal{E}\otimes \mathbb{1}_{\mathcal{H}})(B)\|_1 : \|B\|_1 \leq 1\}, \tag{17}$$

where the maximization is taken over all matrices B on $\mathcal{H}\otimes\mathcal{H}$ satisfying $\|B\|_1 \leq 1$. Below is a useful bound on the diamond-norm distance between two unitary channels.

Lemma 2 (Diamond-norm distance between unitary channels [9, Lemma 7])**.** *Let V and U be unitary matrices, with associated quantum channels $\mathcal{V} : \rho \mapsto V\rho V^\dagger$ and $\mathcal{U} : \rho \mapsto U\rho U^\dagger$. Then,*

$$\|\mathcal{U} - \mathcal{V}\|_{\diamond} \leq 2\,\|U - V\|_{\infty} \,. \tag{18}$$

The sampling-based algorithm (Section 3) produces a channel close to $\mathcal{E}(t,0)(\rho) = \exp_{\mathcal{T}}\big(-i\int_0^t \mathrm{d}\tau H(\tau)\big)\rho \exp^{\dagger}_{\mathcal{T}}\big(-i\int_0^t \mathrm{d}\tau H(\tau)\big)$, and its error is naturally quantified by the diamond-norm distance. Other simulation algorithms such as the Dyson-series approach (Section 4) produce operators that are close to the unitary $\exp_{\mathcal{T}}\big(-i\int_0^t \mathrm{d}\tau H(\tau)\big)$, and we quantify their error in terms of the spectral norm. For a fair comparison one may instead describe all simulation algorithms using quantum channels and use the diamond-norm distance as the unified error metric. By Lemma 2, we lose at most a factor of 2 in this conversion.

2.3 Hamiltonian input models

Quantum simulation algorithms may have different performance depending on the choice of the input model of Hamiltonians. In this section, we describe two input models that are extensively used in previous works: the sparse matrix (SM) model and the linear-combination-of-unitaries (LCU) model. We also discuss other input models that will be used in later sections.

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$. In the SM model, we assume that $H(\tau)$ is d-sparse in the sense that the number of nonzero matrix elements

within each row and column throughout the entire interval $[0, t]$ is at most d. We assume that the locations of the nonzero matrix elements are time independent. Access to the Hamiltonian is given through the oracles

$$\begin{aligned} \mathcal{O}_{\mathrm{loc}}|j, s\rangle &= |j, \mathrm{col}(j, s)\rangle, \\ \mathcal{O}_{\mathrm{val}}|\tau, j, k, z\rangle &= |\tau, j, k, z \oplus H_{jk}(\tau)\rangle. \end{aligned} \tag{19}$$

Here, $\mathrm{col}(j, s)$ returns the column index of the sth element in the jth row that may be nonzero over the entire time interval $[0, t]$. We quantify the complexity of a quantum simulation algorithm by the number of queries it makes to $\mathcal{O}_{\mathrm{loc}}$ and $\mathcal{O}_{\mathrm{val}}$, together with the number of additional elementary gates it requires. Such a model includes many realistic physical systems and is well-motivated from a theoretical perspective [27].

As the following lemma shows, a d-sparse Hamiltonian can be efficiently decomposed as a sum of 1-sparse terms.

Lemma 3 (Decomposition of sparse Hamiltonians [6, Lemma 4.3 and 4.4])**.** *Let H be a time-independent d-sparse Hamiltonian accessed through the oracles $\mathcal{O}_{loc}$ and $\mathcal{O}_{val}$. Then*

1. *there exists a decomposition $H = \sum_{j=1}^{d^2} H_j$, where each H_j is 1-sparse with $\|H_j\|_{\max} \leq \|H\|_{\max}$, and a query to any H_j can be simulated with $O(1)$ queries to H; and*
2. *for any $\gamma > 0$, there exists an approximate decomposition[2] $\|H - \gamma \sum_{j=1}^{\eta} G_j\|_{\max} \leq \sqrt{2}\gamma$, where $\eta = O(d^2 \|H\|_{\max} / \gamma)$, each G_j is 1-sparse with eigenvalues ± 1, and a query to any G_j can be simulated with $O(1)$ queries to H.*

For the LCU model, we suppose that the Hamiltonian $H(\tau)$ admits a decomposition

$$H(\tau) = \sum_{l=1}^{L} \alpha_l(\tau) H_l, \tag{20}$$

where the coefficients $\alpha_l(\tau) \geq 0$ are continuously differentiable and nonzero everywhere, and the matrices H_l are both unitary and Hermitian. We assume that the coefficients $\alpha_l(\tau)$ can be efficiently computed by a classical oracle $\mathcal{O}_{\mathrm{coeff}}$, and we ignore the classical cost of implementing such an oracle. We further assume that each $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes H_l$ can be implemented with gate complexity g_c, and each $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-i\tau H_l}$ for an arbitrarily large τ can be performed with g_e gates. Such a setting is common in the simulation of condensed matter physics and quantum chemistry. We quantify the complexity of a simulation algorithm by the number of elementary gates it uses.

Quantum simulation algorithms can sometimes work in other input models. For example, the qDRIFT protocol introduced in Section 3 requires only that the Hamiltonians have the form

$$H(\tau) = \sum_{l=1}^{L} H_l(\tau), \tag{21}$$

where the Hermitian-valued functions $H_l(\tau)$ are continuous, nonzero everywhere, and can be efficiently exponentiated on a quantum computer. We call this the linear combination

[2] Reference [6] uses [6, Lemma 4.3] and the triangle inequality to show that $\left\|H - \gamma \sum_{j=1}^{\eta} G_j\right\|_{\max} \leq \sqrt{2}\gamma d^2$. However, this bound can be tightened to $\sqrt{2}\gamma$, since the max-norm distance depends on the largest error from rounding off the d^2 1-sparse matrices.

Algorithms	SM	LCU
Monte Carlo estimation (first step) [42]	$\widetilde{O}((d^2 \\|H\\|_{\max,\infty} t)^2 n/\epsilon)$	$O((\\|\alpha\\|_{1,\infty} t)^2 g_e/\epsilon)$
Fractional-query [6]	$\widetilde{O}(d^2 \\|H\\|_{\max,\infty} tn)$	N/A
Dyson series [8, 29, 36]	$\widetilde{O}(d \\|H\\|_{\max,\infty} tn)$	$\widetilde{O}(\\|\alpha\\|_{\infty,\infty} tL^2 g_c)$
Continuous qDRIFT (Section 3.3)	$\widetilde{O}((d^2 \\|H\\|_{\max,1})^2 n/\epsilon)$	$O(\\|\alpha\\|_{1,1}^2 g_e/\epsilon)$
Rescaled Dyson series (Section 4.2)	$\widetilde{O}(d \\|H\\|_{\max,1} n)$	$\widetilde{O}(\\|\alpha\\|_{\infty,1} L^2 g_c)$

Table 1: Complexity comparison of previous algorithms (top three) and the algorithms introduced in this paper (bottom two) for simulating time-dependent Hamiltonians. Logarithmic factors are suppressed by $\widetilde{O}$ notation and the (non-query) gate complexities are compared. The product formula algorithm of [50] is omitted as its gate complexity scales polynomially with high-order derivatives and is not directly comparable to other algorithms in the table. The complexity of the full Monte Carlo estimation algorithm [42] is not analyzed explicitly; only its first step is compared. The fractional-query algorithm [6] does not have an explicit implementation for Hamiltonians in the LCU model, and its implementation in the SM model is streamlined by the Dyson-series approach [8, 29, 36].

(LC) model. On the other hand, the Dyson-series algorithm can be described in terms of the SELECT operation

$$\textsc{Select}(H) := \sum_{l=1}^{L} |l\rangle\langle l| \otimes H_l, \tag{22}$$

irrespective of how this operation is implemented. We consider the SM and LCU models for all the time-dependent simulation algorithms so that we can give a fair comparison of their complexity.

2.4 Simulation algorithms with L^1-norm scaling

We now explain the meaning of L^1-norm scaling in the SM and the LCU models. Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$. We say that an algorithm in the SM model simulates $H(\tau)$ with L^1-norm scaling if, given any continuously differentiable upper bound $\Lambda_{\max}(\tau) \ge \\|H(\tau)\\|_{\max}$, the algorithm has query complexity and gate complexity that scale with $\\|\Lambda_{\max}\\|_1 = \int_0^t \mathrm{d}\tau\, \Lambda_{\max}(\tau)$ up to logarithmic factors. The norm bound $\Lambda_{\max}(\tau)$, together with other auxiliary information, must be accessed by the quantum simulation algorithm; we assume such quantities can be computed efficiently.

In the LCU model, we are given a time-dependent Hamiltonian with the decomposition $H(\tau) = \sum_{l=1}^{L} \alpha_l(\tau) H_l$. We say that an algorithm has L^1-norm scaling if, for any continuously differentiable vector-valued function $\Lambda(t)$ with $\Lambda_l(\tau) \ge \alpha_l(\tau)$, the algorithm has query and gate complexity that scale with $\\|\Lambda\\|_{\infty,1} = \int_0^t \mathrm{d}\tau \max_l \Lambda_l(\tau)$ up to logarithmic factors.

For better readability, we express the complexity of simulation algorithms in terms of the norm of the original Hamiltonian, such as $\\|H\\|_{\max,1}$ and $\\|\alpha\\|_{\infty,1}$, instead of the upper bounds $\\|\Lambda_{\max}\\|_1$ and $\\|\Lambda\\|_{\infty,1}$. We use standard asymptotic notation, with O, Ω, and Θ representing asymptotic upper, lower, and tight bounds, respectively. We also suppress logarithmic factors using the $\widetilde{O}$ notation when the complexity expression becomes too complicated. Table 1 compares the results of this paper with previous results on simulating time-dependent Hamiltonians.

Our goal is to develop simulation algorithms that scale with the L^1-norm with respect

to the time variable τ, for both query complexity and gate complexity. We start by reexamining the fractional-query approach. It was mentioned in [6] that this approach can simulate time-dependent Hamiltonians with L^∞-norm scaling, but we find that its query complexity scales with the L^1 norm. We give this improved analysis in the next section.

2.5 Query complexity with L^1-norm scaling

We begin by reviewing the result of [6] for simulating time-independent Hamiltonians. We assume that the Hamiltonian is given by a linear combination of unitaries $G = \sum_{l=1}^{L} \beta_l G_l$ with nonnegative coefficients β_l. Here, G_l are both unitary and Hermitian, so they are reflections and their eigenvalues are ± 1.

We say that a quantum operation is a fractional-query algorithm if it is of the form

$$U_m Q^{\tau_m} U_{m-1} \cdots U_1 Q^{\tau_1} U_0, \tag{23}$$

where Q is unitary with eigenvalues ± 1, U_j are unitary operations, and $\tau_j \geq 0$. Here, we regard Q as the oracle and U_j as non-query operations, so this algorithm has fractional-query complexity $\sum_{j=1}^{m} \tau_j$. A quantum algorithm that makes (discrete) queries to Q is a fractional-query algorithm with $\tau_j = 1$. Conversely, any fractional-query algorithm can be efficiently simulated in the discrete query model. In particular, an algorithm with fractional-query complexity T can be simulated with error at most ϵ using $O\big(T \frac{\log(T/\epsilon)}{\log\log(T/\epsilon)}\big)$ discrete queries [6, Lemma 3.8].

To apply the fractional-query approach, we approximate the evolution under G using the first-order product formula

$$\left\| e^{-itG} - \left(e^{-i\frac{t}{r}\beta_1 G_1} \cdots e^{-i\frac{t}{r}\beta_L G_L} \right)^r \right\|_\infty = O\left(\frac{(\|\beta\|_1 t)^2}{r} \right). \tag{24}$$

Observe that $e^{-i\pi G_l}$ are unitary operations with eigenvalues ± 1, so $\big(e^{-i\frac{t}{r}\beta_1 G_1} \cdots e^{-i\frac{t}{r}\beta_L G_L}\big)^r$ can be viewed as a fractional-query algorithm with query complexity $O(\|\beta\|_1 t)$, provided that we can make fractional queries to multiple oracles $e^{-i\pi G_1}, \ldots, e^{-i\pi G_L}$. This can be realized by a standard fractional-query algorithm accessing the single oracle

$$\textsc{Select}(\textsc{Exp}\text{-}G) = \sum_{l=1}^{L} |l\rangle\langle l| \otimes e^{-i\pi G_l} \tag{25}$$

with the same query complexity [6, Theorem 4.1].

To simulate with accuracy ϵ, we set $r = O((\|\beta\|_1 t)^2/\epsilon)$ to ensure that

$$\left\| e^{-itG} - \left(e^{-i\frac{t}{r}\beta_1 G_1} \cdots e^{-i\frac{t}{r}\beta_L G_L} \right)^r \right\|_\infty = O(\epsilon). \tag{26}$$

We now convert this multi-oracle algorithm to a single-oracle algorithm with the same fractional-query complexity $T = O(\|\beta\|_1 t)$ and, with precision $O(\epsilon)$, implement it in the discrete query model. Altogether, this approach makes

$$O\left(T \frac{\log(T/\epsilon)}{\log\log(T/\epsilon)} \right) = O\left(\|\beta\|_1 t \frac{\log(\|\beta\|_1 t/\epsilon)}{\log\log(\|\beta\|_1 t/\epsilon)} \right) \tag{27}$$

queries to the operation $\textsc{Select}(\textsc{Exp}\text{-}G) = \sum_{l=1}^{L} |l\rangle\langle l| \otimes e^{-i\pi G_l}$.

As mentioned in [6], the fractional-query approach can also be used to simulate time-dependent Hamiltonians by replacing (24) with a product-formula decomposition of the

time-ordered evolution. However, [6] only gives a brief discussion of this issue and the claimed complexity has only L^∞ scaling. We now give an improved analysis of this algorithm for the SM model, showing that its query complexity achieves L^1-norm scaling.

Theorem 4 (Fractional-query algorithm with L^1-norm scaling (SM))**.** *A d-sparse time-dependent Hamiltonian $H(\tau)$ acting on n qubits can be simulated for time $\tau \in [0,t]$ with accuracy ϵ using*

$$O\left(d^2 \left\|H\right\|_{\max,1} \frac{\log(d^2 \left\|H\right\|_{\max,\infty} t/\epsilon)}{\log\log(d^2 \left\|H\right\|_{\max,\infty} t/\epsilon)}\right) \tag{28}$$

queries to the oracles $\mathcal{O}_{loc}$, $\mathcal{O}_{val}$.

Proof. For readability, we assume that $\left\|H\right\|_{\max,1}$, $\left\|H\right\|_{\max,\infty}$, and $\left\|H'\right\|_{\infty,\infty}$ are the norm upper bounds provided to the algorithm. We first decompose $\exp_{\mathcal{T}}(-i\int_0^t \mathrm{d}\tau\, H(\tau))$ into a product of evolutions of time-independent Hamiltonians $H(kt/r)$, each evolving for time t/r. By Lemma 1, we have

$$\begin{aligned}
\left\|\exp_{\mathcal{T}}\left(-i\int_{\frac{kt}{r}}^{\frac{(k+1)t}{r}} \mathrm{d}\tau\, H(\tau)\right) - e^{-i\frac{t}{r}H\left(\frac{kt}{r}\right)}\right\|_\infty &\le \int_{\frac{kt}{r}}^{\frac{(k+1)t}{r}} \mathrm{d}s \left\|H(s) - H\left(\frac{kt}{r}\right)\right\|_\infty \\
&\le \int_{\frac{kt}{r}}^{\frac{(k+1)t}{r}} \mathrm{d}s \left(s - \frac{kt}{r}\right) \left\|H'\right\|_{\infty,\infty} \\
&= \frac{\left\|H'\right\|_{\infty,\infty} t^2}{2r^2},
\end{aligned} \tag{29}$$

which implies

$$\left\|\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}s\; H(s)\right) - \prod_{k=0}^{r-1} e^{-i\frac{t}{r}H\left(\frac{kt}{r}\right)}\right\|_\infty \le \frac{\left\|H'\right\|_{\infty,\infty} t^2}{2r}. \tag{30}$$

To approximate with precision ϵ, it suffices to choose

$$r = O\left(\frac{\left\|H'\right\|_{\infty,\infty} t^2}{\epsilon}\right). \tag{31}$$

Note that we use O here because we can choose r to be the minimum integer satisfying (30), giving an upper bound on the number of steps that suffice to achieve error at most ϵ.

We then decompose the evolution under each time-independent sparse Hamiltonian $H(kt/r)$ for time t/r with precision $O(\epsilon/r)$. By Lemma 3, $H(kt/r)$ can be decomposed into a sum of $\eta = O(d^2 \left\|H(kt/r)\right\|_{\max}/\gamma)$ terms $G_j(kt/r)$ such that

$$\left\|H(kt/r) - \gamma\sum_{j=1}^{\eta} G_j(kt/r)\right\|_{\max} \le \sqrt{2}\gamma. \tag{32}$$

Furthermore, each G_j is 1-sparse and Hermitian with eigenvalues ± 1 and the value and location of each non-zero matrix element in G_j can be accessed using $O(1)$ queries to H. We choose $\gamma = O(\epsilon/td)$ so that

$$\left\|e^{-i\frac{t}{r}H\left(\frac{kt}{r}\right)} - e^{-i\frac{t}{r}\gamma\sum_{j=1}^{\eta} G_j\left(\frac{kt}{r}\right)}\right\|_\infty = \left\|e^{-i\frac{t}{r}\left[H\left(\frac{kt}{r}\right) - \gamma\sum_{j=1}^{\eta} G_j\left(\frac{kt}{r}\right)\right]} - I\right\|_\infty$$

$$
\begin{aligned}
&\leq \frac{t}{r}\left\| H(kt/r) - \gamma \sum_{j=1}^{\eta} G_j(kt/r) \right\|_{\infty} \\
&\leq \frac{td}{r}\left\| H(kt/r) - \gamma \sum_{j=1}^{\eta} G_j(kt/r) \right\|_{\max} \\
&\leq \frac{td}{r}\sqrt{2}\gamma = O\left(\frac{td}{r}\frac{\epsilon}{td}\right) = O\left(\frac{\epsilon}{r}\right).
\end{aligned}
$$

In the third line we have used the inequality between the spectral norm and max norm, in the fourth line we have used the bound on the max norm (32), and in the fifth line we have used $\gamma = O(\epsilon/td)$. This implies $\eta = O(d^3 \left\|H(kt/r)\right\|_{\max} t/\epsilon)$ and the fractional query complexity is

$$
\eta \frac{t}{r}\gamma = O\left(\frac{d^2 \left\|H(kt/r)\right\|_{\max} t}{r}\right). \tag{33}
$$

We apply the first-order product formula to obtain

$$
\left\| e^{-i\frac{t}{r}\gamma \sum_{j=1}^{\eta} G_j\left(\frac{kt}{r}\right)} - e^{-i\frac{t}{r}\gamma G_1\left(\frac{kt}{r}\right)} \cdots e^{-i\frac{t}{r}\gamma G_\eta\left(\frac{kt}{r}\right)} \right\|_{\infty} = O\left(\frac{(d^2 \left\|H(kt/r)\right\|_{\max} t)^2}{r^2}\right). \tag{34}
$$

Therefore, it is possible to choose r as

$$
r = O\left(\frac{(d^2 \left\|H(kt/r)\right\|_{\max} t)^2}{\epsilon}\right) = O\left(\frac{(d^2 \left\|H\right\|_{\max,\infty} t)^2}{\epsilon}\right), \tag{35}
$$

such that the error of the first-order product-formula decomposition is at most

$$
\left\| e^{-i\frac{t}{r}\gamma \sum_{j=1}^{\eta} G_j\left(\frac{kt}{r}\right)} - e^{-i\frac{t}{r}\gamma G_1\left(\frac{kt}{r}\right)} \cdots e^{-i\frac{t}{r}\gamma G_\eta\left(\frac{kt}{r}\right)} \right\|_{\infty} = O\left(\frac{\epsilon}{r}\right). \tag{36}
$$

By choosing r as the maximum of (31) and (35), we ensure that the error in each of the r time steps is $O(\epsilon/r)$, so the total error is $O(\epsilon)$.

Altogether, we find a fractional-query algorithm with total query complexity

$$
T = O\left(\sum_{k=0}^{r-1} \frac{d^2 \left\|H(kt/r)\right\|_{\max} t}{r}\right) \tag{37}
$$

and error

$$
\left\| \exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}s\ H(s)\right) - \prod_{k=0}^{r-1}\left(e^{-i\frac{t}{rs}\gamma G_1\left(\frac{kt}{r}\right)} \cdots e^{-i\frac{t}{rs}\gamma G_\eta\left(\frac{kt}{r}\right)}\right)^s \right\|_{\infty} \leq \epsilon. \tag{38}
$$

We now convert this multi-oracle algorithm to a single-oracle algorithm with the same fractional-query complexity and, with precision $O(\epsilon)$, implement it in the discrete query model. The single oracle for the standard fractional-query algorithm is now

$$
\textsc{Select}(\textsc{Exp}\text{-}G) = \sum_{k=0}^{r-1} \sum_{l=0}^{\eta(k)-1} |k\rangle\langle k| \otimes |l\rangle\langle l| \otimes e^{-i\pi G_l(kt/r)}. \tag{39}
$$

This oracle encodes the time-dependence of H in an ancilla. The operators U_j in the fractional-query algorithm then need to increment the time register.

Altogether, we make

$$O\Big(T\frac{\log(T/\epsilon)}{\log\log(T/\epsilon)}\Big) = O\Big(\sum_{k=0}^{r-1}\frac{d^2\left\|H(kt/r)\right\|_{\max}t}{r}\frac{\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}{\log\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}\Big) \tag{40}$$

discrete queries.

We now show how the query complexity of this approach achieves L^1-norm scaling. The intuition is that the total query complexity $\widetilde{O}(\sum_{k=0}^{r-1} d^2\left\|H(kt/r)\right\|_{\max}t/r)$ should be close to $\widetilde{O}(d^2\int_0^t \mathrm{d}\tau\left\|H(\tau)\right\|_{\max})$ when r is sufficiently large. Specifically,

$$\begin{aligned}\Big|\sum_{k=0}^{r-1}\Big\|H\Big(\frac{kt}{r}\Big)\Big\|_{\max}\frac{t}{r} - \int_0^t \mathrm{d}\tau\left\|H(\tau)\right\|_{\max}\Big| &\le \sum_{k=0}^{r-1}\int_{\frac{kt}{r}}^{\frac{(k+1)t}{r}}\mathrm{d}\tau\Big|\left\|H(\tau)\right\|_{\max} - \Big\|H\Big(\frac{kt}{r}\Big)\Big\|_{\max}\Big| \\ &\le \sum_{k=0}^{r-1}\int_{\frac{kt}{r}}^{\frac{(k+1)t}{r}}\mathrm{d}\tau\Big\|H(\tau) - H\Big(\frac{kt}{r}\Big)\Big\|_{\max} \\ &\le \frac{\left\|H'\right\|_{\max,\infty}t^2}{2r}.\end{aligned} \tag{41}$$

To achieve an additive error of δ, it suffices to choose

$$r = O\Big(\frac{\left\|H'\right\|_{\max,\infty}t^2}{\delta}\Big). \tag{42}$$

Since δ can be made arbitrarily close to 0, we have the total query complexity of

$$\begin{aligned}&O\Big(d^2\Big(\int_0^t \mathrm{d}\tau\left\|H(\tau)\right\|_{\max} + \delta\Big)\frac{\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}{\log\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}\Big) \\ &= O\Big(d^2\left\|H\right\|_{\max,1}\frac{\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}{\log\log(d^2\left\|H\right\|_{\max,\infty}t/\epsilon)}\Big)\end{aligned} \tag{43}$$

as claimed. □

The above analysis shows that the fractional-query algorithm can simulate a time-dependent Hamiltonian with query complexity that scales with the L^1-norm. However, this approach does not directly give a useful result for the gate complexity. The difficulty arises from the factor of g in [6, Proof of Theorem 1.1], which corresponds to the complexity of applying a sequence of driving operations U_j. These operations need to increment k (indexing the time), as well as l, which takes $\eta(k)$ values depending on k. Applying the sequence of operations U_j therefore requires determining new values of l and k, which can depend on the sum of $\eta(k)$ over a long sequence of values of k. This will introduce significant gate complexity, so a fast algorithm would require a more efficient procedure for implementing the driving operations.

Instead, we develop other quantum algorithms that achieve L^1-norm scaling for not only the query complexity but also the gate complexity. We employ two main techniques: the continuous qDRIFT sampling protocol (Section 3) and a rescaling principle for the Schrödinger equation (Section 4).

3 Continuous qDRIFT

We show in Section 3.2 that continuous qDRIFT is universal, in the sense that any time-independent Hamiltonian simulable by the algorithm of [13] can be simulated by our protocol. We then discuss the simulation complexity in both the SM and the LCU models in Section 3.3.

The continuous qDRIFT protocol also has similarities with the approach of Poulin et al. [43] based on Hamiltonian averaging and Monte Carlo sampling, although their approach does not have L^1-norm scaling. We give a detailed comparison between these two approaches in Appendix A.

3.1 A classical sampler of time-dependent Hamiltonians

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$. For this section only, we relax our requirements on the Hamiltonians: we assume that $H(\tau)$ is nonzero everywhere and is continuous except on a finite number of points. We further suppose that each $H(\tau)$ can be directly exponentiated on a quantum computer. The ideal evolution under $H(\tau)$ for time t is given by $\mathrm{E}(t,0) = \exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right)$ and the corresponding quantum channel is

$$\mathcal{E}(t,0)(\rho) = \mathrm{E}(t,0)\rho\mathrm{E}^\dagger(t,0) = \exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right)\rho\exp^\dagger_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right). \tag{44}$$

The high-level idea of the sampling algorithm is to approximate the ideal channel by a mixed unitary channel

$$\mathcal{U}(t,0)(\rho) = \int_0^t \mathrm{d}\tau\, p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}}\rho e^{i\frac{H(\tau)}{p(\tau)}}, \tag{45}$$

where $p(\tau)$ is a probability density function defined for $0 \leq \tau \leq t$. This channel can be realized by a classical sampling protocol. With a proper choice of $p(\tau)$, this channel approximates the ideal channel and can thus be used for quantum simulation.

We begin with a full definition of $\mathcal{U}(t,0)$. Inspired by [13], we choose $p(\tau)$ to be biased toward those τ with large $\|H(\tau)\|_\infty$. A natural choice is

$$p(\tau) := \frac{\|H(\tau)\|_\infty}{\|H\|_{\infty,1}}. \tag{46}$$

Note that $\mathcal{U}(t,0)$ is a valid quantum channel (in particular, $p(\tau)$ can never be zero). Furthermore, it can be implemented with unit cost: for any input state ρ, we randomly sample a value τ according to $p(\tau)$ and perform $e^{-iH(\tau)/p(\tau)}$. Note also that $H(\tau)/p(\tau)$ in the exponential implicitly depends on t. Indeed, $\|H\|_{\infty,1}$ includes an integral over time, so $p(\tau)$ decreases with the total evolution time t. We call this classical sampling protocol and the channel it implements "continuous qDRIFT".

This protocol assumes that the spectral norm $\|H(\tau)\|_\infty$ is known a priori and that we can efficiently sample from the distribution $p(\tau)$. In practice, it is often easier to obtain a spectral-norm upper bound $\Lambda(\tau) \geq \|H(\tau)\|_\infty$. Such an upper bound can also be used to implement continuous qDRIFT, provided that it has only finitely many discontinuities. Specifically, we define

$$p_\Lambda(\tau) := \frac{\Lambda(\tau)}{\|\Lambda\|_1}, \tag{47}$$

so $p_\Lambda(\tau)$ is a probability density function. Using p_Λ to implement continuous qDRIFT, we obtain the channel

$$\mathcal{U}_\Lambda(t,0)(\rho) := \int_0^t \mathrm{d}\tau\, p_\Lambda(\tau) e^{-i\frac{H(\tau)}{p_\Lambda(\tau)}}\rho e^{i\frac{H(\tau)}{p_\Lambda(\tau)}}, \tag{48}$$

whose analysis is similar to that presented here. For readability, we assume that we can efficiently sample from $p(\tau) = \|H(\tau)\|_\infty / \|H\|_{\infty,1}$ and we analyze $\mathcal{U}(t,0)$.

We show that continuous qDRIFT approximates the ideal channel with error that depends on the L^1-norm.

Theorem 5 (L^1-norm error bound for continuous qDRIFT, short-time version)**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$; assume it is continuous except on a finite number of points and nonzero everywhere. Define $\mathrm{E}(t,0) = \exp_{\mathcal{T}}(-i\int_0^t \mathrm{d}\tau H(\tau))$ and let $\mathcal{E}(t,0)(\cdot) = \mathrm{E}(t,0)(\cdot)\mathrm{E}^\dagger(t,0)$ be the corresponding quantum channel. Let $\mathcal{U}(t,0)$ be the continuous qDRIFT channel*

$$\mathcal{U}(t,0)(\rho) = \int_0^t \mathrm{d}\tau\, p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}} \rho e^{i\frac{H(\tau)}{p(\tau)}}, \tag{49}$$

where $p(\tau) = \|H(\tau)\|_\infty / \|H\|_{\infty,1}$. Then

$$\|\mathcal{E}(t,0) - \mathcal{U}(t,0)\|_\diamond \leq 4\|H\|_{\infty,1}^2 . \tag{50}$$

Note that this bound is only useful when t is small enough that the right-hand side of (50) is less than 1 (the norm $\|H\|_{\infty,1}$ involves an integral over t, so it increases with t).

To prove this theorem, we need a formula that computes the rate at which the evolution operator changes when the Hamiltonian is scaled. To illustrate the idea, consider the degenerate case where the Hamiltonian H is time independent. Then the evolution under H for time t is given by e^{-itH}. A direct calculation shows that

$$\frac{\mathrm{d}}{\mathrm{d}s} e^{-itsH} = -itHe^{-itsH}, \tag{51}$$

so the rate is $-itHe^{-itsH}$ in the time-independent case. This calculation becomes significantly more complicated for a time-dependent Hamiltonian. The following lemma gives an explicit formula for

$$\frac{\mathrm{d}}{\mathrm{d}s} \exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, sH(\tau)\right). \tag{52}$$

We sketch the proof of this formula for completeness, but refer the reader to [21, p. 35] for mathematical justifications that are beyond the scope of this paper.

Lemma 6 (Hamiltonian scaling)**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$ and assume it has finitely many discontinuities. Denote $\mathrm{E}_s(t,v) = \exp_{\mathcal{T}}(-i\int_v^t \mathrm{d}\tau\, sH(\tau))$. Then,*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) = \int_v^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)[-iH(\tau)]\mathrm{E}_s(\tau,v). \tag{53}$$

Proof sketch. We first consider the special case where $H(\tau)$ is continuous in τ. We invoke the variation-of-parameters formula [30, Theorem 4.9] to construct the claimed integral representation for $\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)$. To this end, we need to find a differential equation satisfied by $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)$ and the corresponding initial condition $\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)\big|_{t=v}$. We differentiate the Schrödinger equation $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_s(t,v) = -isH(t)\mathrm{E}_s(t,v)$ with respect to s to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) = -isH(t)\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) - iH(t)\mathrm{E}_s(t,v). \tag{54}$$

Invoking the variation-of-parameters formula, we find an integral representation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) &= \mathrm{E}_s(t,v)\cdot\left[\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)\Big|_{t=v}\right] + \mathrm{E}_s(t,v)\int_v^t \mathrm{d}\tau\, \mathrm{E}_s^\dagger(\tau,v)[-iH(\tau)]\mathrm{E}_s(\tau,v) \\ &= \mathrm{E}_s(t,v)\cdot\left[\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)\Big|_{t=v}\right] + \int_v^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)[-iH(\tau)]\mathrm{E}_s(\tau,v).\end{aligned} \tag{55}$$

It thus remains to find the initial condition $\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)\big|_{t=v}$.

We start from the Schrödinger equation $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_s(t,v) = -isH(t)\mathrm{E}_s(t,v)$ and apply the fundamental theorem of calculus with initial condition $\mathrm{E}_s(v,v) = I$, obtaining the integral representation

$$\mathrm{E}_s(t,v) = I - is\int_v^t \mathrm{d}\tau\, H(\tau)\mathrm{E}_s(\tau,v). \tag{56}$$

Differentiating this equation with respect to s gives

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) = -i\int_v^t \mathrm{d}\tau\, H(\tau)\mathrm{E}_s(t,v) - is\int_v^t \mathrm{d}\tau\, H(\tau)\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(\tau,v), \tag{57}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)\Big|_{t=v} = 0. \tag{58}$$

Combining (55) and (58) establishes the claimed integral representation for $\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v)$.

Now consider the case where $H(\tau)$ is piecewise continuous with one discontinuity at $t_1 \in [v,t]$. We use the multiplicative property to break the evolution at t_1, so that each subevolution is generated by a continuous Hamiltonian. We have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,v) &= \frac{\mathrm{d}}{\mathrm{d}s}\left[\mathrm{E}_s(t,t_1)\mathrm{E}_s(t_1,v)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,t_1)\cdot\mathrm{E}_s(t_1,v) + \mathrm{E}_s(t,t_1)\cdot\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t_1,v) \\
&= \int_{t_1}^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)\big[-iH(\tau)\big]\mathrm{E}_s(\tau,t_1)\cdot\mathrm{E}_s(t_1,v) \\
&\quad + \mathrm{E}_s(t,t_1)\cdot\int_0^{t_1} \mathrm{d}\tau\, \mathrm{E}_s(t_1,\tau)\big[-iH(\tau)\big]\mathrm{E}_s(\tau,v) \\
&= \int_{t_1}^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)\big[-iH(\tau)\big]\mathrm{E}_s(\tau,v) \\
&\quad + \int_v^{t_1} \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)\big[-iH(\tau)\big]\mathrm{E}_s(\tau,v) \\
&= \int_v^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)\big[-iH(\tau)\big]\mathrm{E}_s(\tau,v).
\end{aligned} \tag{59}$$

The general case of finitely many discontinuities follows by induction. □

Note that our argument implicitly assumes the existence of the derivatives and that we can interchange the order of $\frac{\mathrm{d}}{\mathrm{d}s}$ and $\frac{\mathrm{d}}{\mathrm{d}t}$. A rigorous justification of these assumptions is beyond the scope of the paper; we refer the reader to [21, p. 35] for details.

Proof of Theorem 5. Define two parametrized quantum channels

$$\mathcal{E}_s(t,0)(\rho) = \mathrm{E}_s(t,0)\rho\mathrm{E}_s^\dagger(t,0), \qquad \mathcal{U}_s(t,0)(\rho) = \int_0^t \mathrm{d}\tau\, p(\tau)e^{-is\frac{H(\tau)}{p(\tau)}}\rho e^{is\frac{H(\tau)}{p(\tau)}} \tag{60}$$

and observe that

$$\mathcal{E}_0(t,0)(\rho) = \rho, \quad \mathcal{E}_1(t,0)(\rho) = \mathcal{E}(t,0)(\rho), \quad \mathcal{U}_0(t,0)(\rho) = \rho, \quad \mathcal{U}_1(t,0)(\rho) = \mathcal{U}(t,0)(\rho). \tag{61}$$

To bound the diamond-norm error $\|\mathcal{E}_1(t,0) - \mathcal{U}_1(t,0)\|_\diamond$, we should take a state σ on the joint system of the original register and an ancilla register with the same dimension and upper bound $\|(\mathcal{E}_1(t,0)\otimes\mathbb{1})(\sigma) - (\mathcal{U}_1(t,0)\otimes\mathbb{1})(\sigma)\|_1$. For readability, we instead show

how to bound the error $\|\mathcal{E}_1(t,0)(\rho)-\mathcal{U}_1(t,0)(\rho)\|_1$, but the derivation works in exactly the same way for the distance $\|(\mathcal{E}_1(t,0)\otimes\mathbb{1})(\sigma)-(\mathcal{U}_1(t,0)\otimes\mathbb{1})(\sigma)\|_1$ and the resulting bound is the same.

Invoking Lemma 6, we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_s(t,0)\Big|_{s=0} = \int_0^t \mathrm{d}\tau\, \mathrm{E}_s(t,\tau)\Big|_{s=0}[-iH(\tau)]\mathrm{E}_s(\tau,0)\Big|_{s=0} = -i\int_0^t \mathrm{d}\tau\, H(\tau). \tag{62}$$

Thus, the first derivatives of $\mathcal{E}_s(t,0)(\rho)$ and $\mathcal{U}_s(t,0)(\rho)$ at $s=0$ agree with each other:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_s(t,0)(\rho)\Big|_{s=0} &= \Big[-i\int_0^t \mathrm{d}\tau\, H(\tau),\rho\Big] \\ &= \int_0^t \mathrm{d}\tau\, p(\tau)\Big[-i\frac{H(\tau)}{p(\tau)},\rho\Big] = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{U}_s(t,0)(\rho)\Big|_{s=0}.\end{aligned} \tag{63}$$

Applying the fundamental theorem of calculus twice, we obtain

$$\begin{aligned}\mathcal{E}_1(t,0)(\rho)-\mathcal{U}_1(t,0)(\rho) &= (\mathcal{E}_1(t,0)(\rho)-\mathcal{E}_0(t,0)(\rho)) - (\mathcal{U}_1(t,0)(\rho)-\mathcal{U}_0(t,0)(\rho)) \\ &= \int_0^1 \mathrm{d}s\int_0^s \mathrm{d}v\, \frac{\mathrm{d}^2}{\mathrm{d}v^2}[\mathcal{E}_v(t,0)(\rho)-\mathcal{U}_v(t,0)(\rho)] \\ &= \int_0^1 \mathrm{d}s\int_0^s \mathrm{d}v\, \Big\{\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(t,0)\cdot\rho\cdot\mathrm{E}_v^\dagger(t,0) \\ &\quad + 2\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v(t,0)\cdot\rho\cdot\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v^\dagger(t,0) + \mathrm{E}_v(t,0)\cdot\rho\cdot\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v^\dagger(t,0) \\ &\quad - \int_0^t \mathrm{d}\tau\, p(\tau)e^{-iv\frac{H(\tau)}{p(\tau)}}\Big[-i\frac{H(\tau)}{p(\tau)},\Big[-i\frac{H(\tau)}{p(\tau)},\rho\Big]\Big]e^{iv\frac{H(\tau)}{p(\tau)}}\Big\}.\end{aligned} \tag{64}$$

By properties of the Schatten norms and the definition $p(\tau)=\|H(\tau)\|_\infty / \|H\|_{\infty,1}$, we find that

$$\begin{aligned}&\|\mathcal{E}_1(t,0)(\rho)-\mathcal{U}_1(t,0)(\rho)\|_1 \\ &\le \int_0^1 \mathrm{d}s\int_0^s \mathrm{d}v\, \Big\{2\Big\|\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(t,0)\Big\|_\infty + 2\Big\|\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v(t,0)\Big\|_\infty^2 + 4\|H\|_{\infty,1}^2\Big\}.\end{aligned} \tag{65}$$

Lemma 6 immediately yields an upper bound on $\left\|\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v(t,0)\right\|_\infty$:

$$\Big\|\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v(t,0)\Big\|_\infty \le \int_0^t \mathrm{d}\tau\, \|H(\tau)\|_\infty = \|H\|_{\infty,1}. \tag{66}$$

It thus remains to bound $\left\|\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(t,0)\right\|_\infty$.

Using Lemma 6 twice, we have

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(t,0) &= \frac{\mathrm{d}}{\mathrm{d}v}\int_0^t \mathrm{d}\tau\, \mathrm{E}_v(t,\tau)[-iH(\tau)]\mathrm{E}_v(\tau,0) \\ &= \int_0^t \mathrm{d}\tau\int_\tau^t \mathrm{d}\tau'\, \mathrm{E}_v(t,\tau')[-iH(\tau')]\mathrm{E}_v(\tau',\tau)[-iH(\tau)]\mathrm{E}_v(\tau,0) \\ &\quad + \int_0^t \mathrm{d}\tau\, \mathrm{E}_v(t,\tau)[-iH(\tau)]\int_0^\tau \mathrm{d}\tau'\, \mathrm{E}_v(\tau,\tau')[-iH(\tau')]\mathrm{E}_v(\tau',0),\end{aligned} \tag{67}$$

which implies

$$\begin{aligned}\Big\|\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(\tau,0)\Big\|_\infty &\le \int_0^t \mathrm{d}\tau\int_\tau^t \mathrm{d}\tau'\, \|H(\tau')\|_\infty \|H(\tau)\|_\infty + \int_0^t \mathrm{d}\tau\int_0^\tau \mathrm{d}\tau'\, \|H(\tau)\|_\infty \|H(\tau')\|_\infty \\ &= \|H\|_{\infty,1}^2.\end{aligned} \tag{68}$$

We finally obtain the desired bound

$$\|\mathcal{E}_1(t,0)(\rho)-\mathcal{U}_1(t,0)(\rho)\|_1 \le \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \left[2\|H\|_{\infty,1}^2+2\|H\|_{\infty,1}^2+4\|H\|_{\infty,1}^2\right] = 4\|H\|_{\infty,1}^2 \tag{69}$$

as claimed. □

The above error bound works well for a short-time evolution. When t is large, in order to control the error of simulation, we divide the entire evolution into segments $[t_j, t_{j+1}]$ with $0=t_0<t_1<\cdots<t_r=t$ and apply continuous qDRIFT within each. We employ a variable-time scheme to segment the evolution, so that our L^1-norm scaling result can be generalized to a long-time evolution. Specifically, we have:

Theorem 7 (L^1-norm error bound for continuous qDRIFT, long-time version)**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0\le\tau\le t$. Assume that it is continuous except at a finite number of points and nonzero everywhere. Define $\mathrm{E}(t,0)=\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right)$ and let $\mathcal{E}(t,0)(\cdot)=\mathrm{E}(t,0)(\cdot)\mathrm{E}^\dagger(t,0)$ be the corresponding quantum channel. Let $\mathcal{U}(t,0)$ be the continuous qDRIFT channel*

$$\mathcal{U}(t,0)(\rho)=\int_0^t \mathrm{d}\tau\, p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}}\rho e^{i\frac{H(\tau)}{p(\tau)}}, \tag{70}$$

where $p(\tau)=\|H(\tau)\|_\infty/\|H\|_{\infty,1}$. Then, for any positive integer r, there exists a division $0=t_0<t_1<\cdots<t_r=t$, such that

$$\left\|\mathcal{E}(t,0)-\prod_{j=0}^{r-1}\mathcal{U}(t_{j+1},t_j)\right\|_\diamond \le 4\frac{\|H\|_{\infty,1}^2}{r}. \tag{71}$$

To ensure that the simulation error is at most ϵ, it thus suffices to choose

$$r\ge 4\left\lceil\frac{\|H\|_{\infty,1}^2}{\epsilon}\right\rceil. \tag{72}$$

Proof. The times $t_1,\cdots,t_{r-1}$ are selected as follows. We aim to simulate with accuracy

$$4\frac{\|H\|_{\infty,1}^2}{r^2} \tag{73}$$

for each segment. To achieve this, we define $t_1,\cdots,t_{r-1}$ so that

$$\int_0^{t_1}\mathrm{d}\tau\,\|H(\tau)\|_\infty=\int_{t_1}^{t_2}\mathrm{d}\tau\,\|H(\tau)\|_\infty=\cdots=\int_{t_{r-1}}^{t_r}\mathrm{d}\tau\,\|H(\tau)\|_\infty=\frac{1}{r}\int_0^t\mathrm{d}\tau\,\|H(\tau)\|_\infty. \tag{74}$$

The existence of such times is guaranteed by the intermediate value theorem. By telescoping, we find from Theorem 5 that

$$\begin{aligned}\left\|\mathcal{E}(t,0)-\prod_{j=0}^{r-1}\mathcal{U}(t_{j+1},t_j)\right\|_\diamond &\le \sum_{j=0}^{r-1}\|\mathcal{U}(t_{j+1},t_j)-\mathcal{E}(t_{j+1},t_j)\|_\diamond\\ &\le \sum_{j=0}^{r-1}4\left(\int_{t_j}^{t_{j+1}}\mathrm{d}\tau\,\|H(\tau)\|_\infty\right)^2\\ &= 4r\left(\frac{1}{r}\int_0^t\mathrm{d}\tau\,\|H(\tau)\|_\infty\right)^2=4\frac{\|H\|_{\infty,1}^2}{r},\end{aligned} \tag{75}$$

which establishes the claimed error bound. □

3.2 Universality

We now show that the continuous qDRIFT method introduced above can be applied in the far more general LC model where the Hamiltonian is a sum of time-dependent terms. In this sense it can be regarded as a universal method.

Recall from Section 2.3 that in the general LC model, the Hamiltonian can be expressed as

$$H(\tau) = \sum_{l=1}^{L} H_l(\tau), \tag{76}$$

where each $H_l(\tau)$ is continuous and can be efficiently exponentiated on a quantum computer. This includes many familiar models as special cases:

(i) Campbell considered simulating a time-independent Hamiltonian of the form $H = \sum_{l=1}^{L} \alpha_l H_l$, $\|H_l\|_\infty \leq 1$ [13], which is subsumed by the LC model with the time dependence dropped;

(ii) if $H(\tau)$ is a time-dependent d-sparse Hamiltonian, then Lemma 3 shows that it can be decomposed in the form $H(\tau) = \sum_{j=1}^{d^2} H_j(\tau)$, which again belongs to the LC model as the exponentiation of $H_j(\tau)$ can be performed efficiently; and

(iii) the LC model is naturally more general than LCU as each summand is not necessarily unitary.

It is not hard to design a classical sampler for time-dependent Hamiltonians in the LC model. A natural choice is

$$\mathcal{U}(t,0)(\rho) := \sum_{l=1}^{L} \int_0^t \mathrm{d}\tau \, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}, \tag{77}$$

where $p_l(\tau)$ is the probability distribution

$$p_l(\tau) := \frac{\|H_l(\tau)\|_\infty}{\|H\|_{\infty,1,1}}. \tag{78}$$

To analyze the performance of this sampler, we adapt the analysis in Theorem 5 and Theorem 7, which becomes more complicated as we are now sampling a discrete-continuous probability distribution $p_l(\tau)$. Fortunately, a significant amount of effort can be saved with the help of the following universal property.

Theorem 8 (Universality of continuous qDRIFT)**.** *Let $H(\tau) = \sum_{l=1}^{L} H_l(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$ that is nonzero everywhere. Assume that each $H_l(\tau)$ is continuous and nonzero everywhere. Define the probability distribution*

$$p_l(\tau) := \frac{\|H_l(\tau)\|_\infty}{\|H\|_{\infty,1,1}}. \tag{79}$$

Then there exists a time-dependent Hamiltonian $G(\tau)$ defined for $0 \leq \tau \leq t$ with finitely many discontinuities, such that the following correspondence holds:

1. $\|G\|_{\infty,1} = \|H\|_{\infty,1,1}$.
2. $\int_0^t \mathrm{d}\tau \, G(\tau) = \sum_{l=1}^{L} \int_0^t \mathrm{d}\tau \, H_l(\tau)$.

3. $\int_0^t \mathrm{d}\tau\, q(\tau) e^{-i\frac{G(\tau)}{q(\tau)}} \rho e^{i\frac{G(\tau)}{q(\tau)}} = \sum_{l=1}^L \int_0^t \mathrm{d}\tau\, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}$, *where we have the probability distribution* $q(\tau) := \|G(\tau)\|_\infty / \|G\|_{\infty,1}$.

Before presenting the proof, we explain how Theorem 8 can be applied to simulation in the LC model. We expect that the mixed-unitary channel $\sum_{l=1}^L \int_0^t \mathrm{d}\tau\, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}$ approximates the ideal evolution with L^1-norm scaling as in Theorem 5 and Theorem 7, but direct analysis would be considerably more complicated. However, universality (Statement 3 of Theorem 8) shows that this channel is the same as $\int_0^t \mathrm{d}\tau\, q(\tau) e^{-i\frac{G(\tau)}{q(\tau)}} \rho e^{i\frac{G(\tau)}{q(\tau)}}$. Thus, the analysis of Section 3.1 can be applied with the help of Theorem 8.

Proof of Theorem 8. We define $G(\tau)$ to be the piecewise Hamiltonian

$$G(\tau) = \begin{cases} \frac{H_1\left(\frac{\tau}{p_1}\right)}{p_1}, & 0 \le \tau < p_1 t, \\ \frac{H_2\left(\frac{\tau - p_1 t}{p_2}\right)}{p_2}, & p_1 t \le \tau < (p_1+p_2)t, \\ \quad\vdots & \\ \frac{H_L\left(\frac{\tau-(p_1+p_2+\cdots+p_{L-1})t}{p_L}\right)}{p_L}, & (p_1+p_2+\cdots+p_{L-1})t \le \tau \le t, \end{cases} \tag{80}$$

where we use the abbreviation

$$p_l := \|p_l\|_1 = \int_0^t \mathrm{d}\tau\, p_l(\tau) \tag{81}$$

for the marginal probability distribution. Statements 1 and 2 can both be proved by directly evaluating the integrals

$$\begin{aligned} \|G\|_{\infty,1} &= \int_0^{p_1 t} \mathrm{d}\tau \frac{\left\|H_1\left(\frac{\tau}{p_1}\right)\right\|_\infty}{p_1} + \int_{p_1 t}^{(p_1+p_2)t} \mathrm{d}\tau \frac{\left\|H_2\left(\frac{\tau-p_1 t}{p_2}\right)\right\|_\infty}{p_2} \\ &\quad + \cdots + \int_{(p_1+p_2+\cdots+p_{L-1})t}^{t} \mathrm{d}\tau \frac{\left\|H_L\left(\frac{\tau-(p_1+p_2+\cdots+p_{L-1})t}{p_L}\right)\right\|_\infty}{p_L} \\ &= \int_0^t \mathrm{d}\tau \left\|H_1(\tau)\right\|_\infty + \int_0^t \mathrm{d}\tau \left\|H_2(\tau)\right\|_\infty + \cdots + \int_0^t \mathrm{d}\tau \left\|H_L(\tau)\right\|_\infty = \|H\|_{\infty,1,1} \end{aligned} \tag{82}$$

and

$$\begin{aligned} \int_0^t \mathrm{d}\tau\, G(\tau) &= \int_0^{p_1 t} \mathrm{d}\tau \frac{H_1\left(\frac{\tau}{p_1}\right)}{p_1} + \int_{p_1 t}^{(p_1+p_2)t} \mathrm{d}\tau \frac{H_2\left(\frac{\tau-p_1 t}{p_2}\right)}{p_2} \\ &\quad + \cdots + \int_{(p_1+p_2+\cdots+p_{L-1})t}^{t} \mathrm{d}\tau \frac{H_L\left(\frac{\tau-(p_1+p_2+\cdots+p_{L-1})t}{p_L}\right)}{p_L} = \sum_{l=1}^L \int_0^t \mathrm{d}\tau\, H_l(\tau). \end{aligned} \tag{83}$$

We use Statement 1 to deduce that

$$q(\tau) = \frac{\|G(\tau)\|_\infty}{\|G(\tau)\|_{\infty,1}} = \begin{cases} \frac{\left\|H_1\left(\frac{\tau}{p_1}\right)\right\|_\infty}{p_1 \|H\|_{\infty,1,1}}, & 0 \le \tau < p_1 t, \\ \frac{\left\|H_2\left(\frac{\tau-p_1 t}{p_2}\right)\right\|_\infty}{p_2 \|H\|_{\infty,1,1}}, & p_1 t \le \tau < (p_1+p_2)t, \\ \quad\vdots & \\ \frac{\left\|H_L\left(\frac{\tau-(p_1+p_2+\cdots+p_{L-1})t}{p_L}\right)\right\|_\infty}{p_L \|H\|_{\infty,1,1}}, & (p_1+p_2+\cdots+p_{L-1})t \le \tau \le t. \end{cases} \tag{84}$$

Therefore,

$$
\begin{aligned}
&\int_0^t \mathrm{d}\tau\, q(\tau) e^{-i\frac{G(\tau)}{q(\tau)}} \rho e^{i\frac{G(\tau)}{q(\tau)}} \\
&= \int_0^{p_1 t} \mathrm{d}\tau \frac{\left\| H_1(\frac{\tau}{p_1}) \right\|_\infty}{p_1 \left\| H \right\|_{\infty,1,1}} \exp\left(-i \frac{H_1(\frac{\tau}{p_1})}{\left\| H_1(\frac{\tau}{p_1}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \rho \exp\left(i \frac{H_1(\frac{\tau}{p_1})}{\left\| H_1(\frac{\tau}{p_1}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \\
&+ \int_{p_1 t}^{(p_1+p_2)t} \mathrm{d}\tau \frac{\left\| H_2(\frac{\tau - p_1 t}{p_2}) \right\|_\infty}{p_2 \left\| H \right\|_{\infty,1,1}} \exp\left(-i \frac{H_2(\frac{\tau - p_1 t}{p_2})}{\left\| H_2(\frac{\tau - p_1 t}{p_2}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \rho \exp\left(i \frac{H_2(\frac{\tau - p_1 t}{p_2})}{\left\| H_2(\frac{\tau - p_1 t}{p_2}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \\
&+ \cdots + \int_{(p_1+p_2+\cdots+p_{L-1})t}^{t} \mathrm{d}\tau \frac{\left\| H_L(\frac{\tau - (p_1+p_2+\cdots+p_{L-1})t}{p_L}) \right\|_\infty}{p_L \left\| H \right\|_{\infty,1,1}} \\
&\cdot \exp\left(-i \frac{H_L(\frac{\tau - (p_1+p_2+\cdots+p_{L-1})t}{p_L})}{\left\| H_L(\frac{\tau - (p_1+p_2+\cdots+p_{L-1})t}{p_L}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \rho \exp\left(i \frac{H_L(\frac{\tau - (p_1+p_2+\cdots+p_{L-1})t}{p_L})}{\left\| H_L(\frac{\tau - (p_1+p_2+\cdots+p_{L-1})t}{p_L}) \right\|_\infty} \left\| H \right\|_{\infty,1,1} \right) \\
&= \sum_{l=1}^{L} \int_0^t \mathrm{d}\tau\, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}, \qquad (85)
\end{aligned}
$$

which completes the proof of Statement 3. □

Theorem 5′ (L^1-norm error bound for continuous qDRIFT (LC), short-time version)**.** *Let $H(\tau) = \sum_{l=1}^{L} H_l(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ that is nonzero everywhere. Assume that each $H_l(\tau)$ is continuous and nonzero everywhere. Define $\mathrm{E}(t,0) = \exp_{\mathcal{T}}\left(-i \int_0^t \mathrm{d}\tau\, H(\tau) \right)$ and let $\mathcal{E}(t,0)(\cdot) = \mathrm{E}(t,0)(\cdot)\mathrm{E}^\dagger(t,0)$ be the corresponding quantum channel. Let $\mathcal{U}(t,0)$ be the continuous qDRIFT channel*

$$
\mathcal{U}(t,0)(\rho) := \sum_{l=1}^{L} \int_0^t \mathrm{d}\tau\, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}, \qquad (86)
$$

where $p_l(\tau)$ is the probability distribution $p_l(\tau) := \left\| H_l(\tau) \right\|_\infty / \left\| H \right\|_{\infty,1,1}$. Then,

$$
\left\| \mathcal{E}(t,0) - \mathcal{U}(t,0) \right\|_\diamond \le 4 \left\| H \right\|_{\infty,1,1}^2 . \qquad (87)
$$

In the special case where $H = \sum_{l=1}^{L} H_l$ is time independent, our bound reduces to

$$
\left\| \mathcal{E}(t,0) - \mathcal{U}(t,0) \right\|_\diamond \le 4 \Big(\sum_l \left\| H_l \right\|_\infty \Big)^2 t^2. \qquad (88)
$$

This tightens a bound due to Campbell [13, Eq. (B12)] by a multiplicative factor from a tail bound. Note that [13] considered the distance $\left\| \mathcal{E}(t,0) - \mathcal{U}(t,0) \right\|_\diamond / 2$, which is different from our definition of the diamond-norm distance $\left\| \mathcal{E}(t,0) - \mathcal{U}(t,0) \right\|_\diamond$.

Proof of Theorem 5′. Consider the channel

$$
\mathcal{G}(t,0)(\rho) := \int_0^t \mathrm{d}\tau\, q(\tau) e^{-i\frac{G(\tau)}{q(\tau)}} \rho e^{i\frac{G(\tau)}{q(\tau)}}, \qquad (89)
$$

where $q(\tau) := \left\| G(\tau) \right\|_\infty / \left\| G \right\|_{\infty,1}$ and $G(\tau)$ is defined by (80). By Theorem 8, it suffices to bound $\left\| \mathcal{E}(t,0) - \mathcal{G}(t,0) \right\|_\diamond$.

Define two parametrized quantum channels

$$\mathcal{E}_s(t,0)(\rho) = \mathrm{E}_s(t,0)\rho\mathrm{E}_s^\dagger(t,0), \qquad \mathcal{G}_s(t,0)(\rho) = \int_0^t \mathrm{d}\tau\, q(\tau) e^{-is\frac{G(\tau)}{q(\tau)}} \rho e^{is\frac{G(\tau)}{q(\tau)}} \tag{90}$$

and observe that

$$\mathcal{E}_0(t,0)(\rho) = \rho \qquad \mathcal{E}_1(t,0)(\rho) = \mathcal{E}(t,0)(\rho) \qquad \mathcal{G}_0(t,0)(\rho) = \rho \qquad \mathcal{G}_1(t,0)(\rho) = \mathcal{G}(t,0)(\rho). \tag{91}$$

For readability, we only consider the trace norm $\|\mathcal{E}_1(t,0)(\rho) - \mathcal{G}_1(t,0)(\rho)\|_1$, whose analysis can be easily adapted to bound $\|(\mathcal{E}_1(t,0)\otimes\mathbb{1})(\sigma) - (\mathcal{G}_1(t,0)\otimes\mathbb{1})(\sigma)\|_1$ and thus the diamond-norm distance $\|\mathcal{E}_1(t,0) - \mathcal{G}_1(t,0)\|_\diamond$.

By Lemma 6 and Theorem 8, we find that the first derivatives of $\mathcal{E}_s(t,0)(\rho)$ and $\mathcal{G}_s(t,0)(\rho)$ at $s=0$ agree with each other:

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_s(t,0)(\rho)\Big|_{s=0} = \Big[-i\int_0^t \mathrm{d}\tau\, H(\tau), \rho\Big] = \Big[-i\int_0^t \mathrm{d}\tau\, G(\tau), \rho\Big] = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{G}_s(t,0)(\rho)\Big|_{s=0}. \tag{92}$$

Thus, we can apply the fundamental theorem of calculus twice and obtain

$$\begin{aligned}
&\mathcal{E}_1(t,0)(\rho) - \mathcal{G}_1(t,0)(\rho) \\
&= (\mathcal{E}_1(t,0)(\rho) - \mathcal{E}_0(t,0)(\rho)) - (\mathcal{G}_1(t,0)(\rho) - \mathcal{G}_0(t,0)(\rho)) \\
&= \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v\, \frac{\mathrm{d}^2}{\mathrm{d}v^2}[\mathcal{E}_v(t,0)(\rho) - \mathcal{G}_v(t,0)(\rho)] \\
&= \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \Big\{ \frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v(t,0)\cdot\rho\cdot\mathrm{E}_v^\dagger(t,0) \\
&\qquad + 2\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v(t,0)\cdot\rho\cdot\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}_v^\dagger(t,0) \\
&\qquad + \mathrm{E}_v(t,0)\cdot\rho\cdot\frac{\mathrm{d}^2}{\mathrm{d}v^2}\mathrm{E}_v^\dagger(t,0) \\
&\qquad - \int_0^t \mathrm{d}\tau\, q(\tau) e^{-iv\frac{G(\tau)}{q(\tau)}}\Big[-i\frac{G(\tau)}{q(\tau)}, \Big[-i\frac{G(\tau)}{q(\tau)}, \rho\Big]\Big] e^{iv\frac{G(\tau)}{q(\tau)}} \Big\},
\end{aligned} \tag{93}$$

which implies

$$\begin{aligned}
&\|\mathcal{E}_1(t,0)(\rho) - \mathcal{G}_1(t,0)(\rho)\|_1 \\
&\le \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \left\{ 2\|H\|_{\infty,1,1}^2 + 2\|H\|_{\infty,1,1}^2 + 4\|G\|_{\infty,1}^2 \right\} = 4\|H\|_{\infty,1,1}^2 .
\end{aligned} \tag{94}$$

□

Theorem 7′ (L^1-norm error bound for continuous qDRIFT (LC), long-time version)**.** *Let $H(\tau) = \sum_{l=1}^L H_l(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ that is nonzero everywhere. Assume that each $H_l(\tau)$ is continuous and nonzero everywhere. Define $\mathrm{E}(t,0) = \exp_{\mathcal{T}}(-i\int_0^t \mathrm{d}\tau\, H(\tau))$ and let $\mathcal{E}(t,0)(\cdot) = \mathrm{E}(t,0)(\cdot)\mathrm{E}^\dagger(t,0)$ be the corresponding quantum channel. Let $\mathcal{U}(t,0)$ be the continuous qDRIFT channel*

$$\mathcal{U}(t,0)(\rho) := \sum_{l=1}^L \int_0^t \mathrm{d}\tau\, p_l(\tau) e^{-i\frac{H_l(\tau)}{p_l(\tau)}} \rho e^{i\frac{H_l(\tau)}{p_l(\tau)}}, \tag{95}$$

where $p_l(\tau)$ is the probability distribution $p_l(\tau) := \|H_l(\tau)\|_\infty / \|H\|_{\infty,1,1}$. Then, for any positive integer r, there exists a division $0 = t_0 < t_1 < \cdots < t_r = t$, such that

$$\left\| \mathcal{E}(t,0) - \prod_{j=0}^{r-1} \mathcal{U}(t_{j+1}, t_j) \right\|_\diamond \leq 4 \frac{\|H\|_{\infty,1,1}^2}{r}. \tag{96}$$

To ensure that the simulation error is at most ϵ, it thus suffices to choose

$$r \geq 4 \left\lceil \frac{\|H\|_{\infty,1,1}^2}{\epsilon} \right\rceil. \tag{97}$$

The proof of Theorem 7′ follows from Theorem 5′ using the same reasoning as that used to prove Theorem 7.

3.3 Complexity of the continuous qDRIFT algorithm

As an immediate consequence of universality, we obtain the complexity of the continuous qDRIFT algorithm for simulating time-dependent Hamiltonians in both the SM and the LCU models.

Corollary 9 (Continuous qDRIFT algorithm with L^1-norm scaling (SM))**.** *A d-sparse time-dependent Hamiltonian $H(\tau)$ acting on n qubits can be simulated for time $\tau \in [0,t]$ with accuracy ϵ using*

$$O\left(\frac{d^4 \|H\|_{\max,1}^2}{\epsilon} \right) \tag{98}$$

queries to $\mathcal{O}_{loc}$, $\mathcal{O}_{val}$ and an additional

$$\widetilde{O}\left(\frac{d^4 \|H\|_{\max,1}^2}{\epsilon} n \right) \tag{99}$$

gates, assuming that the probability distribution $p_j(\tau) := \|H(\tau)\|_{\max} / d^2 \|H\|_{\max,1}$, $j \in \{1, \ldots, d^2\}$ can be efficiently sampled.

Proof. For any $\tau \in [0,t]$, Lemma 3 shows that $H(\tau)$ admits a decomposition $H(\tau) = \sum_{j=1}^{d^2} H_j(\tau)$, where each $H_j(\tau)$ is 1-sparse and a query to any $H_j(\tau)$ can be simulated with $O(1)$ queries to $H(\tau)$. We use the continuous qDRIFT algorithm to simulate $H(\tau) = \sum_{j=1}^{d^2} H_j(\tau)$. We estimate

$$\begin{aligned} \|H\|_{\infty,1,1} &= \sum_{j=1}^{d^2} \int_0^t \mathrm{d}\tau \, \|H_j(\tau)\|_\infty = \sum_{j=1}^{d^2} \int_0^t \mathrm{d}\tau \, \|H_j(\tau)\|_{\max} \\ &\leq d^2 \int_0^t \mathrm{d}\tau \, \|H(\tau)\|_{\max} = d^2 \|H\|_{\max,1}, \end{aligned} \tag{100}$$

where the second equality follows because $H_j(\tau)$ is 1-sparse, and the inequality follows from Lemma 3. Assuming $\|H_j(\tau)\|_\infty / \|H\|_{\infty,1,1}$ can be sampled efficiently, Theorem 7′ implies that the algorithm has sample complexity and thus query complexity

$$O\left(\frac{d^4 \|H\|_{\max,1}^2}{\epsilon} \right). \tag{101}$$

For each elementary exponentiation, we initialize a quantum register in the computational basis state $|\tau, j\rangle$ and use it to control the 1-sparse term we need to simulate. This

can be done with gate complexity $\widetilde{O}(n)$. Since the number of 1-sparse simulations is the query complexity, we obtain the gate complexity

$$\widetilde{O}\Big(\frac{d^4 \|H\|_{\max,1}^2}{\epsilon} n\Big) \tag{102}$$

as claimed.

Our above argument assumes that $\|H_j(\tau)\|_\infty$ is known a priori and that the distribution $\|H_j(\tau)\|_\infty / \|H\|_{\infty,1,1}$ can be efficiently sampled. However, the argument still works if we replace each $\|H_j(\tau)\|_\infty$ by the upper bound

$$\|H_j(\tau)\|_\infty = \|H_j(\tau)\|_{\max} \leq \|H(\tau)\|_{\max}\,, \tag{103}$$

which means we sample the distribution $p_j(\tau) := \|H(\tau)\|_{\max} / d^2 \|H\|_{\max,1}$, $j \in \{1, \ldots, d^2\}$. The claimed query and gate complexities follow from a similar analysis. □

Corollary 9′ (Continuous qDRIFT algorithm with L^1-norm scaling (LCU))**.** *A time-dependent Hamiltonian with the LCU decomposition $H(\tau) = \sum_{l=1}^L \alpha_l(\tau) H_l$, where the controlled exponentiation of each H_l can be performed with g_e gates, can be simulated for time $\tau \in [0, t]$ with accuracy ϵ with gate complexity*

$$4\left\lceil \frac{\|\alpha\|_{1,1}^2}{\epsilon} \right\rceil g_e, \tag{104}$$

assuming that the probability distribution $p_l(\tau) := \alpha_l(\tau) / \|\alpha\|_{1,1}$ can be efficiently sampled.

Proof. For any $H(\tau) = \sum_{l=1}^L \alpha_l(\tau) H_l$, we estimate

$$\|H\|_{\infty,1,1} = \sum_{l=1}^L \int_0^t \mathrm{d}\tau\, \alpha_l(\tau)\, \|H_l\|_\infty = \|\alpha\|_{1,1}\,. \tag{105}$$

The claimed complexity then follows from Theorem 7′. □

4 Rescaled Dyson-series algorithm

In this section, we propose a general principle for rescaling the Schrödinger equation (Section 4.1). We then apply this principle to improve the Dyson-series algorithm (Section 4.2) to achieve L^1-norm scaling.

4.1 A rescaling principle for the Schrödinger equation

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$. The evolution under $H(\tau)$ for time t is given by the unitary operator $\mathrm{E}(t, 0) = \exp_{\mathcal{T}}(-i \int_0^t \mathrm{d}\tau\, H(\tau))$, which satisfies the Schrödinger equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(t, 0) = -iH(t)\mathrm{E}(t, 0). \tag{106}$$

We now propose a rescaling principle that helps to achieve L^1-norm scaling. The goal is to effectively have a Hamiltonian with constant spectral norm. Recall that for a time-independent Hamiltonian one can multiply the time by a constant and divide the Hamiltonian by the same constant and obtain the same time evolution. We can achieve

something similar with a time-dependent Hamiltonian by rescaling the total evolution time to

$$s = f(t) := \int_0^t \mathrm{d}\tau \, \|H(\tau)\|_\infty \tag{107}$$

and using the rescaled Hamiltonian

$$\widetilde{H}(\varsigma) := H(f^{-1}(\varsigma)) / \left\|H(f^{-1}(\varsigma))\right\|_\infty . \tag{108}$$

From this definition, it is obvious that the Hamiltonian has constant norm, because

$$\|\widetilde{H}(\varsigma)\|_\infty = \left\| \frac{H(f^{-1}(\varsigma))}{\|H(f^{-1}(\varsigma))\|_\infty} \right\|_\infty = 1. \tag{109}$$

Moreover, we find that the time-evolution operator satisfies

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}(t,0) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(t,0) \cdot \frac{\mathrm{d}t}{\mathrm{d}s} \\ &= -iH(t)\mathrm{E}(t,0) \cdot \frac{1}{\|H(t)\|_\infty} \\ &= -i\widetilde{H}(s)\mathrm{E}(t,0). \end{aligned} \tag{110}$$

Solving this equation shows that we can obtain exactly the same time-evolution operator using the rescaled time and Hamiltonian:

$$\mathrm{E}(t,0) = \exp_{\mathcal{T}}\left(-i \int_0^s \mathrm{d}\varsigma \, \widetilde{H}(\varsigma)\right). \tag{111}$$

We also have the norm equality

$$s \max_{\varsigma\in[0,s]} \|\widetilde{H}(\varsigma)\|_\infty = s = \|H\|_{\infty,1} , \tag{112}$$

so any algorithm that simulates the rescaled Hamiltonian $\widetilde{H}(\varsigma)$ with complexity that scales with the L^∞ norm can simulate the original Hamiltonian with L^1-norm scaling.

While our above discussion considers the spectral norm $\|\cdot\|_\infty$, other norms may be used depending on the input model of the Hamiltonian. Indeed, in the analysis for the SM model below we use the max-norm instead of the spectral norm.

Note that it may be hard in practice to compute the exact value of $\|H(\tau)\|$. However, we can instead use the change of variable

$$s = f(t) := \int_0^t \mathrm{d}\tau \, \Lambda(\tau), \tag{113}$$

where $\Lambda(\tau) \geq \|H(\tau)\|$ is any upper bound on the norm that can be efficiently computed.

4.2 Complexity of the rescaled Dyson-series algorithm

In this section, we show how the Dyson-series algorithm [8, 29, 36] can be rescaled to have L^1-norm scaling. We address this first for the SM model of Hamiltonian access before handling the LCU model (see Section 2.3 for definitions of these models).

Unlike continuous qDRIFT, the rescaled Dyson-series algorithm requires additional oracle access to the input Hamiltonian. Specifically, we need oracles that implement the inverse change-of-variable

$$\mathcal{O}_{\mathrm{var}}|\varsigma, z\rangle = |\varsigma, z \oplus f^{-1}(\varsigma)\rangle \tag{114}$$

and compute the max-norm

$$\mathcal{O}_{\text{norm}}|\tau, z\rangle = |\tau, z \oplus \|H(\tau)\|_{\max}\rangle. \tag{115}$$

A quantum computer with access to these oracles can simulate time-dependent Hamiltonians with L^1-norm scaling. Note that because $f(\tau)$ increases monotonically, we can use binary search to compute $f^{-1}(\varsigma)$ up to precision δ using $O(\log(t/\delta))$ queries to f, so we expect it to be straightforward to implement the oracle $\mathcal{O}_{\text{var}}$ in practice.

Theorem 10 (Rescaled Dyson-series algorithm with L^1-norm scaling (SM))**.** *For $\tau \in [0,t]$, let $H(\tau)$ be a d-sparse Hamiltonian acting on n qubits. Let $f(t) := \int_0^t \mathrm{d}\tau \, \|H(\tau)\|_{\max}$, and suppose we have an upper bound on the max-norm, denoted $\left\|H(f^{-1}(\varsigma))\right\|_{\max}$, that is positive and continuously differentiable. Then H can be simulated for time t with accuracy ϵ using*

$$O\left(d\,\|H\|_{\max,1}\,\frac{\log(d\,\|H\|_{\max,1}/\epsilon)}{\log\log(d\,\|H\|_{\max,1}/\epsilon)}\right) \tag{116}$$

queries to the oracles $\mathcal{O}_{loc}$, $\mathcal{O}_{val}$, $\mathcal{O}_{var}$, $\mathcal{O}_{norm}$ and an additional

$$\widetilde{O}(d\,\|H\|_{\max,1}\,n) \tag{117}$$

gates.

Proof. We simulate the rescaled Hamiltonian $\widetilde{H}(\varsigma) := H(f^{-1}(\varsigma))/\left\|H(f^{-1}(\varsigma))\right\|_{\max}$ for a total time of $\|H\|_{\max,1} := \int_0^t \mathrm{d}\tau \, \|H(\tau)\|_{\max}$ using the rescaling function

$$f(t) := \int_0^t \mathrm{d}\tau \, \|H(\tau)\|_{\max}\,. \tag{118}$$

Following [36, Theorem 9], we construct a unitary operation that block-encodes

$$\sum_{\varsigma\in[0,t/M,2t/M,\ldots,(M-1)t/M]} |\varsigma\rangle\langle\varsigma| \otimes \frac{\widetilde{H}(\varsigma)}{d\|\widetilde{H}\|_{\max,\infty}} = \sum_{\varsigma} |\varsigma\rangle\langle\varsigma| \otimes \frac{\widetilde{H}(\varsigma)}{d}. \tag{119}$$

This construction is similar to [36, Lemma 8], except that the Hamiltonian is rescaled. Specifically, we use oracles $\mathcal{O}_{\text{var}}$ and $\mathcal{O}_{\text{norm}}$ to implement the transformation

$$|\varsigma, 0, 0\rangle \mapsto |\varsigma, f^{-1}(\varsigma), \left\|H(f^{-1}(\varsigma))\right\|_{\max}\rangle, \tag{120}$$

from which we obtain the rescaled Hamiltonian by querying $\mathcal{O}_{\text{val}}$ and re-normalizing the result with $\left\|H(f^{-1}(\varsigma))\right\|_{\max}$ to compute $\widetilde{H}_{jk}(\varsigma)$:

$$|f^{-1}(\varsigma), \left\|H(f^{-1}(\varsigma))\right\|_{\max}, j, k, z\rangle \mapsto |f^{-1}(\varsigma), \left\|H(f^{-1}(\varsigma))\right\|_{\max}, j, k, z \oplus \widetilde{H}_{jk}(\varsigma)\rangle. \tag{121}$$

We then uncompute the ancilla registers storing $f^{-1}(\varsigma)$ and $\left\|H(f^{-1}(\varsigma))\right\|_{\max}$. Overall, this implements a rescaled oracle

$$\widetilde{\mathcal{O}}_{\text{val}}|\varsigma, j, k, z\rangle = |\varsigma, j, k, z \oplus \widetilde{H}_{jk}(\varsigma)\rangle. \tag{122}$$

The remaining algorithm proceeds as in [36]. As the implementation of each $\widetilde{\mathcal{O}}_{\text{val}}$ requires $O(1)$ queries to the oracles $\mathcal{O}_{\text{loc}}$, $\mathcal{O}_{\text{val}}$, $\mathcal{O}_{\text{var}}$, $\mathcal{O}_{\text{norm}}$, the overall query complexity is obtained by applying [36, Theorem 9] to the rescaled Hamiltonian, giving query complexity

$$O\left(T\frac{\log(T/\epsilon)}{\log\log(T/\epsilon)}\right) \tag{123}$$

where

$$T = d\|\widetilde{H}\|_{\max,\infty} s = d\,\|H\|_{\max,1}\,. \tag{124}$$

Using this expression for T gives the query complexity in (116).

We now analyze the gate complexity. If the entries of the Hamiltonian are given to within precision

$$O\left(\frac{\epsilon}{td}\right), \tag{125}$$

then the overall error due to the finite precision is $O(\epsilon)$. Since the maximum value of any matrix entry of H is $\|H\|_{\max,\infty}$, the number of bits required is

$$n_p \in \Theta\left(\log\left(\frac{d\|H\|_{\max,\infty} t}{\epsilon}\right)\right). \tag{126}$$

The implementation involves performing arithmetic on these values, which can be performed with complexity[3] $O(n_p^2)$. Since this is a logarithmic gate cost for each oracle query, it gives a contribution to the gate complexity of $\widetilde{O}(d\,\|H\|_{\max,1})$.

The number of time steps is [36, Corollary 4]

$$M \in \Theta\left(\frac{t}{\alpha\epsilon}\left(\frac{\|\widetilde{H}'\|_{\infty,1}}{t} + \|\widetilde{H}\|_{\infty,\infty}^2\right)\right) \tag{127}$$

where $\alpha = d\|\widetilde{H}\|_{\max,\infty}$. The complexity to prepare the time registers is $\log M$ times the query complexity. We may ignore the second term in (127), because it is negligible compared to the complexity of the arithmetic.

We have

$$\|\widetilde{H}'\|_{\infty,1} = \int_0^s \mathrm{d}\varsigma \left\|\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}\varsigma}\right\|_\infty. \tag{128}$$

Evaluating this derivative, we get

$$\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}\varsigma} = \frac{\mathrm{d}\tau}{\mathrm{d}\varsigma}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{H(\tau)}{\|H(\tau)\|_{\max}}\right) = \frac{\mathrm{d}\tau}{\mathrm{d}\varsigma}\left(\frac{H'(\tau)}{\|H(\tau)\|_{\max}} - \frac{H(\tau)\,\|H(\tau)\|'_{\max}}{\|H(\tau)\|_{\max}^2}\right), \tag{129}$$

so we obtain

$$\begin{aligned}
\|\widetilde{H}'\|_{\infty,1} &= \int_0^t \mathrm{d}\tau \left\|\frac{H'(\tau)}{\|H(\tau)\|_{\max}} - \frac{H(\tau)\,\|H(\tau)\|'_{\max}}{\|H(\tau)\|_{\max}^2}\right\|_\infty \\
&\leq \frac{\|H'(\tau)\|_{\infty,1}}{\min_{\tau\in[0,t]}\|H(\tau)\|_{\max}} + \frac{\left\|\|H(\tau)\|_\infty\,\|H(\tau)\|'_{\max}\right\|_1}{\min_{\tau\in[0,t]}\|H(\tau)\|_{\max}^2} \\
&\leq \frac{\|H'(\tau)\|_{\infty,1}}{\min_{\tau\in[0,t]}\|H(\tau)\|_{\max}} + \frac{\|H(\tau)\|_{\infty,2}\left\|\|H(\tau)\|'_{\max}\right\|_2}{\min_{\tau\in[0,t]}\|H(\tau)\|_{\max}^2}.
\end{aligned} \tag{130}$$

[3] In [29] and [36], the most complicated operations used are additions, which can be performed with complexity $O(n_p)$. Here we are normalizing the Hamiltonian, so we must also perform multiplication and/or division, for which the straightforward approach has complexity $O(n_p^2)$. While it is possible to perform multiplication and division with lower asymptotic complexity, such algorithms are only advantageous for very large instances, and do not affect the result as presented in Theorem 10, where logarithmic contributions to the gate complexity are suppressed.

The gate complexity of the preparation of the time registers is $\log M$ times the query complexity, where we have shown that

$$M \in \Theta\left(\frac{\|\widetilde{H}'\|_{\infty,1}}{\epsilon d \|H\|_{\max,\infty}}\right), \tag{131}$$

where $\|\widetilde{H}'\|_{\infty,1}$ is polynomial in norms of H and its derivative. Since this a logarithmic cost, the contribution to the complexity from preparation of the time registers is $\widetilde{O}(d\,\|H\|_{\max,1})$.

The remaining contribution to the gate complexity comes from acting on the system itself. The cost of this is $O(n)$ for each of the oracle queries, which gives gate complexity $\widetilde{O}(d\,\|H\|_{\max,1}\,n)$ (this is the dominant cost in (117)). □

Thus, the rescaled Dyson-series algorithm can simulate time-dependent Hamiltonians in the SM model with L^1-norm scaling. Next we turn our attention to the LCU model. For an input Hamiltonian $H(\tau) = \sum_{l=1}^{L} \alpha_l(\tau) H_l$, this approach assumes quantum access to the coefficient oracle

$$O_{\text{coeff}}|\tau, l, z\rangle = |\tau, l, z \oplus \alpha_l(\tau)\rangle, \tag{132}$$

in contrast to the continuous qDRIFT which only needs classical access. Given a classical circuit that computes the coefficients $\alpha_l(\tau)$, we can express it as a sequence of elementary gates and construct a corresponding quantum circuit with the same gate complexity. In our analysis, we ignore the implementation details and count the number of uses of the quantum oracle $\mathcal{O}_{\text{coeff}}$. The definitions of $\mathcal{O}_{\text{var}}$ and $\mathcal{O}_{\text{norm}}$ are similar to the SM case, except that the norm $\|H(\tau)\|_{\max}$ is replaced by $\|\alpha(\tau)\|_\infty$.

Theorem 10′ (Rescaled Dyson-series algorithm with L^1-norm scaling (LCU))**.** *For $\tau \in [0,t]$, let H be a time-dependent Hamiltonian with the decomposition $H(\tau) = \sum_{l=1}^{L} \alpha_l(\tau) H_l$, where each controlled H_l can be performed with g_c gates. Let $f(t) := \int_0^t \mathrm{d}\tau\, \|\alpha(\tau)\|_\infty$, and suppose we have an upper bound on the ℓ_∞ norm of the coefficients, denoted $\|\alpha(f^{-1}(\varsigma))\|_\infty$, that is continuously differentiable. Then H can be simulated for time t with accuracy ϵ using*

$$O\left(L\,\|\alpha\|_{\infty,1}\,\frac{\log(L\,\|\alpha\|_{\infty,1}/\epsilon)}{\log\log(L\,\|\alpha\|_{\infty,1}/\epsilon)}\right) \tag{133}$$

queries to the oracles $\mathcal{O}_{coeff}$, $\mathcal{O}_{var}$, $\mathcal{O}_{norm}$ and an additional

$$\widetilde{O}(\|\alpha\|_{\infty,1}\,L^2 g_c) \tag{134}$$

gates.

Proof. We simulate the rescaled Hamiltonian $\widetilde{H}(\varsigma) := H(f^{-1}(\varsigma))/\|\alpha(f^{-1}(\varsigma))\|_\infty$ for time $\|\alpha\|_{\infty,1}$ using the rescaling function

$$f(t) := \int_0^t \mathrm{d}\tau\, \|\alpha(\tau)\|_\infty . \tag{135}$$

The rescaled Hamiltonian takes the form

$$\widetilde{H}(\varsigma) = \sum_{l=1}^{L} \widetilde{\alpha}_l(f^{-1}(\varsigma)) H_l, \tag{136}$$

where $\widetilde{\alpha}_l(\tau) := \alpha_l(\tau)/\|\alpha(\tau)\|_\infty$. Therefore, we have a global upper bound on the absolute value of the coefficients

$$\|\widetilde{\alpha}\|_{\infty,\infty} \leq 1. \tag{137}$$

The remaining construction is similar to [29, Section V C], except that the Hamiltonian is rescaled. Specifically, we use oracles $\mathcal{O}_{\text{var}}$ and $\mathcal{O}_{\text{norm}}$ to implement the transformation

$$|\varsigma, 0, 0\rangle \mapsto |\varsigma, f^{-1}(\varsigma), \left\|\alpha(f^{-1}(\varsigma))\right\|_\infty\rangle, \tag{138}$$

from which we obtain the rescaled coefficients by querying $\mathcal{O}_{\text{coeff}}$ and doing arithmetic, giving

$$|f^{-1}(\varsigma), \left\|\alpha(f^{-1}(\varsigma))\right\|_\infty, l, z\rangle \mapsto |f^{-1}(\varsigma), \left\|\alpha(f^{-1}(\varsigma))\right\|_\infty, l, z \oplus \widetilde{\alpha}_l(\varsigma)\rangle. \tag{139}$$

We then uncompute the ancilla registers storing $f^{-1}(\varsigma)$ and $\|\alpha(f^{-1}(\varsigma))\|_\infty$. Overall, this implements a rescaled oracle

$$\widetilde{\mathcal{O}}_{\text{coeff}}|\varsigma, l, z\rangle = |\varsigma, l, z \oplus \widetilde{\alpha}_l(\varsigma)\rangle. \tag{140}$$

The remaining algorithm proceeds as in [29]. As the implementation of each $\widetilde{\mathcal{O}}_{\text{coeff}}$ requires $O(1)$ queries to the oracles $\mathcal{O}_{\text{coeff}}$, $\mathcal{O}_{\text{var}}$, $\mathcal{O}_{\text{norm}}$, the overall query complexity is obtained by applying [29, Theorem 2] to the rescaled Hamiltonian. The analysis of the gate complexity proceeds along similar lines to that of Theorem 10. The multiplicative factor of Lg_c is the cost of implementing the SELECT operation

$$\textsc{Select}(H) = \sum_{l=1}^{L} |l\rangle\langle l| \otimes H_l. \tag{141}$$

That complexity may be obtained from [18, Lemma G.7], or the unary iteration procedure of [2, Section III A]. □

5 Applications to chemistry and scattering theory

There are numerous cases in physics where one needs to simulate time-dependent quantum systems. Indeed, the pulse sequences that constitute individual quantum gates or adiabatic sweeps are described by time-dependent Hamiltonians. Here, we look at the particular case of simulating semi-classical scattering of molecules within a chemical reaction as an example of time-dependent Hamiltonian dynamics [24, 46].

Chemical scattering problems involve colliding reagents. As the molecules move closer, the electronic configuration changes due to strengthening Coulomb interactions, which is ultimately responsible for either the reagents forming a bond or flying apart depending on the initial conditions and the nature of the reagents. In the non-relativistic case, the Hamiltonian for two colliding atoms A and B at positions x_A and x_B, respectively, and M electrons with positions x_m for $m = 1, \ldots, M$, can be expressed as

$$\begin{aligned} H &= H_{\text{nuc}} + H_{\text{elec}} \\ H_{\text{nuc}} &= \frac{p_A^2}{2m_A} + \frac{p_B^2}{2m_B} + \frac{Z_A Z_B}{|x_A - x_B|} \\ H_{\text{elec}} &= \sum_{m=1}^{M} \frac{p_m^2}{2m_e} - \sum_{m=1}^{M} \frac{Z_A}{|x_m - x_A|} - \sum_{m=1}^{M} \frac{Z_B}{|x_m - x_B|} + \sum_{m<m'} \frac{1}{|x_m - x_{m'}|}. \end{aligned} \tag{142}$$

Here $p_m = [[p_m]_x, [p_m]_y, [p_m]_z]$ and $x_m = [[x_m]_x, [x_m]_y, [x_m]_z]$ are three-dimensional vectors of operators, whereas the corresponding nuclear terms (such as x_A and x_B) are three-dimensional vectors of scalars. We further define $|x_m - x'_m|$ to be the operator

$$|x_m - x_{m'}| := \sqrt{([x_m]_x - [x_{m'}]_x)^2 + ([x_m]_y - [x_{m'}]_y)^2 + ([x_m]_z - [x_{m'}]_z)^2}.$$

The wave function can be thought of as having a nuclear as well as an electronic component. First, we assume that the nuclear and the electronic wave functions are decoupled [26, 49]:

$$\psi(x_A, x_B, x_1, \ldots, x_M; t) \approx \psi_{\text{nuc}}(x_A, x_B; t)\psi_{\text{elec}}(x_1, \ldots, x_M; t). \tag{143}$$

This approximation is justified by the fact that the nuclear mass is substantially greater than the electronic mass. We then follow the time-dependent self-consistent field (TDSCF) approximation, which further treats x_A and x_B as classical degrees of freedom $x_A(t)$ and $x_B(t)$ with conjugate momenta $(p_A(t), p_B(t))$. This simplification is justified by Ehrenfest's theorem, which states that for a sufficiently narrow quantum wave packet, the equation of motion for the centroid follows the classical trajectory (to leading order in $\hbar$). Under this approximation, the electronic dynamics satisfy

$$i\partial_t|\psi_{\text{elec}}(t)\rangle = \left(\sum_{m=1}^{M} \frac{p_m^2}{2m_e} - \frac{Z_A}{|x_m - x_A(t)|} - \frac{Z_B}{|x_m - x_B(t)|} + \sum_{m<m'} \frac{1}{|x_m - x_{m'}|}\right)|\psi_{\text{elec}}(t)\rangle, \tag{144}$$

where we have suppressed the implicit dependence of the electronic wave function on $x_1, \ldots, x_M$. The equation of motion for the two nuclear positions in the time-dependent self-consistent field approximation within the Ehrenfest method is given by the Hamilton-Jacobi equation:

$$\begin{aligned} \partial_t[p_A(t)]_i &= -\partial_{[x_A]_i}\langle\psi_{\text{elec}}(t)|H_{\text{elec}}|\psi_{\text{elec}}(t)\rangle - \partial_{[x_A]_i}H_{\text{nuc}}(t), \\ \partial_t[x_A(t)]_i &= \partial_{[p_A]_i}H_{\text{nuc}}(t), \end{aligned} \tag{145}$$

and similarly for x_B. The function $H_{\text{nuc}}(t)$ here is simply the Hamiltonian H_{nuc} with the classical substitutions $x_A \to x_A(t)$, $p_A \to p_A(t)$ and similarly for x_B and p_B. Similarly, we define $H_{\text{elec}}(t)$ to be the electronic Hamiltonian under this classical substitution.

The evolution in the Ehrenfest method is governed by a pair of tightly coupled quantum and classical dynamical equations, wherein the full Schrödinger equation only needs to be solved to understand part of the dynamics for the system. Indeed, as the Born-Oppenheimer approximation instantaneously holds under the above approximations, we can further express the electronic dynamics within a second-quantized framework with respect to a basis of molecular orbitals as

$$H_{\text{elec}}(t) = \sum_{pq} h_{pq}(t)a_p^\dagger a_q + \frac{1}{2}\sum_{pqrs} h_{pqrs}(t)a_p^\dagger a_q^\dagger a_r a_s, \tag{146}$$

where for some basis of orthonormal molecular orbitals $\psi_p(\vec{x}; t)$ (which are implicitly time dependent if these basis functions are chosen to be functions of the nuclear positions, as would be appropriate for an atomic orbital basis),

$$h_{pq}(t) = \iint \mathrm{d}_{\vec{x}_1}\mathrm{d}_{\vec{x}_2}\psi_p^*(\vec{x}_1; t)\left(\sum_{m=1}^{M} \frac{p_m^2}{2m_e} - \frac{Z_A}{|x_m - x_A(t)|} - \frac{Z_B}{|x_m - x_B(t)|}\right)\psi_q(\vec{x}_2; t) \tag{147}$$

$$h_{pqrs}(t) = \iiiint \mathrm{d}_{\vec{x}_1}\mathrm{d}_{\vec{x}_2}\mathrm{d}_{\vec{x}_3}\mathrm{d}_{\vec{x}_4}\psi_p^*(\vec{x}_1; t)\psi_q^*(\vec{x}_2; t)\sum_{m<m'} \frac{1}{|x_m - x_{m'}|}\psi_r(\vec{x}_3; t)\psi_s(\vec{x}_4; t). \tag{148}$$

Thus under the above approximations, the dynamics that need to be simulated take the form of a standard second quantized simulation of chemistry, except the Hamiltonian is

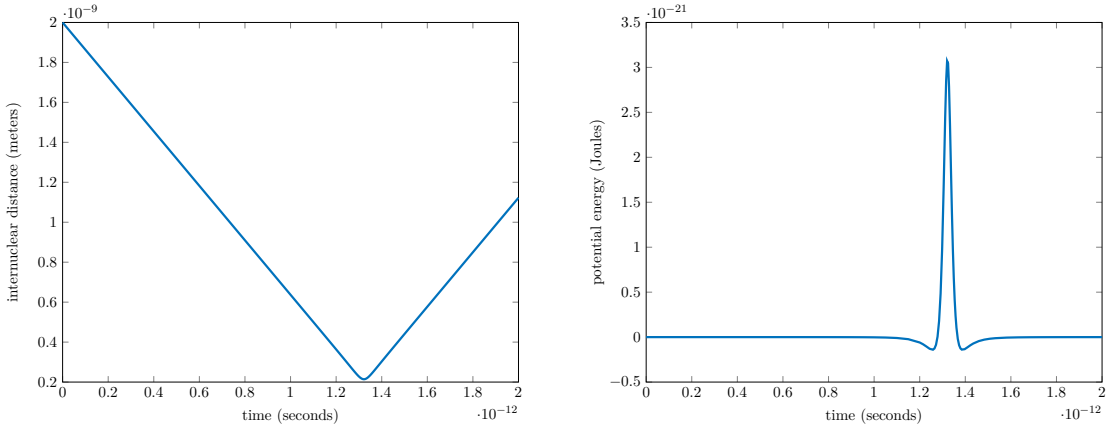

Figure 1: Trajectory for two helium atoms colliding head on and interacting according to a Lennard-Jones potential with an initial separation of 20 nm and a velocity of 1350 m/s.

time dependent. The generalization of this to multiple nuclei is similarly straightforward, with the summation over two nuclear positions replaced by summation over all L positions.

Consider the case where two reagents move towards each other from distant points with large momenta. To get an intuitive understanding of this evolution, it is instructive to examine the case of two molecules colliding using a classical force field. This will give us an expression that is qualitatively accurate for $x_A(t)$ and $x_B(t)$. To do this, we use a Lennard-Jones potential to model the interaction between two helium nuclei. The potential as a function of separation between the nuclei $r(t) = |x_A(t) - x_B(t)|$ is assumed to be of the form $V(r) = \epsilon \left(\frac{r_m^{12}}{r^{12}} - \frac{2r_m^6}{r^6} \right)$, where $\epsilon \approx 10$ K and $r_m \approx 2.6$ Å [44]. Setting the initial radial velocity to be approximately the root mean square (RMS) velocity of helium at 25 °C, we solve the classical equations of motion to find the trajectory shown in Figure 1.

From Figure 1, we find that the interaction appears as a brief but intense kick between the two systems. As a result, the norm of the Hamiltonian changes dramatically throughout the evolution and we expect simulation algorithms with L^1-norm scaling to be advantageous over previous approaches. We leave a detailed study of such an advantage as a subject for future work.

6 Discussion

We have shown that a time-dependent Hamiltonian $H(\tau)$ can be simulated for the time interval $0 \leq \tau \leq t$ with gate complexity that scales according to the L^1 norm $\int_0^t \mathrm{d}\tau \, \|H(\tau)\|$. We designed new algorithms based on classical sampling and improved the previous Dyson-series approach to achieve this scaling. This is a polynomial speedup in terms of the norm dependence, an advantage that can be favorable in practice. In particular, our result has potential applications to simulating scattering processes in quantum chemistry. Our analysis also matches the intuition that the difficulty of simulating a quantum system should depend on the norm of the Hamiltonian instantaneously. This dual interpretation suggests that the L^1-norm dependence of our result cannot be significantly improved. However, further speedup might be possible if we know a priori the energy range of the initial state, as is suggested in [34, 43].

The rescaled Dyson-series approach is nearly optimal with respect to all parameters of interest. Indeed, a lower bound of $\Omega\big(d \, \|H\|_{\max} \, t + \frac{\log(1/\epsilon)}{\log\log(1/\epsilon)}\big)$ was given in [9, Theorem 2] for simulating time-independent sparse Hamiltonians, which of course also holds for the more

general time-dependent case. The query complexity (116) of the rescaled Dyson-series approach matches this dependence on $d\|H\|_{\max} t$ and on ϵ, except that it scales as the product of the two terms instead of the sum (so, as in all quantum simulation algorithms prior to the advent of quantum signal processing [33], it does not achieve the optimal tradeoff between t and ϵ). However, this approach requires computing the rescaling function (114) and the Hamiltonian norm (115) in quantum superposition, which may introduce large overhead in practice. In comparison, continuous qDRIFT relies on classical sampling and may be better suited to near-term simulation. Its complexity has no dependence on the parameter L in the LCU decomposition (Corollary 9′), which is advantageous for Hamiltonians consisting of many terms.

For most of our analysis, we have assumed that the Hamiltonian $H(\tau)$ is continuously differentiable. This assumption can be relaxed to allow finitely many discontinuities. Indeed, if $H(\tau)$ is discontinuous at the times $0 = t_0 < t_1 < \cdots < t_r = t$ but otherwise continuously differentiable, we may divide the evolution into r segments and apply a time-dependent Hamiltonian simulation algorithm within each time interval $[t_j, t_{j+1}]$. For the Dyson-series approach, the complexity depends linearly on the L^1 norm, so concatenation gives a simulation of the entire evolution with L^1-norm scaling. The assumptions about the Hamiltonian can be even further relaxed: the continuous qDRIFT algorithm works properly provided only that $H(\tau)$ is Lebesgue integrable. Further discussion of this point is beyond the scope of this paper, and we refer the reader to [21] for details.

Our analysis can also be adapted to simulate time-dependent Hamiltonians that have countably many zeros. Indeed, since the equation $H(\tau) = 0$ has at most countably many solutions, we can find $c \in \mathbb{R}$ such that $H(\tau) + cI$ is nonzero everywhere. Then, $\exp_{\mathcal{T}}(-i\int_0^t \mathrm{d}\tau\,(H(\tau) + cI)) = e^{-ict}\exp_{\mathcal{T}}(-i\int_0^t \mathrm{d}\tau\, H(\tau))$, so the result is only off by a global phase. Note that this assumption can be completely dropped if we use continuous qDRIFT: we define the exceptional set

$$\mathcal{B}_0 := p^{-1}(0) = \{\tau : p(\tau) = 0\} = \{\tau : \|H(\tau)\|_\infty = 0\} = \{\tau : H(\tau) = 0\} \tag{149}$$

and redefine $\mathcal{U}(t,0)$ as

$$\mathcal{U}(t,0)(\rho) := \int_{[0,t]\setminus\mathcal{B}_0} \mathrm{d}\tau\; p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}} \rho e^{i\frac{H(\tau)}{p(\tau)}}, \qquad p(\tau) := \frac{\|H(\tau)\|_\infty}{\|H\|_{\infty,1}}. \tag{150}$$

We note that $\mathcal{U}(t,0)$ is a valid quantum channel and can be implemented with unit cost. Indeed, for any input state ρ, we randomly sample a value τ according to $p(\tau)$ and perform $e^{-iH(\tau)/p(\tau)}$ if $\tau \in [0,t]\setminus\mathcal{B}_0$, and the identity operation otherwise. This implements

$$\int_{[0,t]\setminus\mathcal{B}_0} \mathrm{d}\tau\; p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}} \rho e^{i\frac{H(\tau)}{p(\tau)}} + \int_{\mathcal{B}_0} \mathrm{d}\tau\; p(\tau)\rho = \mathcal{U}(t,0)(\rho). \tag{151}$$

The remaining analysis proceeds as in Section 3.

The qDRIFT protocol that we analyzed here only achieves first-order accuracy. It is natural to ask if sampling a different probability distribution could lead to an algorithm with better performance. The answer seems to be "no" if we only use a univariate distribution. To see this, consider the discrete case where $H = \sum_{l=1}^L H_l$ is a Hamiltonian consisting of L terms. We sample according to a probability vector $p \in [0,1]^L$. Upon getting outcome l, we perform the unitary e^{-itH_l/p_l}. Effectively, we implement the quantum channel $\mathcal{U}(t)(\rho) := \sum_{l=1}^L p_l e^{-it\frac{H_l}{p_l}} \rho e^{it\frac{H_l}{p_l}}$, which is a first-order approximation to the ideal evolution $\mathcal{E}(t)(\rho) := e^{-it\sum_{l=1}^L H_l} \rho e^{it\sum_{l=1}^L H_l}$. In particular, the difference between $\mathcal{U}(t)(\rho)$

and $\mathcal{E}(t)(\rho)$ admits an integral representation

$$\begin{aligned}\mathcal{U}(t)(\rho)-\mathcal{E}(t)(\rho)=\int_0^t \mathrm{d}u\int_0^u \mathrm{d}v\ \Bigg\{&\sum_{l=1}^L p_l e^{-iv\frac{H_l}{p_l}}\Big[-i\frac{H_l}{p_l},\Big[-i\frac{H_l}{p_l},\rho\Big]\Big]e^{iv\frac{H_l}{p_l}}\\ &-e^{-iv\sum_{l=1}^L H_l}\Big[-i\sum_{l=1}^L H_l,\Big[-i\sum_{l=1}^L H_l,\rho\Big]\Big]e^{iv\sum_{l=1}^L H_l}\Bigg\}.\end{aligned} \tag{152}$$

To estimate the diamond-norm error $\|\mathcal{U}(t)-\mathcal{E}(t)\|_\diamond$, we take σ to be a state on the joint system of the original register and an ancilla register with the same dimension. We compute

$$\begin{aligned}\|(\mathcal{U}(t)\otimes\mathbb{1})(\sigma)-(\mathcal{E}(t)\otimes\mathbb{1})(\sigma)\|_1 &\le \int_0^t \mathrm{d}u\int_0^u \mathrm{d}v\ \Bigg\{\sum_{l=1}^L p_l\left\|\Big[-i\frac{H_l}{p_l}\otimes\mathbb{1},\Big[-i\frac{H_l}{p_l}\otimes\mathbb{1},\sigma\Big]\Big]\right\|_1\\ &\qquad+\left\|\Big[-i\sum_{l=1}^L H_l\otimes\mathbb{1},\Big[-i\sum_{l=1}^L H_l\otimes\mathbb{1},\sigma\Big]\Big]\right\|_1\Bigg\}\\ &\le 2t^2\left(\sum_{l=1}^L\frac{\|H_l\|_\infty^2}{p_l}+\|H\|_{\infty,1}^2\right).\end{aligned} \tag{153}$$

By Jensen's inequality,

$$\sum_{l=1}^L\frac{\|H_l\|_\infty^2}{p_l}=\sum_{l=1}^L p_l\left(\frac{\|H_l\|_\infty}{p_l}\right)^2\ge\left(\sum_{l=1}^L p_l\frac{\|H_l\|_\infty}{p_l}\right)^2=\|H\|_{\infty,1}^2\,, \tag{154}$$

with equality if and only if all $\|H_l\|_\infty/p_l$ are equal, implying that the probability distribution $p_l:=\|H_l\|_\infty/\|H\|_{\infty,1}$ is optimal. A similar optimality result holds for continuous qDRIFT (though the proof is more involved).

However, this does not preclude the existence of a higher-order qDRIFT protocol using more complicated sampling [39]. For example, besides the basic evolutions e^{-itH_l/p_l}, one could evolve under commutators $[H_j,H_k]$ or anticommutators $\{H_j,H_k\}$. We could also use a multivariate distribution and correlate different steps of the qDRIFT protocol. For future work, it would be interesting to find a higher-order protocol, or prove that such a protocol cannot exist.

The fractional-query algorithm described in Section 2.5 provides a natural approach to simulating time-dependent Hamiltonians whose query complexity scales with the L^1-norm. While we believe such a scaling also holds for the gate complexity, it would be highly nontrivial to give an explicit implementation. In any case, the fractional-query approach is streamlined by the Dyson-series approach and the latter can be rescaled to achieve L^1-norm scaling.

The rescaling principle that we proposed can potentially be applied to improve other quantum simulation algorithms. For example, we can use the product-formula algorithm [50] to simulate the rescaled Hamiltonian $\widetilde{H}(\varsigma):=H(f^{-1}(\varsigma))/\|H(f^{-1}(\varsigma))\|_\infty$ for time $s=\|H\|_{\infty,1}$. The difficulty here is that the derivative of the rescaled Hamiltonian can be larger than the original one, making the rescaled algorithm perform worse. We leave a thorough study of this issue as a subject for future work.

Finally, it would be interesting to identify further applications of our L^1-norm scaling result, such as to designing new quantum algorithms and to improving the performance of quantum chemistry simulation. It might also be of interest to demonstrate these approaches experimentally, for applications such as implementing adiabatic algorithms with quantum circuits.

Acknowledgements

We thank Ryan Babbush, Earl Campbell, András Gilyén, Mária Kieferová, Robin Kothari, Tongyang Li, Guang Hao Low, Jarrod R. McClean, Yuval Sanders, and Leonard Wossnig for helpful discussions. DWB is funded by Australian Research Council Discovery Projects DP160102426 and DP190102633. This work was also supported in part by the Army Research Office (MURI award W911NF-16-1-0349), the Canadian Institute for Advanced Research, the National Science Foundation (grant 1813814), and the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Quantum Algorithms Teams and Quantum Testbed Pathfinder programs. In addition, YS is supported by the Google Ph.D. Fellowship program. XW acknowledges support from the Department of Defense. NW is supported in part by a Google Quantum Research Award.

References

[1] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the 35th ACM Symposium on Theory of Computing*, pages 20–29, 2003. DOI: 10.1145/780542.780546. arXiv:quant-ph/0301023.

[2] Ryan Babbush, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven. Encoding electronic spectra in quantum circuits with linear T complexity. *Physical Review X*, 8:041015, Oct 2018. DOI: 10.1103/PhysRevX.8.041015. arXiv:1805.03662.

[3] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin-Lic Chan. Low-depth quantum simulation of materials. *Physical Review X*, 8:011044, Mar 2018. DOI: 10.1103/PhysRevX.8.011044. arXiv:1706.00023.

[4] Dominic W. Berry and Andrew M. Childs. Black-box Hamiltonian simulation and unitary implementation. *Quantum Information and Computation*, 12(1-2):29–62, 2012. arXiv:0910.4157.

[5] Dominic W. Berry, Graeme Ahokas, Richard Cleve, and Barry C. Sanders. Efficient quantum algorithms for simulating sparse Hamiltonians. *Communications in Mathematical Physics*, 270(2):359–371, 2007. DOI: 10.1007/s00220-006-0150-x. arXiv:quant-ph/0508139.

[6] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Exponential improvement in precision for simulating sparse Hamiltonians. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 283–292, 2014. DOI: 10.1145/2591796.2591854. arXiv:1312.1414.

[7] Dominic W. Berry, Richard Cleve, and Sevag Gharibian. Gate-efficient discrete simulations of continuous-time quantum query algorithms. *Quantum Information and Computation*, 14(1-2):1–30, January 2014. arXiv:1211.4637.

[8] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Physical Review Letters*, 114(9):090502, 2015. DOI: 10.1103/PhysRevLett.114.090502. arXiv:1412.4687.

[9] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Proceedings of the 56th IEEE Symposium on Foundations of Computer Science*, pages 792–809, 2015. DOI: 10.1109/FOCS.2015.54. arXiv:1501.01715.

[10] Dominic W. Berry, Andrew M. Childs, Aaron Ostrander, and Guoming Wang. Quantum algorithm for linear differential equations with exponentially improved depen-

dence on precision. *Communications in Mathematical Physics*, 356(3):1057–1081, Dec 2017. ISSN 1432-0916. DOI: 10.1007/s00220-017-3002-y. arXiv:1701.03684.

[11] Fernando G. S. L. Brandao and Krysta M. Svore. Quantum speed-ups for solving semidefinite programs. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science*, pages 415–426, 2017. DOI: 10.1109/FOCS.2017.45. arXiv:1609.05537.

[12] Laurie J. Butler. Chemical reaction dynamics beyond the Born-Oppenheimer approximation. *Annual Review of Physical Chemistry*, 49(1):125–171, 1998. DOI: 10.1146/annurev.physchem.49.1.125.

[13] Earl Campbell. Random compiler for fast Hamiltonian simulation. *Physical Review Letters*, 123:070503, Aug 2019. DOI: 10.1103/PhysRevLett.123.070503. arXiv:1811.08017.

[14] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, 2019. DOI: 10.1021/acs.chemrev.8b00803. arXiv:1812.09976.

[15] Andrew M. Childs and Robin Kothari. Limitations on the simulation of non-sparse Hamiltonians. *Quantum Information and Computation*, 10(7-8):669–684, 2010. arXiv:0908.4398.

[16] Andrew M. Childs and Yuan Su. Nearly optimal lattice simulation by product formulas. *Physical Review Letters*, 123:050503, Aug 2019. DOI: 10.1103/PhysRevLett.123.050503. arXiv:1901.00564.

[17] Andrew M. Childs, Richard Cleve, Enrico Deotto, Edward Farhi, Sam Gutmann, and Daniel A. Spielman. Exponential algorithmic speedup by quantum walk. In *Proceedings of the 35th ACM Symposium on Theory of Computing*, pages 59–68, 2003. DOI: 10.1145/780542.780552. arXiv:quant-ph/0209131.

[18] Andrew M. Childs, Dmitri Maslov, Yunseong Nam, Neil J. Ross, and Yuan Su. Toward the first quantum simulation with quantum speedup. *Proceedings of the National Academy of Sciences*, 115(38):9456–9461, 2018. DOI: 10.1073/pnas.1801723115. arXiv:1711.10980.

[19] Andrew M. Childs, Aaron Ostrander, and Yuan Su. Faster quantum simulation by randomization. *Quantum*, 3:182, September 2019. DOI: 10.22331/q-2019-09-02-182. arXiv:1805.08385.

[20] Anirban Narayan Chowdhury and Rolando D. Somma. Quantum algorithms for Gibbs sampling and hitting-time estimation. *Quantum Information and Computation*, 17(1-2):41–64, 2017. arXiv:1603.02940.

[21] John Day Dollard and Charles N. Friedman. *Product Integration with Application to Differential Equations*. Cambridge University Press, 1984. DOI: 10.1017/CBO9781107340701.

[22] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292(5516):472–475, 2001. DOI: 10.1126/science.1057726. arXiv:quant-ph/0104129.

[23] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum algorithm for the Hamiltonian NAND tree. *Theory of Computing*, 4(1):169–190, 2008. DOI: 10.4086/toc.2008.v004a008. quant-ph/0702144.

[24] Antonio Fernández-Ramos, James A Miller, Stephen J Klippenstein, and Donald G

Truhlar. Modeling the kinetics of bimolecular reactions. *Chemical reviews*, 106(11): 4518–4584, 2006. DOI: 10.1021/cr050205w.

[25] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982. DOI: 10.1007/BF02650179.

[26] Robert B. Gerber, Victoria Buch, and Mark A. Ratner. Time-dependent self-consistent field approximation for intramolecular energy transfer. I. formulation and application to dissociation of van der Waals molecules. *Journal of Chemical Physics*, 77(6):3022–3030, 1982. DOI: 10.1063/1.444225.

[27] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15):150502, 2009. DOI: 10.1103/PhysRevLett.103.150502. arXiv:0811.3171.

[28] Stephen P. Jordan, Keith S. M. Lee, and John Preskill. Quantum algorithms for quantum field theories. *Science*, 336(6085):1130–1133, 2012. DOI: 10.1126/science.1217069. arXiv:1111.3633.

[29] Mária Kieferová, Artur Scherer, and Dominic Berry. Simulating the dynamics of time-dependent Hamiltonians with a truncated Dyson series. *Physical Review A*, 99:042314, 2019. DOI: 10.1103/PhysRevA.99.042314. arXiv:1805.00582.

[30] Anthony W. Knapp. *Basic Real Analysis*. Birkhëuser, 2005. DOI: 10.3792/euclid/9781429799997.

[31] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996. DOI: 10.1126/science.273.5278.1073.

[32] Guang Hao Low. Hamiltonian simulation with nearly optimal dependence on spectral norm. In *Proceedings of the 51th ACM Symposium on Theory of Computing*, pages 491–502. ACM, 2019. DOI: 10.1145/3313276.3316386. arXiv:1807.03967.

[33] Guang Hao Low and Isaac L. Chuang. Optimal Hamiltonian simulation by quantum signal processing. *Physical Review Letters*, 118:010501, 2017. DOI: 10.1103/PhysRevLett.118.010501. arXiv:1606.02685.

[34] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by uniform spectral amplification, 2017. arXiv:1707.05391.

[35] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, July 2019. DOI: 10.22331/q-2019-07-12-163. arXiv:1610.06546.

[36] Guang Hao Low and Nathan Wiebe. Hamiltonian simulation in the interaction picture, 2018. arXiv:1805.00675.

[37] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Reviews of Modern Physics*, 92:015003, Mar 2020. DOI: 10.1103/RevModPhys.92.015003. arXiv:1808.10402.

[38] Michael A. Nielsen, Mark R. Dowling, Mile Gu, and Andrew C. Doherty. Optimal control, geometry, and quantum computing. *Physical Review A*, 73:062323, Jun 2006. DOI: 10.1103/PhysRevA.73.062323. arXiv:quant-ph/0603160.

[39] Yingkai Ouyang, David R. White, and Earl T. Campbell. Compilation by stochastic Hamiltonian sparsification. *Quantum*, 4:235, February 2020. DOI: 10.22331/q-2020-02-27-235. arXiv:1910.06255.

[40] Shengshi Pang and Andrew N. Jordan. Optimal adaptive control for quantum metrology with time-dependent Hamiltonians. *Nature Communications*, 8:14695, 2017. DOI: 10.1038/ncomms14695. arXiv:1606.02166.

[41] David Poulin and Pawel Wocjan. Preparing ground states of quantum many-body systems on a quantum computer. *Physical Review Letters*, 102:130503, Apr 2009. DOI: 10.1103/PhysRevLett.102.130503. arXiv:0809.2705.

[42] David Poulin, Angie Qarry, Rolando D. Somma, and Frank Verstraete. Quantum simulation of time-dependent Hamiltonians and the convenient illusion of Hilbert space. *Physical Review Letters*, 106(17):170501, 2011. DOI: 10.1103/PhysRevLett.106.170501. arXiv:1102.1360.

[43] David Poulin, Matthew B. Hastings, Dave Wecker, Nathan Wiebe, Andrew C. Doherty, and Matthias Troyer. The Trotter step size required for accurate quantum simulation of quantum chemistry. *Quantum Information and Computation*, 15(5-6): 361–384, 2015. arXiv:1406.4920.

[44] Orhan Talu and Alan L Myers. Reference potentials for adsorption of helium, argon, methane, and krypton in high-silica zeolites. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 187:83–93, 2001. DOI: 10.1016/S0927-7757(01)00628-8.

[45] Minh C. Tran, Andrew Y. Guo, Yuan Su, James R. Garrison, Zachary Eldredge, Michael Foss-Feig, Andrew M. Childs, and Alexey V. Gorshkov. Locality and digital quantum simulation of power-law interactions. *Physical Review X*, 9:031006, Jul 2019. DOI: 10.1103/PhysRevX.9.031006. arXiv:1808.05225.

[46] John C. Tully. Mixed quantum–classical dynamics. *Faraday Discussions*, 110:407–419, 1998. DOI: 10.1039/a801824c.

[47] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. DOI: 10.1017/9781316848142.

[48] James D. Whitfield, Jacob Biamonte, and Alán Aspuru-Guzik. Simulation of electronic structure Hamiltonians using quantum computers. *Molecular Physics*, 109(5):735–750, 2011. DOI: 10.1080/00268976.2011.552441. arXiv:1001.3855.

[49] Gregory S. Whittier and John C. Light. Quantum/classical time-dependent self-consistent field treatment of Ar+HCO inelastic and dissociative scattering. *Journal of Chemical Physics*, 110(9):4280–4290, 1999. DOI: 10.1063/1.478291.

[50] Nathan Wiebe, Dominic Berry, Peter Høyer, and Barry C Sanders. Higher order decompositions of ordered operator exponentials. *Journal of Physics A*, 43(6):065203, 2010. DOI: 10.1088/1751-8113/43/6/065203. arXiv:0812.0562.

[51] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2017. DOI: 10.1017/9781316809976.

A Continuous qDRIFT and Hamiltonian averaging

Poulin, Qarry, Somma, and Verstraete developed an algorithm for time-dependent Hamiltonian simulation based on techniques of Hamiltonian averaging and Monte Carlo estimation [42]. In this section, we discuss the relation between their algorithm and our continuous qDRIFT.

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \leq \tau \leq t$. Assume that $H(\tau)$ is continuous, nonzero everywhere, and efficiently simulable for each particular τ. Then, Poulin et al.'s approach simulates $H(\tau)$ for time $\tau \in [0, t]$ in two steps: (i) they replace the evolution $\exp_{\mathcal{T}}\left(-i\int_0^t \mathrm{d}\tau\, H(\tau)\right)$ by an ordinary matrix exponential $e^{-itH_{\mathrm{av}}}$ of the average Hamiltonian $H_{\mathrm{av}} := \frac{1}{t}\int_0^t \mathrm{d}\tau\, H(\tau)$ with an error that scales like $O((t\,\|H\|_{\infty,\infty})^2)$; (ii) they further implement $\int_0^t \mathrm{d}\tau\, H(\tau)$ with Monte Carlo estimation by picking m random times and approximating $\int_0^t \mathrm{d}\tau\; H(\tau) \approx \frac{1}{m}\sum_{k=1}^m H(\tau_k)$ with error $O\big(t\,\|H\|_{\infty,\infty}/\sqrt{m}\big)$, the result of which is further approximated by product formulas.

The approach of [42] is essentially a sampling-based algorithm and thus similar in spirit to our continuous qDRIFT, except for a notable difference: their algorithm scales with the L^∞ norm instead of the L^1 norm. Unfortunately, this drawback cannot be remedied merely

by a better analysis of the same algorithm. Indeed, they use a uniform distribution to pick random times during the Monte Carlo estimation. This sampling ignores the instantaneous norm $\|H(\tau)\|_\infty$ of the Hamiltonian and therefore the resulting algorithm cannot scale with the L^1 norm $\int_0^t \mathrm{d}\tau \, \|H(\tau)\|_\infty$.

Instead, continuous qDRIFT uses a probability distribution that biases toward those times with larger instantaneous norm. In Section 3, we proved that such a sampling gives a direct simulation of time-dependent Hamiltonians with complexity that scales with the L^1 norm. We now give an indirect implementation: (i') we show in Appendix A.1 that the error of replacing $\exp_{\mathcal{T}}\big(-i\int_0^t \mathrm{d}\tau\, H(\tau)\big)$ by an ordinary matrix exponential of H_{av} scales like $O\big(\|H\|_{\infty,1}^2\big)$, improving the analysis of [42]; (ii') we further prove in Appendix A.2 that the average Hamiltonian can be simulated by continuous qDRIFT with L^1-norm scaling. Combining these two steps, we see that the Monte Carlo estimation approach of [42] is superseded by continuous qDRIFT.

A.1 Hamiltonian averaging

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ and assume that it is continuous and nonzero everywhere. Define

$$\mathrm{E}(s,0) := \exp_{\mathcal{T}}\bigg(-i\int_0^s \mathrm{d}\tau\, H(\tau)\bigg), \qquad \mathrm{E}_{\mathrm{av}}(s) := e^{-isH_{\mathrm{av}}}, \tag{155}$$

where $H_{\mathrm{av}} := \frac{1}{t}\int_0^t \mathrm{d}\tau\, H(\tau)$ is the average Hamiltonian. Our goal is to bound the distance between $\mathrm{E}(s,0)$ and $\mathrm{E}_{\mathrm{av}}(s)$ at $s=t$. Using the initial condition $\mathrm{E}(0,0) = \mathrm{E}_{\mathrm{av}}(0) = I$, we have

$$\|\mathrm{E}(t,0) - \mathrm{E}_{\mathrm{av}}(t)\|_\infty = \Big\|\mathrm{E}_{\mathrm{av}}^\dagger(t)\mathrm{E}(t,0) - I\Big\|_\infty = \bigg\|\int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}\big[\mathrm{E}_{\mathrm{av}}^\dagger(s)\mathrm{E}(s,0)\big]\bigg\|_\infty . \tag{156}$$

By the Schrödinger equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}_{\mathrm{av}}(s) = -iH_{\mathrm{av}}\mathrm{E}_{\mathrm{av}}(s), \qquad \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}(s,0) = -iH(s)\mathrm{E}(s,0), \tag{157}$$

we obtain

$$\begin{aligned}
\int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}\big[\mathrm{E}_{\mathrm{av}}^\dagger(s)\mathrm{E}(s,0)\big] &= \int_0^t \mathrm{d}s \Big\{\mathrm{E}_{\mathrm{av}}^\dagger(s)\big[iH_{\mathrm{av}}\big]\mathrm{E}(s,0) + \mathrm{E}_{\mathrm{av}}^\dagger(s)\big[-iH(s)\big]\mathrm{E}(s,0)\Big\} \\
&= \frac{1}{t}\int_0^t \mathrm{d}s \int_0^t \mathrm{d}\tau \Big\{\mathrm{E}_{\mathrm{av}}^\dagger(s)\big[iH(\tau)\big]\mathrm{E}(s,0) + \mathrm{E}_{\mathrm{av}}^\dagger(s)\big[-iH(s)\big]\mathrm{E}(s,0)\Big\} \\
&= \frac{1}{t}\int_0^t \mathrm{d}s \int_0^t \mathrm{d}\tau \Big\{\mathrm{E}_{\mathrm{av}}^\dagger(s)\big[iH(\tau)\big]\mathrm{E}(s,0) + \mathrm{E}_{\mathrm{av}}^\dagger(\tau)\big[-iH(\tau)\big]\mathrm{E}(\tau,0)\Big\},
\end{aligned} \tag{158}$$

which implies, by telescoping, that

$$\begin{aligned}
\bigg\|\int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}\big[\mathrm{E}_{\mathrm{av}}^\dagger(s)\mathrm{E}(s,0)\big]\bigg\|_\infty \le \frac{1}{t}\int_0^t \mathrm{d}s \int_0^t \mathrm{d}\tau \, \big(&\|\mathrm{E}_{\mathrm{av}}(s) - \mathrm{E}_{\mathrm{av}}(\tau)\|_\infty \, \|H(\tau)\|_\infty \\
&+ \|\mathrm{E}(s,0) - \mathrm{E}(\tau,0)\|_\infty \, \|H(\tau)\|_\infty \big).
\end{aligned} \tag{159}$$

By the fundamental theorem of calculus, the first term of the integrand can be bounded as

$$\|\mathrm{E}_{\mathrm{av}}(s) - \mathrm{E}_{\mathrm{av}}(\tau)\|_\infty \le \|H_{\mathrm{av}}\|_\infty \, |s-\tau| \le \frac{1}{t}\int_0^t \mathrm{d}u \, \|H(u)\|_\infty \, |s-\tau| \le \|H\|_{\infty,1} \, . \tag{160}$$

To handle the second term, we use Lemma 1. Observe that the generator of $\mathrm{E}(s,0)$ is $H(u),\ 0 \le u \le s$, whereas the generator of $\mathrm{E}(\tau,0)$ is $H(u),\ 0 \le u \le \tau$. So they only differ on the interval $[\min\{s,\tau\}, \max\{s,\tau\}]$. Consequently,

$$\|\mathrm{E}(s,0) - \mathrm{E}(\tau,0)\|_\infty \le \int_{\min\{s,\tau\}}^{\max\{s,\tau\}} \mathrm{d}u \, \|H(u)\|_\infty \le \int_0^t \mathrm{d}u \, \|H(u)\|_\infty = \|H\|_{\infty,1} . \tag{161}$$

Altogether, we have

$$\begin{aligned} \|\mathrm{E}(t,0) - \mathrm{E}_{\mathrm{av}}(t)\|_\infty &= \left\| \int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s} [\mathrm{E}_{\mathrm{av}}(s)^\dagger \mathrm{E}(s,0)] \right\|_\infty \\ &\le \frac{1}{t} \int_0^t \mathrm{d}s \int_0^t \mathrm{d}\tau \ \left(\|\mathrm{E}_{\mathrm{av}}(s) - \mathrm{E}_{\mathrm{av}}(\tau)\|_\infty \, \|H(\tau)\|_\infty \right. \\ &\qquad \left. + \|\mathrm{E}(s,0) - \mathrm{E}(\tau,0)\|_\infty \, \|H(\tau)\|_\infty \right) \\ &\le 2 \, \|H\|_{\infty,1}^2 . \end{aligned} \tag{162}$$

Theorem 11 (Hamiltonian simulation by averaging (spectral-norm distance))**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ and assume that it is continuous and nonzero everywhere. Define $\mathrm{E}(t,0) := \exp_{\mathcal{T}} \left(-i \int_0^t \mathrm{d}\tau \, H(\tau) \right)$ and $\mathrm{E}_{av}(t) := e^{-itH_{av}}$, where $H_{av} := \frac{1}{t} \int_0^t \mathrm{d}\tau \, H(\tau)$ is the average Hamiltonian. Then,*

$$\|\mathrm{E}(t,0) - \mathrm{E}_{av}(t)\|_\infty \le 2 \, \|H\|_{\infty,1}^2 . \tag{163}$$

The above bound on the spectral-norm error can be converted to a bound on the diamond-norm error using Lemma 2.

Theorem 11′ (Hamiltonian simulation by averaging (diamond-norm distance))**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ and assume that it is continuous and nonzero everywhere. Define unitary operators $\mathrm{E}(t,0) := \exp_{\mathcal{T}} \left(-i \int_0^t \mathrm{d}\tau \ H(\tau) \right)$, $\mathrm{E}_{av}(t) := e^{-itH_{av}}$ and let $\mathcal{E}(t,0)(\cdot) := \mathrm{E}(t,0)(\cdot)\mathrm{E}(t,0)^\dagger$, $\mathcal{E}_{av}(t)(\cdot) = \mathrm{E}_{av}(t)(\cdot)\mathrm{E}_{av}(t)^\dagger$ be the corresponding channels. Then,*

$$\|\mathcal{E}(t,0) - \mathcal{E}_{av}(t)\|_\diamond \le 4 \, \|H\|_{\infty,1}^2 . \tag{164}$$

A.2 Implementing Hamiltonian averaging by continuous qDRIFT

Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ and assume that it is continuous and nonzero everywhere. We have showed that the ideal evolution can be approximated by an evolution under the average Hamiltonian with error that scales with the L^1 norm. We now show that such a Hamiltonian averaging can be implemented by continuous qDRIFT, again with L^1-norm scaling. This improves over the algorithm of [42] which scales with the L^∞ norm.

Theorem 12 (Hamiltonian averaging by continuous qDRIFT)**.** *Let $H(\tau)$ be a time-dependent Hamiltonian defined for $0 \le \tau \le t$ and assume that it is continuous and nonzero everywhere. Define $\mathrm{E}_{av}(t) := e^{-itH_{av}}$ and let $\mathcal{E}_{av}(t)(\cdot) = \mathrm{E}_{av}(t)(\cdot)\mathrm{E}_{av}(t)^\dagger$ be the corresponding channels. Let $\mathcal{U}(t,0)$ be the continuous qDRIFT channel*

$$\mathcal{U}(t,0)(\rho) = \int_0^t \mathrm{d}\tau \ p(\tau) e^{-i\frac{H(\tau)}{p(\tau)}} \rho e^{i\frac{H(\tau)}{p(\tau)}} , \tag{165}$$

where $p(\tau) = \|H(\tau)\|_\infty / \|H\|_{\infty,1}$. Then,

$$\|\mathcal{E}_{av}(t) - \mathcal{U}(t,0)\|_\diamond \le 4 \, \|H\|_{\infty,1}^2 . \tag{166}$$

Note that by applying the triangle inequality to Theorem 5 and Theorem 11′, we obtain

$$\|\mathcal{E}_{\mathrm{av}}(t) - \mathcal{U}(t,0)\|_{\diamond} \leq 8\,\|H\|_{\infty,1}^2\,. \tag{167}$$

Theorem 12 improves the constant prefactor from 8 to 4.

Proof of Theorem 12. We parametrize the two channels $\mathcal{E}_{\mathrm{av}}(t)$, $\mathcal{U}(t)$ and define

$$\mathcal{E}_{\mathrm{av},u}(t)(\rho) := e^{-iu\int_0^t \mathrm{d}\tau H(\tau)}\rho e^{iu\int_0^t \mathrm{d}\tau H(\tau)}, \quad \mathcal{U}_u(t,0)(\rho) := \int_0^t \mathrm{d}\tau\; p(\tau) e^{-iu\frac{H(\tau)}{p(\tau)}}\rho e^{iu\frac{H(\tau)}{p(\tau)}}. \tag{168}$$

Since $\mathcal{E}_{\mathrm{av},0}(t)(\rho) = \rho$, $\mathcal{E}_{\mathrm{av},1}(t)(\rho) = \mathcal{E}_{\mathrm{av}}(t)(\rho)$, $\mathcal{U}_0(t,0)(\rho) = \rho$, and $\mathcal{U}_1(t,0)(\rho) = \mathcal{U}(\rho)$, the first derivative of $\mathcal{E}_{\mathrm{av},u}(t)(\rho)$ and $\mathcal{U}_u(t,0)(\rho)$ agrees with each other at $u = 0$

$$\frac{\mathrm{d}}{\mathrm{d}u}\mathcal{E}_{\mathrm{av},u}(t)(\rho)\Big|_{u=0} = \Big[-i\int_0^t \mathrm{d}\tau\, H(\tau), \rho\Big] = \frac{\mathrm{d}}{\mathrm{d}u}\mathcal{U}_u(t,0)(\rho)\Big|_{u=0}. \tag{169}$$

Applying the fundamental theorem of calculus twice, we obtain

$$\begin{aligned}
&\mathcal{E}_{\mathrm{av}}(t)(\rho) - \mathcal{U}(t,0)(\rho)\\
&= (\mathcal{E}_{\mathrm{av},1}(t)(\rho) - \mathcal{E}_{\mathrm{av},0}(t)(\rho)) - (\mathcal{U}_1(t,0)(\rho) - \mathcal{U}_0(t,0)(\rho))\\
&= \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}v\; \frac{\mathrm{d}^2}{\mathrm{d}v^2}[\mathcal{E}_{\mathrm{av},v}(t)(\rho) - \mathcal{U}_v(t,0)(\rho)]\\
&= \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}v\; \Big\{e^{-iv\int_0^t \mathrm{d}\tau H(\tau)}\Big[-i\int_0^t \mathrm{d}\tau H(\tau), \Big[-i\int_0^t \mathrm{d}\tau H(\tau), \rho\Big]\Big] e^{iv\int_0^t \mathrm{d}\tau H(\tau)}\\
&\qquad\qquad - \int_0^t \mathrm{d}\tau\; p(\tau) e^{-iv\frac{H(\tau)}{p(\tau)}}\Big[-i\frac{H(\tau)}{p(\tau)}, \Big[-i\frac{H(\tau)}{p(\tau)}, \rho\Big]\Big] e^{iv\frac{H(\tau)}{p(\tau)}}\Big\}.
\end{aligned} \tag{170}$$

We take σ to be a state on the joint system of the original register and an ancilla register with the same dimension. Using properties of the Schatten norms, we have

$$\begin{aligned}
&\|(\mathcal{E}_{\mathrm{av}}(t)\otimes\mathbb{1})(\sigma) - (\mathcal{U}(t)\otimes\mathbb{1})(\sigma)\|_1\\
&\leq \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}v\; \Big\{\Big\|\Big[-i\int_0^t \mathrm{d}\tau H(\tau)\otimes\mathbb{1}, \Big[-i\int_0^t \mathrm{d}\tau H(\tau)\otimes\mathbb{1}, \sigma\Big]\Big]\Big\|_1\\
&\qquad\qquad + \int_0^t \mathrm{d}\tau\; p(\tau)\Big\|\Big[-i\frac{H(\tau)}{p(\tau)}\otimes\mathbb{1}, \Big[-i\frac{H(\tau)}{p(\tau)}\otimes\mathbb{1}, \sigma\Big]\Big]\Big\|_1\Big\}\\
&\leq \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}v\; \Big[4\,\|H\|_{\infty,1}^2 + 4\int_0^t \mathrm{d}\tau\; \frac{\|H(\tau)\|_\infty^2}{p(\tau)}\Big].
\end{aligned} \tag{171}$$

Using the definition $p(\tau) = \|H(\tau)\|_\infty / \|H\|_{\infty,1}$, the second term of the integrand can be further simplified as

$$\int_0^t \mathrm{d}\tau\; \frac{\|H(\tau)\|_\infty^2}{p(\tau)} = \|H\|_{\infty,1}^2\,, \tag{172}$$

giving

$$\|(\mathcal{E}_{\mathrm{av}}(t)\otimes\mathbb{1})(\sigma) - (\mathcal{U}(t,0)\otimes\mathbb{1})(\sigma)\|_1 \leq 4\,\|H\|_{\infty,1}^2\,. \tag{173}$$

Optimizing over σ proves the claimed bound. □