Nonlinear dynamics as a ground-state solution on quantum computers

Albert J. Pool [1,2] Alejandro D. Somoza [1,2] Conor Mc Keever [3] Michael Lubasch [3,*] and Birger Horstmann [1,2,4,†]

[1]*Institute of Engineering Thermodynamics, German Aerospace Center (DLR), Wilhelm-Runge-Str. 10, 89081 Ulm, Germany*
[2]*Helmholtz Institute Ulm, Helmholtzstr. 11, 89081 Ulm, Germany*
[3]*Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom*
[4]*Department of Physics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany*

For the solution of time-dependent nonlinear differential equations, we present variational quantum algorithms (VQAs) that encode both space and time in qubit registers. The spacetime encoding enables us to obtain the entire time evolution from a single ground-state computation. We describe a general procedure to construct efficient quantum circuits for the cost function evaluation required by VQAs. To mitigate the barren plateau problem during the optimization, we propose an adaptive multigrid strategy. The approach is illustrated for the nonlinear Burgers equation. We classically optimize quantum circuits to represent the desired ground-state solutions, run them on IBM Q System One and Quantinuum System Model H1, and demonstrate that current quantum computers are capable of accurately reproducing the exact results.

I. INTRODUCTION

Methods to solve partial differential equations (PDEs), particularly those which are nonlinear, have long been of central importance in fields such as aerospace engineering [1] and energy science [2]. Prominent applications include solving the Navier-Stokes equation in computational fluid dynamics and the numerical integration of continuum models in battery research [3,4]. Despite their efficacy, conventional numerical methods encounter substantial constraints when tackling large-scale three-dimensional models [5].

Numerous quantum algorithms have been proposed to integrate PDEs with an exponential advantage over their classical counterparts in theory [6]. This includes quantum linear systems algorithms (QLSAs) for linear [7–16] and nonlinear [17–21] PDEs, as well as quantum algorithms based on Hamiltonian simulation [11,17,21–27]. Furthermore, techniques based on quantum amplitude and phase estimation have been put forth to integrate PDEs [28–32] (which might be realizable using approximations with low quantum hardware requirements [33–35]). In practice, however, these algorithms do not yet give a quantum advantage. Various challenges, including the input-output problem, hinder an actual quantum advantage in some cases [36,37]. On current noisy intermediate-scale quantum (NISQ) devices [38], these

algorithms are affected by strong limitations on the depth of the quantum circuits that can be coherently executed.

On NISQ devices, variational quantum algorithms (VQAs) have become a popular approach to solving PDEs [39]. One famous example is the Variational Quantum Linear Solver, which is a generic solver for linear systems of equations and can be applied to linear PDEs [40–42]. A standard benchmark is the Poisson equation [43–47]. For nonlinear equations, different algorithmic primitives have been proposed [48–52].

One way of solving time-dependent PDEs using VQAs proceeds one time step at a time, analogous to established classical numerical solvers [53]. In this time-stepping approach, however, a large number of time steps can lead to an accumulation of errors. Also note that time stepping leads to a computational complexity that is at least linear in the number of time steps.

In this paper, our goal is to develop alternative VQAs that do not rely on time stepping. To that end, we use the formalism of Feynman [54,55] and Kitaev *et al.* [56], in which time is encoded in a clock qubit register and the solution to a time-dependent problem at all points in time is contained in the ground state of a Hamiltonian. A proof-of-principle demonstration of this spacetime approach was introduced by McClean *et al.* for quantum chemical problems in 2013 [57], which was extended to open quantum systems by Tempel and Aspuru-Guzik in 2014 [58]. A recent implementation of this method has been put forth by Barison *et al.* [59] for the dynamics of quantum systems, providing numerical evidence for a favorable scaling.

We present methods that extend the Feynman-Kitaev formalism to incorporate nonlinear dynamics [60]. We show how to evaluate the Feynman-Kitaev Hamiltonian using quantum nonlinear processing units (QNPUs) [49,61,62]. To mitigate barren plateaus, which are a serious problem for variational algorithms [63], and scale the variational algorithm to a larger
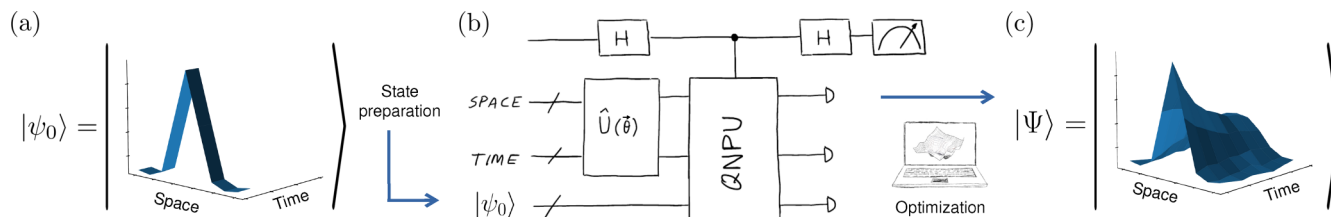

FIG. 1. Schematic summarizing the proposed approach. (a) The quantum state $|\psi_0\rangle$ encodes the initial condition at $t = 0$. (b) The nonlinear Hamiltonian is evaluated using QNPUs. The solution to the PDE at all times corresponds to the ground state of the nonlinear Hamiltonian, which can be calculated via the variational optimization of an ansatz $U(\vec{\theta})$ with $n_x + n_t$ qubits and variational parameters $\vec{\theta}$. (c) 3D plot of the solution corresponding to the result from the Quantinuum H1-1 computer. See Fig. 3(a) for more details.

number of qubits, we propose a multigrid optimization strategy with a customized ansatz structure [64]. In this approach, a converged ansatz is passed from a coarser to a finer grid, which allows us to speed up the optimization by limiting the number of variational parameters. We apply the approach to the Burgers equation as a paradigmatic nonlinear equation. We show that the number of circuits required to evaluate the Feynman-Kitaev Hamiltonian only depends on the number of terms in the equation and the order of approximation in the time step, not on the number of time steps or spatial points, and that these circuits have a depth that scales at most linearly with the number of qubits. A schematic of the approach is depicted in Fig. 1.

This paper has the following structure. In Sec. II, we first present the modified Feynman-Kitaev cost function and its implementation in terms of QNPUs. We then show how to evaluate this cost function on NISQ devices and how to optimize it using VQAs. In Sec. III, we present results from the IBMQ and Quantinuum devices. We then show the cost achieved for different variational ansatz structures and discuss the scalability of the algorithm. We provide our conclusions in Sec. IV and technical details in the Appendixes.

II. METHODS

Here we describe how we use the Feynman-Kitaev formalism to solve nonlinear PDEs on NISQ devices. In Sec. II A, we derive the Feynman-Kitaev Hamiltonian for nonlinear systems. In Sec. II B, we show how to construct QNPU circuits to evaluate it. In Sec. II C, we describe the variational procedure which we use to find the ground state of the Hamiltonian. Finally, we discuss in Sec. II D what is necessary to implement this Hamiltonian on a NISQ device.

A. Feynman-Kitaev Hamiltonian for nonlinear systems

We want to apply the Feynman-Kitaev formalism to time-dependent differential equations of the form

$$\frac{d}{dt} f(x, t) = \mathcal{L}[f(x, t)] f(x, t), \qquad (1)$$

where $\mathcal{L}$ is a (linear or nonlinear) differential operator acting on the function $f(x, t)$. To implement nonlinear equations, the operator $\mathcal{L}$ may have a functional dependence on $f(x, t)$. Such an equation can be solved classically by defining a propagator [53]

$$\hat{T}(dt) = e^{dt \mathcal{L}}. \qquad (2)$$

The exponential of the operator can then be approximated using the Taylor series

$$\hat{T}(dt) = \mathbb{1} + dt \, \mathcal{L} + (dt \, \mathcal{L})^2/2 + \mathcal{O}(dt^3), \qquad (3)$$

where the time step is assumed to be small, i.e., $dt \ll 1$. Although this is formulated for a time-independent problem, the generalization to time-dependent problems using a time-ordered propagator is straightforward [57]. In contrast to quantum dynamics [59], the propagator $\hat{T}$ is not unitary for classical systems in general.

To encode the function $f(x, t)$ in a quantum state $|\Psi\rangle$, we define a spatial register of n_x qubits and a time register of n_t qubits. We then discretize the function $f(x, t)$ on a grid of $2^{n_x} = N_x + 1$ points in space and $2^{n_t} = N_t + 1$ in time and use the amplitude encoding [23,59] given by

$$|\Psi\rangle = \sum_{i=0}^{N_x} \sum_{j=0}^{N_t} \psi_{i,j} |\text{binary}(i)\rangle \otimes |\text{binary}(j)\rangle, \qquad (4)$$

where binary(i) is the binary encoding of the number i as a bitstring on the space or time qubits. This quantum state is normalized as $\sum_{i,j} |\psi_{i,j}|^2 = 1$. We now define the Feynman-Kitaev Hamiltonian,

$$\hat{H} = c_0 \hat{C}_0 + \hat{X}^\dagger \hat{X}, \qquad (5)$$

where the operator $\hat{C}_0$, which enforces the initial condition $|\psi_0\rangle = f(x, t_0)/\mathcal{N}$, is given by

$$\hat{C}_0 = [\hat{I} - |\psi_0\rangle\langle\psi_0|] \otimes |0\rangle\langle 0|. \qquad (6)$$

The constant c_0 can be tuned to speed up optimization; the ground state of the Hamiltonian, for which $\langle \hat{C}_0 \rangle = 0$, does not depend on it. We use $c_0 = 2$ in all cases presented in this paper. The normalization constant $\mathcal{N} = \|f(x, t_0)\|_2$ is used to rescale the results such that the norm of the first time point matches $\mathcal{N}$ and the initial condition is obeyed. The operator $\hat{X}$, which describes the time evolution, is given by

$$\hat{X} = \sum_{i=0}^{N_t-1} \hat{T}(-dt) \otimes |i\rangle\langle i+1| - \hat{I} \otimes |i\rangle\langle i|. \qquad (7)$$

This $\hat{X} \approx dt(\partial_t - \mathcal{L})$ describes an implicit (backward Euler) time integration scheme, known to be more stable for stiff
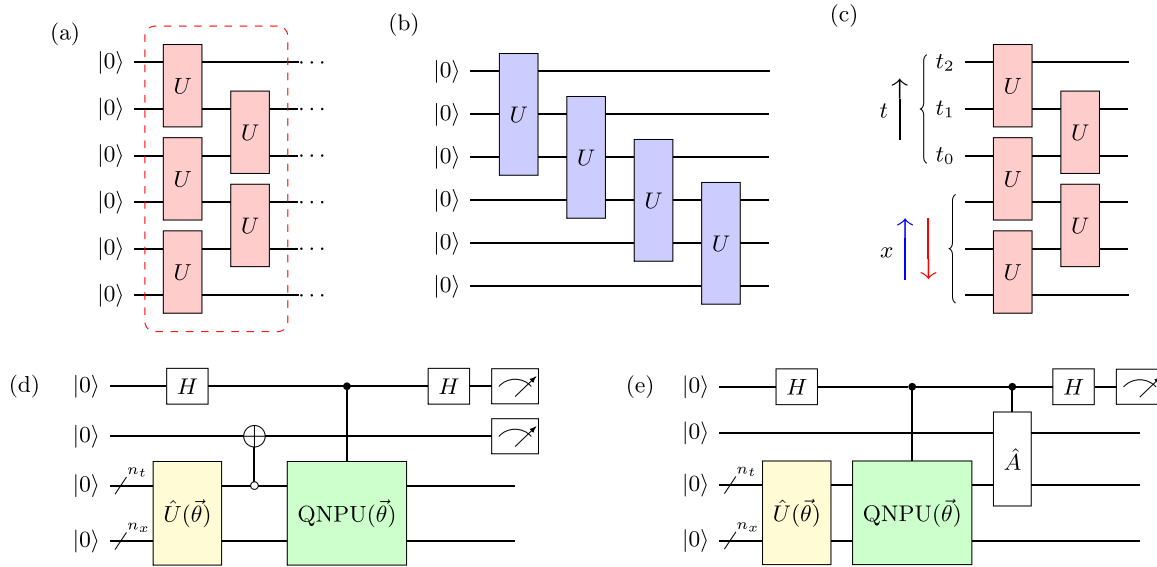
FIG. 2. Parametrizable quantum circuits employed throughout this work. (a) One layer of a brick-wall ansatz, which can be repeated to increase the depth of the circuit. Each unitary U (red) represents a generic two-qubit gate. (b) Generic quantum MPS with bond dimension $\chi = 4$ (Sec. II C) where each unitary U (blue) represents a generic three-qubit gate. (c) Sequential (blue) and reversed space (red) ordering of space and time qubits in the register. The arrows point from the most to the least significant qubit for time and space. [(d), (e)] Example QNPU circuits to calculate the cost function, using the adder circuit $\hat{A}$ [49] for the time shift operator, where $\hat{U}(\vec{\theta})$ is the ansatz circuit and H the Hadamard gate. Of these, (d) is used for the $\hat{C}_1$ term and (e) for the $\hat{C}_2$ term. In these two figures, the notation QNPU($\vec{\theta}$) is used as shorthand for the QNPU itself and any duplicates of the ansatz, which depends on $\vec{\theta}$, encoded on ancilla qubits. See Appendix B for details about the implementation of these circuits.

PDEs than explicit schemes [53]. The product $\hat{X}^\dagger \hat{X}$ in Eq. (5) can be further decomposed into $\hat{C}_1 - \hat{C}_2$ with

$$\hat{C}_1 = \sum_{i=0}^{N_t-1} \hat{I} \otimes |i\rangle\langle i| + \hat{T}^\dagger(-dt)\hat{T}(-dt) \otimes |i+1\rangle\langle i+1|, \quad (8)$$

$$\hat{C}_2 = \sum_{i=0}^{N_t-1} \hat{T}(-dt) \otimes |i\rangle\langle i+1| + \text{H.c.} \quad (9)$$

We now integrate the PDE of Eq. (1) for all time steps by finding the zero-energy ground state of $\hat{H}$. This ground state is given by the *history state*

$$|\Psi\rangle = \frac{1}{\sqrt{N_t+1}} \sum_{i=0}^{N_t} [\hat{T}(dt)]^i |\psi_0\rangle \otimes |i\rangle. \quad (10)$$

We show numerically in Appendix A that the history state is indeed the ground state and that it is nondegenerate. Further details on the derivation of the cost function can also be found there.

B. Evaluation of the Hamiltonian

We evaluate the Feynman-Kitaev Hamiltonian $\langle \hat{H} \rangle$ described in the previous section using the QNPU formalism introduced by Lubasch *et al.* [49]. A QNPU implements nonlinear terms by pointwise multiplication. Lubasch *et al.* presented an efficient implementation of spatial derivatives and periodic boundary conditions using a quantum *adder* circuit [49]. Nonlinear terms such as $f(x,t)\frac{\partial f(x,t)}{\partial x}$ in the Burgers equation (12) are represented by combining these two concepts with a duplicate of the ansatz.

We go beyond the approach of Lubasch *et al.* [49] by including the time register. Because our problem is not periodic in time, we add a time qubit to the register before applying the adder circuit $\hat{A}$ for the time derivative, so the time does not wrap around from $|N_t\rangle$ to $|0\rangle$ [see Fig. 2(e)]. Using QNPUs, we can estimate the expectation value $\langle \hat{H} \rangle$ in a number of circuits that does not depend on the number of discretization points, with a depth linear in the number of space and time qubits. In this way, we ensure the scalability of the QNPU implementation. The number of circuits to realize $\hat{T}$ and its product $\hat{T}^\dagger \hat{T}$ as they appear in Eqs. (8) and (9) depends on the number of terms $\mathcal{L} = \mathcal{A} + \mathcal{B} + \ldots$ in the PDE and the powers of $\mathcal{L}$ in the Taylor expansion of the time evolution operator $\hat{T}$ in Eq. (3). Every product of the terms $\mathcal{A}$, $\mathcal{B}$, etc. in the powers of $\mathcal{L}$ gives rise to a small number of QNPUs; see, for example, Eq. (B3). The number of circuits to implement the Hamiltonian for the Burgers equation depends on the order of the expansion in dt of $\hat{T}$ in $\hat{C}_2$ and $\hat{T}^\dagger \hat{T}$ in $\hat{C}_1$. Using $\hat{T}$ up to first order, which means $\hat{T}^\dagger \hat{T}$ contains up to dt^2, one can implement the Hamiltonian in 18 circuits, as shown in Appendix B. It is possible to reduce this to 11 circuits by cutting off seven circuits encoding terms of order dt^2 in $\hat{T}^\dagger \hat{T}$. Notably, the number of circuits does not depend on the number of time or space qubits.

The $\hat{C}_0$ term enforcing the initial condition can be measured in a single circuit with a swap test [65], for which efficient *destructive* implementations have recently been proposed [66], if one is able to prepare $|\psi_0\rangle$ as a quantum state. While preparing a function in amplitude encoding can require deep circuits in general, most engineering problems have a simple initial state that can be encoded in a shallow circuit on a quantum computer. Alternatively, one

could generate this initial state using dissipative engineering [67,68].

C. Variational ansatz circuits

Throughout this paper, we parametrize the solution using a variational ansatz circuit

$$|\Psi\rangle = \hat{U}(\vec{\theta})|0\rangle, \tag{11}$$

where $\hat{U}(\vec{\theta})$ is a parametrizable quantum circuit of the forms shown in Fig. 2. We consider two different parametrizable quantum circuits as variational ansatz to find the ground state of the Hamiltonian $\hat{H}$. First, we consider a short and dense brick-wall ansatz of various depths, of which a single layer can be seen in Fig. 2(a). The brick-wall ansatz, also referred to as the checkerboard ansatz [46], is especially suited for NISQ devices [69]. It has been used before in the context of the variational quantum eigensolver (VQE) [70] as well as simulations of quantum dynamics [71]. Second, we employ a tensor-inspired quantum matrix product state (MPS) ansatz, as seen in Fig. 2(b). In classical simulations of dynamical systems, matrix product states represent a promising technique to improve the algorithmic efficiency thanks to their efficient representation of multidimensional problems [72,73]. Quantum MPS have been employed successfully for the numerical simulation of one-dimensional systems with relatively small correlations [74]. Our ansatz consists of successive three-qubit unitaries, which corresponds to a sparse representation of a classical MPS with bond dimension $\chi = 4$. By varying the depth of the individual unitaries, we obtain an ansatz with 13, 26, or 39 CNOT gates. This is described in more detail in Appendix C.

To identify an optimal ordering of spacetime qubits that may take advantage of the underlying entanglement structure, we investigate two different ways to order the qubits in the register: a sequential structure with separate time and space registers [blue arrow in Fig. 2(c)], and one where the space register is reversed to have the qubits representing the coarsest structures in space and time next to each other [red arrow in Fig. 2(c)]. This ordering is particularly suitable for the multigrid optimization method discussed below and provides a low-rank tensor representation [64].

To accelerate the minimization procedure and avoid barren plateaus, we initially use small values for coefficients such as the D or β in Eq. (12). We use the Adam optimizer [75] to find the ground state of $\hat{H}$ for this simpler problem. Then, we improve the solution in steps with the L-BFGS-B optimizer [76], as described in Appendix D. In Sec. III C, we discuss a multigrid approach to further improve the scalability, limiting the number of variational parameters.

D. Variational optimization on NISQ devices

We want to implement the nonlinear Hamiltonian (5) on NISQ devices such as IBMQ Ehningen and Quantinuum H1-1. IBMQ Ehningen is a superconducting quantum computer with 27 qubits, based on the IBM Quantum Falcon processor which allows for two-qubit gate infidelities below 10^{-2} and coherence times around $100\,\mu s$ [77]. It features fast computation times, suitable for optimization problems with

many iterations. Quantinuum H1-1 is an ion trap quantum computer with 20 qubits. It features all-to-all connectivity, a two-qubit gate infidelity around 2×10^{-3} [78] and coherence times in the order of seconds [79]. Therefore, complicated circuits are evaluated at high accuracy.

After performing the VQE optimization on the Qiskit statevector simulator [80], we sample the resulting state on both quantum computers. On IBMQ Ehningen, we use readout error mitigation from the M3 library [81] and transpile the circuits using the highest optimization level. This includes dynamical decoupling [82] to avoid decoherence during the idle time of the qubits. As a result, the two-qubit gate infidelity becomes the main source of errors. On Quantinuum H1-1, we run our circuit without error mitigation. The main sources of errors on this machine are qubit depolarization and the two-qubit gate infidelity [79].

We have validated the QNPU implementation for a $2 + 2$-qubit problem on a noiseless simulator. For further benchmarking of our nonlinear Hamiltonian $\hat{H}$, we decompose it into Pauli strings because this is more feasible for a low number of qubits. Nonlinear terms such as $f(x,t)\frac{\partial f(x,t)}{\partial x}$ in Eq. (12) depend on the ansatz state itself. Therefore, we perform a Pauli decomposition of a linearized Hamiltonian at each iteration of the optimization. While this is possible for small systems, the decomposition of an observable into Pauli strings scales exponentially with the number of qubits [83].

III. RESULTS AND DISCUSSION

In Sec. III A, we perform simulations on the Quantinuum and IBMQ quantum computers to show that our ansatz circuits are viable on NISQ devices. In Sec. III B, we investigate the performance of the different ansatz and entanglement structures in the minimization of the Hamiltonian $\hat{H}$ and numerically demonstrate the scaling. Finally, we discuss promising optimization strategies for our hybrid algorithm in Sec. III C and the applicability of our methods beyond variational algorithms in Sec. III D.

A. IBMQ and Quantinuum

We want to solve the Burgers equation,

$$\frac{\partial f(x,t)}{\partial t} = D\frac{\partial^2 f(x,t)}{\partial x^2} - \beta f(x,t)\frac{\partial f(x,t)}{\partial x}, \tag{12}$$

a nonlinear PDE that describes the convection-diffusion of particles in a viscous medium. Here, D is the diffusion coefficient and β determines the strength of the nonlinearity. If one sets $\beta = 0$, the nonlinear term vanishes and the diffusion or heat equation is obtained [84].

First, we study this equation for $3 + 3$ qubits, which corresponds to a spacetime grid of $2^3 = 8$ points in time and $2^3 = 8$ points in space. We use $\beta = 1$ and $D = 0.05$ with periodic boundary conditions, with a time step of $\Delta t = 0.05$. A detailed discussion of the discretization error in spacetime is given in Appendix E. To emphasize the nonlinearity in the form of a shock wave, we choose a Gaussian profile $f(x,t_0) = \exp(-(2\pi x - \pi)^2)$ for $x \in [0,1]$ as initial state. On a noiseless state-vector simulator, we optimize for the VQE solution $|\psi_{\text{VQE}}\rangle$ using a four-layer brick-wall ansatz
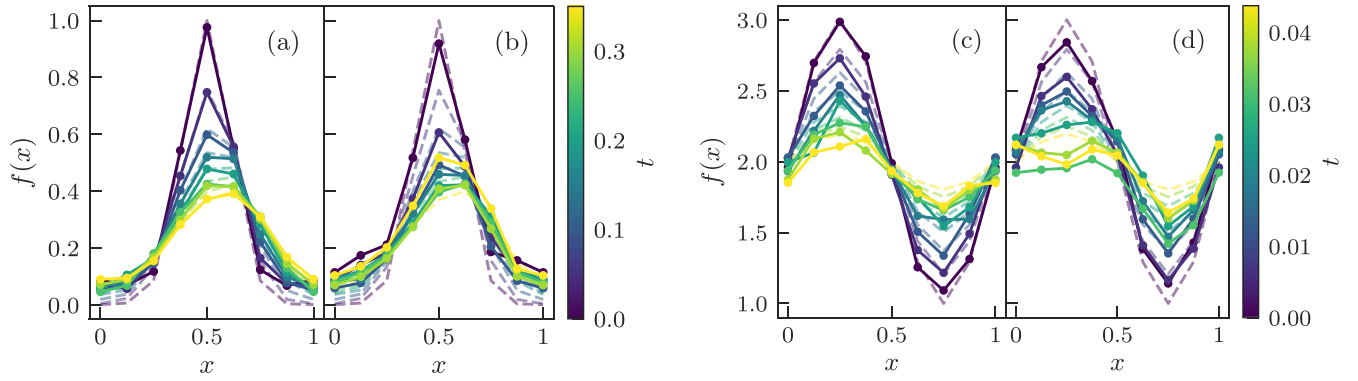
FIG. 3. Solutions to the Burgers and diffusion equation (12) sampled on NISQ machines. We use $3 + 3$ qubits ($2^3 = 8$ points in space and in time) and a reversed space *brick-wall ansatz* (Sec. II C) with four layers (Burgers) or three layers (diffusion). (a) Burgers equation, Quantinuum H1-1. (b) Burgers equation, IBMQ Ehningen. (c) Diffusion equation, Quantinuum H1-1. (d) Diffusion equation, IBMQ Ehningen. The IBMQ Ehningen experiments were run on August 17, 2023, and the Quantinuum H1-1 experiments on September 15 and 20, 2023. The dashed lines correspond to a converged circuit from a noiseless simulator. This noiseless result has a cost $\langle \hat{H} \rangle$ of 2.7×10^{-4} (Burgers) or 4.7×10^{-13} (diffusion) and a very good overlap with a numerical solution [infidelity of 3.3×10^{-4} (Burgers) or 3.2×10^{-7} (diffusion)]. More details can be found in Sec. III A.

(circuit depth 8) with reversed space entanglement structure [Fig. 2(c)]. The cost $\langle \hat{H} \rangle$ of this state is low, 2.7×10^{-4}. The infidelity $1 - \langle \psi_{\text{num}} | \psi_{\text{VQE}} \rangle = 3.3 \times 10^{-4}$, where $|\psi_{\text{num}}\rangle$ is the solution from the ODE solver from the SCIPY library [85], proves that our VQE optimization has converged.

We sample this spacetime state $|\psi_{\text{VQE}}\rangle$ on the IBMQ Ehningen and Quantinuum H1-1 quantum computers [see Figs. 3(a) and 3(b)]. On Quantinuum H1-1, we take 2.5×10^4 shots, whereas on IBMQ Ehningen, we take 2×10^6 shots (in 20 series of 10^5 shots), which has a similar execution time. We observe that the Quantinuum computer is able to reproduce the rightwards shift of the wave, characteristic for the Burgers equation, whereas the IBM result is more affected by noise, especially in the yellow curve representing the last time point. The analogous simulation with a quantum MPS ansatz is plotted in Figs. 15(a) and 15(b), where the Quantinuum result stems from an emulator which models the noise of the Quantinuum H1-1 machine.

Next, we discard the nonlinear term from Eq. (12) by setting $\beta = 0$ and $D = 1$ with a time step of $\Delta t = 0.00625$, which gives us the one-dimensional diffusion or heat equation. As initial condition at $t_0 = 0$, we consider a sinusoidal profile $f(x, t_0) = 2 + \sin(2\pi x)$ for $x \in [0, 1]$, with periodic boundary conditions. We have chosen the constant 2 instead of 1 for the experiments on both quantum computers because we find that, in the presence of noise, we measure some counts for every bitstring, including the ones that correspond to values close to zero. This makes it difficult to distinguish values close to zero in amplitude encoding. This shift does not change the dynamics of the diffusion equation. For comparison, we show the unshifted simulations in Fig. 16 in Appendix E.

We calculate a VQE solution on the Qiskit state-vector simulator using a three-layer brick-wall ansatz (circuit depth 6). The cost $\langle \hat{H} \rangle$ of this state is 4.7×10^{-13}. To verify that this state with a low cost indeed matches the desired solution, we calculate the infidelity with regards to a solution from numerical integration $1 - \langle \psi_{\text{num}} | \psi_{\text{VQE}} \rangle = 3.2 \times 10^{-7}$. These cost and infidelity values indicate that the upwards shift of the

initial condition does not prevent convergence of the VQE. In Figs. 3(c) and 3(d), we sample it on the IBMQ Ehningen and Quantinuum H1-1 quantum computers. The analogous simulation using a quantum MPS ansatz is depicted in Figs. 15(c) and 15(d).

B. Circuit depth requirements

We show the achieved cost of the VQE solutions for all ansatz and entanglement structures in Fig. 4. Note that the initial condition for the diffusion equation is $f(x, t_0) = 1 + \sin(2\pi x)$ here. The circuit depth in this figure is defined by counting the consecutive layers of CNOT gates; for the brick-wall ansatz with six qubits, this is the number of CNOT gates divided by 2.5 because one brick-wall layer contains five blocks, but has a depth of 2. For the quantum MPS ansatz, it is equal to the number of CNOT gates. We find in Fig. 4(c) that our diffusion problem is very well described using a shallow reversed space ansatz with three or four brick-wall layers. The low cost which we observe for these depths, and the subsequent increase in cost for deeper circuits, indicates that these ansatzes can be optimized in less iterations than the deeper ones. The sequential ordering [see Fig. 2(c)] requires a few more layers and is more sensitive to random initializations. For the nonlinear problem in Fig. 4(a), we see that the deeper ansatzes provide an advantage in expressibility, with only small differences between the entanglement structures. The quantum MPS ansatz requires a slightly higher number of CNOT gates, but is less sensitive to the ordering of the qubits in the register, as seen in Figs. 4(b) and 4(d). We conclude that there are particularly efficient (shallow) representations of our diffusion problem using the reversed space entanglement structure; however, the cost for these shallow circuits does show a significant spread depending on the random initial condition, indicating there are local minima to the cost function. This is consistent with literature about the trainability of shallow circuits [86,87]. Our classical MPS calculations indicate that the entanglement entropy in the ansatz and the necessary ansatz depth depend significantly on the
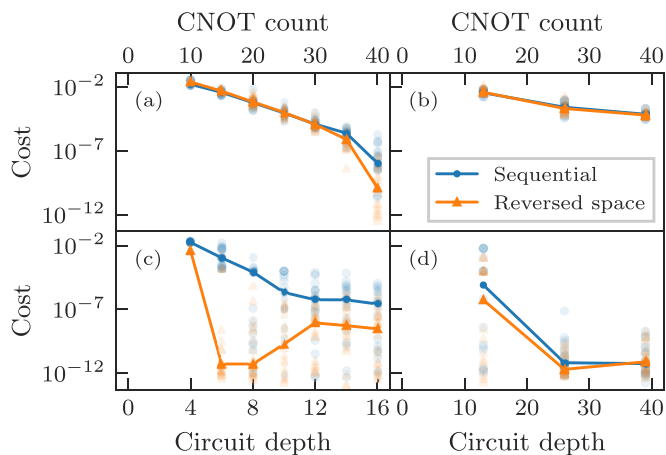
FIG. 4. Cost $\langle \hat{H} \rangle$ with $\hat{H}$ from Eq. (5) as a function of the circuit depth on a noiseless simulator for $3 + 3$ qubits. (a) Burgers equation, brick wall; (b) Burgers equation, quantum MPS; (c) diffusion, brick wall; (d) diffusion, quantum MPS. The data points in (a) and (c) represent two to eight layers, those in (b) and (d) three quantum MPS ansatzes with 13 to 39 CNOT gates as detailed in Appendix C. The top and bottom axes indicate that the CNOT count and circuit depth are equivalent for the quantum MPS ansatz but different for brick-wall circuits. The lines represent the median of the cost resulting from 20 different random initializations, which are plotted as scattered points. The maximum number of iterations is 4×2500 as described in Appendix D.
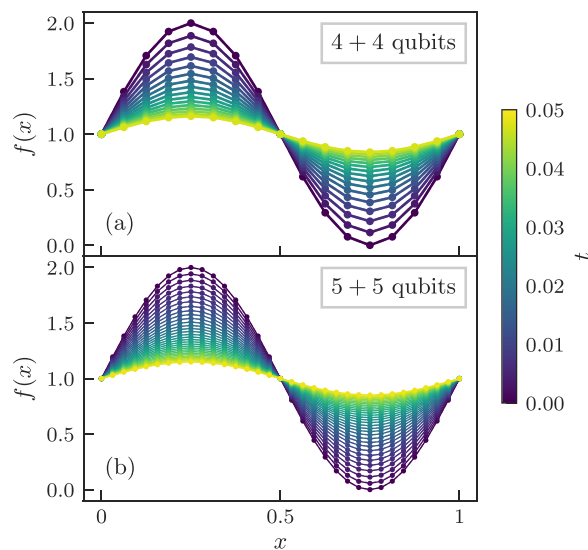

FIG. 5. Results for the diffusion equation [Eq. (12) with $\beta = 0$] obtained from a noiseless simulator for different numbers of qubits. (a) Obtained solution for $4 + 4$ qubits ($2^4 = 16$ points in space and time) with $D = 1$ and $\Delta t = 1/320$. Reversed space brick-wall ansatz with four layers, with a cost of 1.6×10^{-9} and an infidelity of 9.3×10^{-8}. (b) $5 + 5$ qubits ($2^5 = 32$ points in space and time) with $D = 1$ and $\Delta t = 1/640$. Reversed space brick-wall ansatz with six layers, with a cost of 7.0×10^{-7} and an infidelity of 2.9×10^{-7}.

smoothness of the initial condition and the spectral content of its evolution.

We study the scaling of our algorithm. To this aim, we analyze the required circuit depth increasing the system size from six to ten qubits ($n_x + n_t = 3 + 3, 4 + 4$ and $5 + 5$). This gives us a measure for the necessary coherence time required to run any of these circuits on a real quantum computer. The results for a four-layer brick-wall ansatz with $4 + 4$ qubits and a six-layer brick-wall ansatz with $5 + 5$ qubits are shown in Fig. 5. The costs for brick-wall and quantum MPS ansatzes of different depths are shown in Fig. 6. We normalize them by dividing by the energy E_1 of the first excited state of the Hamiltonian for the respective number of qubits, such that we can use the distance to the first excited state as a measure of the quality of the solution, irrespective of the number of qubits. We observe in this figure that the low cost solutions for $3 + 3$ qubits require an ansatz with more parameters than the dimension of the Hilbert space, whereas the solutions from Fig. 5 do not. From this analysis, we conclude that a crossover exists between coarse grids requiring an over-parametrized ansatz for convergence (six qubits) and finer grids (eight qubits or more) that can be efficiently encoded with less variational parameters than the dimension of the corresponding Hilbert space.

C. Barren plateaus and multigrid optimization

Variational cost functions commonly exhibit barren plateaus: gradients of the cost function exponentially vanish as the number of qubits increases when optimizing using random initial parameters [63]. Figure 4(c) indicates that it

can be beneficial to optimize for a smaller depth first and then increase the depth if necessary. This motivates our approach of first solving a simpler problem with a small D and/or β and then ramping up these physical constants (see Appendix D). We calculate the variance of the gradient for different random initializations in Appendix F to evaluate this strategy. We do not observe any barren plateaus for the system sizes studied here.

Now, we analyze a multigrid approach [64] for further improvements: First, the solution is found for a coarse grid using a combination of the Adam and L-BFGS optimizers, as described in Appendix D; second, optimization on a finer grid is started with the optimal parameters of the coarse grid. We use the reversed space entanglement structure. In this way, it becomes possible to add new time or space qubits while maintaining the ansatz structure for the coarser grid, as shown in Fig. 7(a). This strategy is inspired by the fact that the entanglement entropy is greatest for the tensors encoding the coarsest partitions of the grid in a matrix product state representation [64].

To stay close to the ground state and avoid detrimental jumps in the cost when transitioning from a coarse to a fine grid, we want to implement a step profile as shown in Fig. 7(b). We achieve this by slightly modifying the brick-wall ansatz so every 2-qubit unitary contains two CNOT gates instead of one, and six rotations after every CNOT, duplicating the unit seen in Fig. 11(d) in Appendix C. There, we also show that this modification of the ansatz does not negatively affect the required circuit depth because less layers are required if every layer contains twice the number of gates. By initializing all of these angles to zero, except for the very last rotations on the new qubits, the new qubits representing fine structures
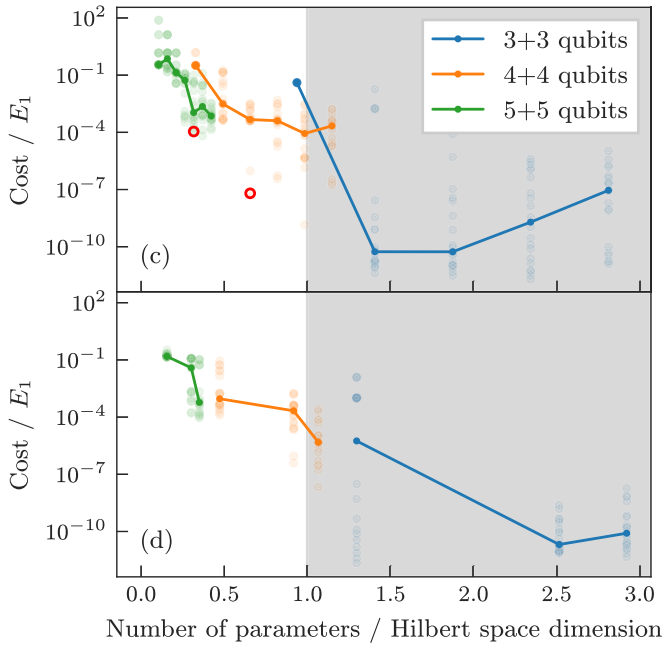
FIG. 6. Cost $\langle \hat{H} \rangle$ with $\hat{H}$ from Eq. (5) achieved for the diffusion equation with different numbers of qubits, divided by the energy E_1 of the first excited state, plotted as a function of number of parameters divided by Hilbert space dimension (64 for six qubits, 256 for eight qubits, 1024 for ten qubits.) The shaded area indicates where the number of parameters exceeds the dimension of the Hilbert space. The dots in the background represent the 20 individual runs for which the line shows the median. (a) Brick-wall ansatz of two to eight layers. The two states from Fig. 5 are marked with red circles. (b) Three quantum MPS circuits, as detailed in Appendix C. A reversed space entanglement structure is used in both cases. We have increased the number of iterations from 2500 to 5000 for every value of D for 8 qubits, and to 10 000 for every value of D for 10 qubits.

in space and time are not immediately entangled with the existing coarse ones. Importantly, the rotations at the end of the ansatz are initialized such that they represent a Hadamard gate. This results in a step profile after initialization of the fine grid ansatz as shown in Fig. 7(b). We slightly shift the initial condition along the x axis to make the generated step profile a better match. In Fig. 8(a), we depict the cost after transitioning from a coarse to a fine grid. Our step profile algorithm results in a smaller cost function than when initializing the new qubits randomly or in a zero state. It may be possible to engineer other initial states, such that they come even closer to the desired solution.

After adding the new qubits, we optimize for the new parameters and the blocks affecting the qubits x_2 and t_2 directly adjacent to the new ones, as shown in Fig. 7(a). This guides the optimizer towards the solution of the refined problem. After this optimization round, we add the blocks affecting x_1 and t_1 and optimize again. Finally, we optimize all blocks. We compare the cost as a function of the number of iterations of this method with a direct optimization for $4 + 4$ qubits in Fig. 8(b), either using Adam only, without ramping up D on the way, or using Adam and L-BFGS. This figure shows that the multigrid strategy, with a transition from $3 + 3$ to $4 + 4$

qubits, needs a similar number of iterations as a direct optimization for $4 + 4$ qubits with Adam and L-BFGS, whereby the multigrid method is faster because it optimizes less parameters. It is also clear that either method provides a significant advantage compared to solving with Adam only, which ends up in a local minimum almost two orders of magnitude higher than the other methods in this example. We conclude that the multigrid strategy described here is a promising method to solve problems on fine grids.

D. Beyond variational algorithms

The nonlinear Feynman-Kitaev Hamiltonian is applicable beyond the variational methods discussed in this paper. Besides the VQE, on which our methods are based, there are many other variational and nonvariational techniques for finding the ground state. Imaginary time evolution (ITE), which we use in Appendix A 2 to study the stability of the solution, has a quantum analog that is applicable on NISQ devices [88]. This quantum imaginary time evolution, which has previously been used to solve differential equations [89], is still variational. A nonvariational strategy can be found in adiabatic quantum computing [90], which has similarities with the optimization strategy we used of increasing D and/or β in steps. An implementation on NISQ devices has been recently shown [91] and promising proposals exist to further compress the corresponding quantum circuits [92]. As such, we expect that this can also be used to solve our ground-state problem.

IV. CONCLUSIONS

In this paper, we present an extension of the Feynman-Kitaev formalism, originally developed for quantum dynamics, that is tailored to the integration of arbitrary PDEs with nonlinearities. We provide proof-of-principle calculations that demonstrate that nonlinear dissipative processes are well reproduced within this framework. Here, the full spacetime solution of the system can be retrieved in a single optimization routine of an appropriate cost function, which prevents the accumulation of errors of iterative time marching schemes. We see this as a necessary step to solve PDEs on quantum computers, as stability conditions (see Appendix A 2 for details) require that an increase in spatial resolution goes together with an increase in time resolution. Our results indicate that solutions to nonlinear PDEs can be obtained with shallow circuits using either a brick-wall ansatz with a few layers or a sparse formulation of MPSs as a quantum circuit. We adopt a multigrid strategy where the ansatz is adaptively optimized for finer and finer solutions of the discretized PDE, thereby making the spacetime quantum solver potentially scalable in the number of qubits. We show how to adapt the ansatz so it can be expanded with additional qubits while keeping the coarser result. We expect this multigrid strategy to become more powerful as the resolution, and thereby smoothness of the function, increases. We adapt the QNPU [49] to evaluate the cost function. To implement time evolution, we show how to implement derivatives without periodic boundary conditions. Notably, the number of circuits to be measured is only dependent on the number of (non)linear terms in the PDE and the desired accuracy of the time evolution operator and not on the number of qubits, and the depth of the circuits is linear in
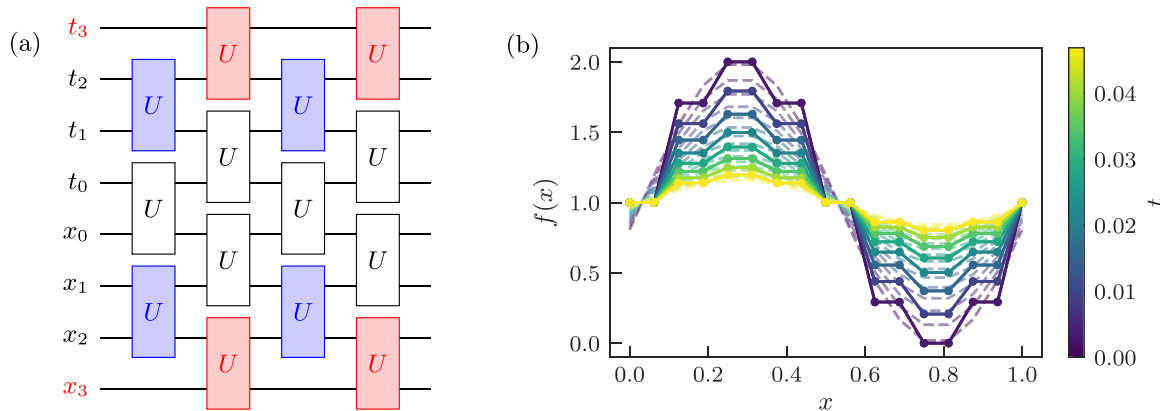
FIG. 7. Multigrid optimization strategy. (a) The reversed space ansatz of two brick-wall layers is expanded from six to eight qubits. The red blocks correspond to the new spacetime qubits (x_3, t_3) and the optimization is initially only performed for the angles within the red and blue blocks, while the unitaries of the qubits encoding the coarser structures (x_0, t_0) are fixed to the parameters of the converged circuit with six qubits. (b) Converged $3 + 3$-qubit result expanded to $4 + 4$ qubits, showing the "step" profile that is created by initializing the new qubits with a Hadamard gate. The dashed lines correspond to the solution obtained via numerical integration.

the number of qubits. Depending on the nature of the problem and the available quantum resources, it can thus be favorable to use a first- or second-order propagator and a fine grid in time instead of a higher order integration scheme on a coarser time grid, which is commonly done in classical simulations [53]. Alternatively, one could start with a low order solution and use that as input for a higher order optimization with more circuits.

Current classical methods for nonlinear PDEs in fluid dynamics allow around 10^6 spatial points [1,72]. This amounts to about 10^2 points in every dimension for a 3D grid and is equivalent to 20 spatial qubits. Hence we believe that only around 100 qubits in total (including time qubits and duplicates of the ansatz for nonlinearities) would be necessary to go beyond current classical simulations. Our methods are also relevant in classical computing, as recent developments in tensor network representations of nonlinear problems have shown that it is possible to devise efficient classical algorithms inspired by

quantum algorithms for PDEs [72,73,93]. Finally, we remark that our methodology does not require a variational approach for finding the ground state of the Hamiltonian, and alternative techniques can be readily employed.
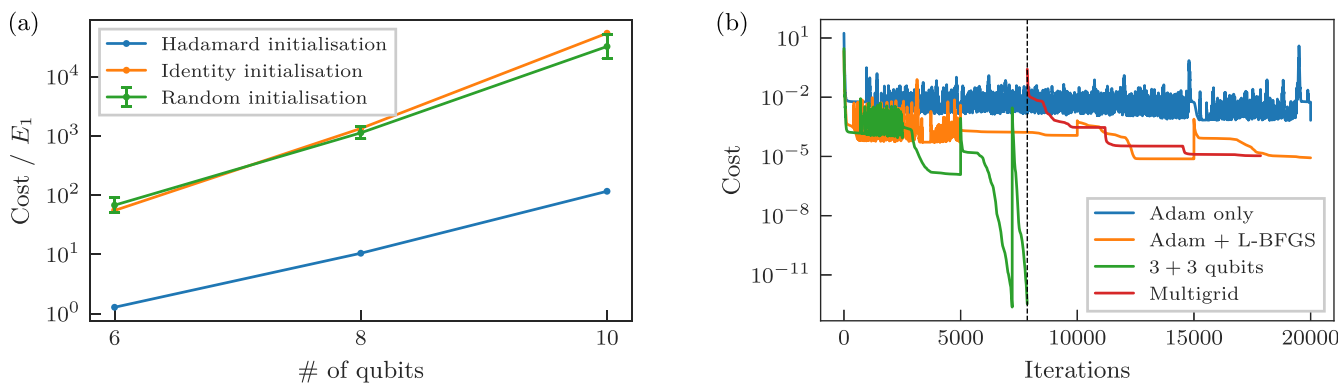
FIG. 8. (a) Increase in cost caused by adding more qubits to a converged solution of the diffusion equation, normalized by the energy of the first excited state E_1, for different initializations of the new qubits. (b) Comparison of the number of iterations versus cost for a $4 + 4$-qubit problem using different optimization procedures, showing that the multigrid ansatz has a performance comparable to the Adam + L-BFGS procedure. For all three lines, the best result out of 20 random initializations was chosen. The green line describes an Adam + L-BFGS optimization for $3 + 3$ qubits, followed by a transition to $4 + 4$ qubits marked by the dashed black line, and subsequent L-BFGS optimization in red. This illustrates our multigrid strategy. The green and red lines for the multigrid strategy are shorter because the $3 + 3$-qubit simulation considered only took 7857 iterations instead of the total 10 000.

APPENDIX A: DERIVATION OF THE COST FUNCTION

We start with the general form of the differential equation as given by Eq. (1), where the differential operator $\mathcal{L}$ (linear or nonlinear) is the generator of the underlying dynamics. For simplicity, let us assume that $\mathcal{L}$ is time independent. We encode the values of the function in a quantum state using the amplitude encoding of Eq. (4). This way we can define a quantum version of Eq. (1) as

$$\frac{d}{dt}|\Psi\rangle = \mathcal{L}|\Psi\rangle, \tag{A1}$$

where the amplitudes $\psi(x,t) = \mathcal{N}f(x,t)$ of the state $|\Psi\rangle$ represent the function values up to a certain normalization constant $\mathcal{N}$. A general solution to this equation can be written in terms of the time evolution operator (propagator),

$$\vec{\psi}(t) = \hat{T}(t)\vec{\psi}_0, \tag{A2}$$

where $\vec{\psi}(t)$ is the state in space for a specific time t. We use this propagator to define an implicit integration scheme

$$dt(\partial_t - \mathcal{L})\vec{\psi}(t_i) \approx \hat{T}(-dt)\vec{\psi}(t_{i+1}) - \vec{\psi}(t_i), \tag{A3}$$

where the vector notation $\vec{\psi}(t_i)$ is used for the state in space at time step $|i\rangle$ such that $\vec{\psi}(t_i) = \langle i|\Psi\rangle$. For a time-dependent $\mathcal{L}$, one would replace the backwards propagator $\hat{T}(-dt)$ by its time-dependent equivalent $\hat{T}(t + dt, t)$. If $\hat{T}$ is approximated to first order in dt, this is equivalent to a backward Euler integration scheme. We turn this into a spacetime representation by multiplying on the left with $|i\rangle$ and taking the sum over all i, which leads to

$$dt(\partial_t - \mathcal{L})|\Psi\rangle \approx \sum_{i=0}^{N_t-1}(\hat{T}(-dt) \otimes |i\rangle\langle i+1| - \hat{I} \otimes |i\rangle\langle i|)|\Psi\rangle$$

$$\equiv \hat{X}|\Psi\rangle, \tag{A4}$$

where $\hat{X}$ is as defined in Eq. (7). This means that $\hat{X} \approx dt(\partial_t - \mathcal{L})$. We multiply with the adjoint to get a Hermitian operator $\hat{X}^\dagger\hat{X} \approx dt^2\|\partial_t - \mathcal{L}\|^2$, and use this operator to define a ground-state problem for Eq. (1). Combined with the $\hat{C}_0$ term for the initial condition $|\psi(t=0)\rangle = |\psi_0\rangle$, one obtains the Hamiltonian in Eq. (5).

For a better physical understanding, we rewrite $\hat{X}^\dagger\hat{X}$ as

$$\hat{X}^\dagger\hat{X} = \left(\sum_{j=0}^{N_t-1}\hat{T}^\dagger(-dt) \otimes |j+1\rangle\langle j| - \hat{I} \otimes |j\rangle\langle j|\right)\left(\sum_{i=0}^{N_t-1}\hat{T}(-dt) \otimes |i\rangle\langle i+1| - \hat{I} \otimes |i\rangle\langle i|\right), \tag{A5a}$$

$$= \sum_{i=0}^{N_t-1}(\hat{I} \otimes |i\rangle\langle i| + \hat{T}^\dagger(-dt)\hat{T}(-dt) \otimes |i+1\rangle\langle i+1|) - \sum_{i=0}^{N_t-1}(\hat{T}^\dagger(-dt) \otimes |i+1\rangle\langle i| + \hat{T}(-dt) \otimes |i\rangle\langle i+1|). \tag{A5b}$$

We define the two terms of equation (A5b) as $\hat{C}_1$ and $\hat{C}_2$. We can then alternatively write Eq. (5) as

$$\hat{H} = c_0\hat{C}_0 + \hat{C}_1 - \hat{C}_2, \tag{A6}$$

with the three terms given by

$$\hat{C}_0 = [\hat{I} - |\psi_0\rangle\langle\psi_0|] \otimes |0\rangle\langle 0|, \tag{A7a}$$

$$\hat{C}_1 = \sum_{i=0}^{N_t-1}[\hat{I} \otimes |i\rangle\langle i| + \hat{T}^\dagger(-dt)\hat{T}(-dt) \otimes |i+1\rangle\langle i+1|], \tag{A7b}$$

$$\hat{C}_2 = \sum_{i=0}^{N_t-1}[\hat{T}^\dagger(-dt) \otimes |i+1\rangle\langle i| + \hat{T}(-dt) \otimes |i\rangle\langle i+1|]. \tag{A7c}$$

Here $\hat{C}_0$ ensures the initial condition is kept, $\hat{C}_1$ ensures that all time points are present in the solution, and $\hat{C}_2$ ensures that the steps between them correspond to the application of the differential operator $\mathcal{L}$. For quantum systems, it is possible to derive this cost function from a time-embedded discrete variational principle that is equivalent to the Dirac-Frenkel-McLachlan time-dependent variational principle in the limit of infinitesimal time [57].

We show here that the history state (10) is an eigenstate of this Hamiltonian with an eigenvalue of 0. To this end, we calculate the action of all three terms of Eq. (A7) on this state.

The first term gives us

$$\hat{C}_0|\Psi\rangle = \frac{1}{\sqrt{N_t+1}}[\hat{I} - |\psi_0\rangle\langle\psi_0|]|\psi_0\rangle \otimes |0\rangle$$

$$= 0|\Psi\rangle. \tag{A8}$$

The second and third terms give us

$$\hat{C}_1|\Psi\rangle = \frac{1}{\sqrt{N_t+1}}\sum_{i=0}^{N_t-1}[\hat{T}^i(dt)|\psi_0\rangle \otimes |i\rangle$$

$$+ \hat{T}^\dagger(-dt)\hat{T}(-dt)\hat{T}^{i+1}(dt)|\psi_0\rangle \otimes |i+1\rangle], \tag{A9a}$$

$$\hat{C}_2|\Psi\rangle = \frac{1}{\sqrt{N_t+1}}\sum_{i=0}^{N_t-1}[\hat{T}(-dt)\hat{T}^{i+1}(dt)|\psi_0\rangle \otimes |i\rangle$$

$$+ \hat{T}^\dagger(-dt)\hat{T}^i(dt)|\psi_0\rangle \otimes |i+1\rangle]. \tag{A9b}$$

Using the property $\hat{T}(-dt)\hat{T}(dt) = \hat{I}$, one can see that $\hat{C}_1|\Psi\rangle = \hat{C}_2|\Psi\rangle$. Therefore,

$$\hat{H}|\Psi\rangle = (c_0\hat{C}_0 + \hat{C}_1 - \hat{C}_2)|\Psi\rangle = 0|\Psi\rangle, \tag{A10}$$

which means that $|\Psi\rangle$ is an eigenstate of $\hat{H}$ with eigenvalue 0. We numerically show in Appendix A 2 that this is the ground state, and that there is a gap with the first eigenstate which makes this ground state unique.
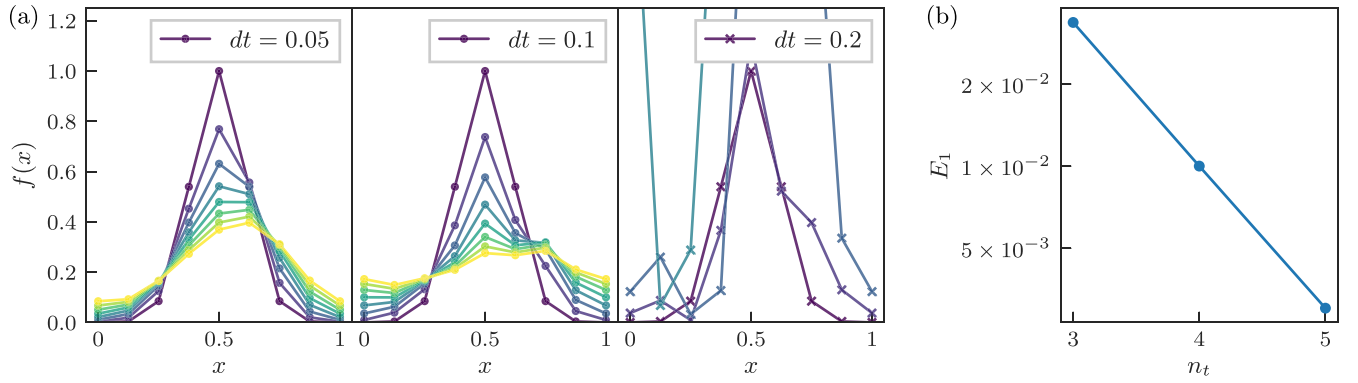
FIG. 9. (a) Ground-state profiles of $\hat{H}$ for the Burgers equation ($D = 0.05$, $\beta = 1$) calculated with imaginary time evolution (ITE). Converged solutions can be achieved for $dt = 0.05$ and $dt = 0.1$, but the breakdown of stable solutions to the discretized PDE with $dt = 0.2$ results in unphysical states. (b) Energies of the first excited state E_1 (A12) of $\hat{H}$ as a function of the number of time qubits n_t for a fixed time step of $dt = 0.05$ and a spatial register of $n_x = 3$ qubits.

1. Note on implicit integration

Using the implicit integration scheme of Eq. (A3), the $\hat{C}_1$ term (A7b) contains a term with $\hat{T}^\dagger(-dt)\hat{T}(-dt)$. If one performs the derivation with an explicit integration scheme, there is a term $\hat{T}^\dagger(dt)\hat{T}(dt)$ instead. For quantum dynamics [57,59], the propagator is unitary, which means that both terms become an identity. In this case, there is no difference between an implicit and an explicit integration scheme. For our nonunitary $\hat{T}$, we have chosen to use an implicit scheme.

2. Uniqueness and stability

To produce stable solutions in a finite-difference integration scheme, the distance traveled by the solution in one time step must be less than the distance between two points in the grid [53]. For the Burgers equation (12), this stability condition translates into the requirement that $D\,dt/dx^2 < 1/2$. With our model parameters ($D = 0.05$, $\beta = 1$), we expect stable solutions for $dt \lesssim 0.15$ for three space qubits ($dx = 1/8$). To obtain stable solutions from the ground-state problem given by the Feynman-Kitaev Hamiltonian $\hat{H}$, it is necessary that the ground state of $\hat{H}$ is unique. This is ensured if there is an energy gap between the ground state and the first excited state. In this section, we show numerical simulations with ITE using the MPNUM library [94] to evaluate the existence of this gap under different model parameters.

The ground and first excited states of $\hat{H}$ must satisfy the conditions

$$\hat{H}(|\Psi_0\rangle)|\Psi_0\rangle = E_0|\Psi_0\rangle, \tag{A11}$$

$$\hat{H}(|\Psi_1\rangle)|\Psi_1\rangle = E_1|\Psi_1\rangle. \tag{A12}$$

Due to the nonlinear nature of the Hamiltonian, the states $|\Psi_0\rangle$, $|\Psi_1\rangle$ may not be orthogonal to each other. Furthermore, the lowest eigenvector $|E_0'\rangle$ of $\hat{H}(|\phi\rangle)$ for an arbitrary state $|\phi\rangle$ may not coincide with $|\Psi_0\rangle$.

The ITE method is based on the fact that the operator $e^{-t\hat{H}}$ becomes a projector onto the ground state of $\hat{H}$ in the limit $t \to \infty$. For an arbitrary, initial state $|\phi\rangle$ with a non-negligible overlap $|\langle\Psi_0|\phi\rangle| > 0$ with the ground state, the ITE operator

$e^{-t\hat{H}}$ acting on $|\phi\rangle$ converges to the ground state

$$|\Psi_0\rangle \propto \lim_{t\to\infty}(e^{-t\hat{H}})|\phi\rangle. \tag{A13}$$

Starting from a random state $|\phi\rangle$, we choose an Euler integration scheme for ITE by iterating
 (1) $|\phi'(t+\tau)\rangle = |\phi(t)\rangle - \tau\hat{H}|\phi(t)\rangle$,
 (2) $|\phi(t+\tau)\rangle = |\phi'(t+\tau)\rangle/\sqrt{\langle\phi'(t+\tau)|\phi'(t+\tau)\rangle}$,
until convergence is reached for $\hat{H}(|\Psi_0\rangle)|\Psi_0\rangle = E_0|\Psi_0\rangle$ with $E_0 \lesssim 10^{-14}$. We normalize the state after each time step because the norm $\langle\phi(t)|\phi(t)\rangle$ diverges exponentially with t.

We turn the calculation of the first excited state $|\Psi_1\rangle$ (A12) into another ground-state problem. This can be achieved by performing ITE with a modified Hamiltonian $\hat{H}'$, whose ground-state energy is raised above the energy of its first excited state. One may be tempted to use the "shifted" Hamiltonian

$$\hat{H}' = \hat{H} + c|\Psi_0\rangle\langle\Psi_0|, \tag{A14}$$

where c is a positive constant and $|\Psi_0\rangle$ is the ground state calculated with ITE as described above. However, for nonlinear Hamiltonians, the ground and excited states as defined in Eqs. (A11), (A12) may not be orthogonal ($\langle\Psi_0|\Psi_1\rangle \neq 0$) and the shift $c|\Psi_0\rangle\langle\Psi_0|$ in Eq. (A14) not only modifies the lowest energy subspace of $\hat{H}(|\phi\rangle)$ but also perturbs the rest of the spectrum. To avoid this, we perform ITE simulations with the modified Hamiltonian

$$\hat{H}'(|\phi\rangle) = \hat{H}(|\phi\rangle) + c|E_0'\rangle\langle E_0'|, \tag{A15}$$

where the second term raises the energy of the lowest eigenstate $|E_0'\rangle$ of $\hat{H}(|\phi(t)\rangle)$, which changes with every ITE iteration. This way, the ITE is directed towards a state that solves Eq. (A12). The only requisite for the constant c is that it should be larger than the second eigenvalue E_1' of $\hat{H}(|\phi\rangle)$ for any $|\phi\rangle$. We use this method to calculate the energy gap of the nonlinear Hamiltonian $\hat{H}$ associated with the Burgers equation (12).

In Fig. 9(a), we calculate the ground state for the Burgers equation (12) using the same parameters as in the main text ($D = 0.05$, $\beta = 1$), for a spacetime grid with $3 + 3$ qubits ($n_x = n_t = 3$) and three different time steps ($dt = 0.05$, 0.1, 0.2). Our results demonstrate a stable solution for
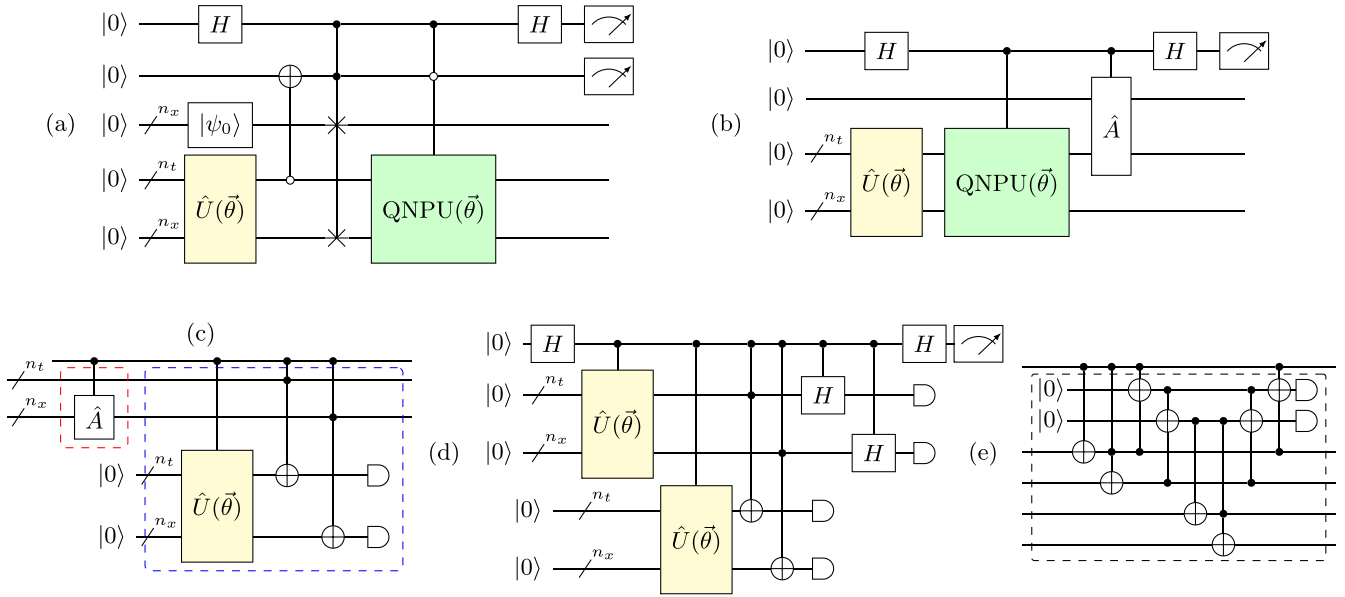
FIG. 10. (a) Circuit to evaluate $\hat{C}_0$ and the $\hat{T}^\dagger\hat{T}$ term in $\hat{C}_1$ simultaneously, using an ancilla qubit to determine whether to apply a swap test for $\hat{C}_0$ if the time is $|0\rangle$ or a QNPU for $\hat{C}_1$ otherwise. (b) Circuit to evaluate $\hat{C}_2$, with an additional qubit for the time-shift operator to make it nonperiodic. This qubit represents the most significant bit of the adder. In these two figures, the notation $\mathrm{QNPU}(\vec{\theta})$ is used as shorthand for the QNPU itself and any duplicates of the ansatz, which depends on $\vec{\theta}$, encoded on ancilla qubits. (c) Building blocks of QNPU circuits, with the $\hat{A}$ operator red and the pointwise multiplication $\hat{c}$ in blue. The multicontrolled gates need to be understood as to be applied pairwise: qubit 0 on one time or space register controls qubit 0 on the other register, and so forth for every time and space qubit. (d) Circuit calculating $\Re(\sum_i \psi_i^2)/\sqrt{N_t N_x}$ with $\Re$ the real value. (e) Adder circuit for four qubits [49].

the cases with $dt = 0.05$ and $dt = 0.1$. However, for $dt = 0.2$, the ITE trajectories converge to an unphysical state, which is in accordance with the stability condition. In this scenario, the coefficients ψ_{ij} (4) corresponding to later times become overrepresented. In Fig. 9(b), we show E_1 as a function of the number of time qubits for a fixed time step of $dt = 0.05$ and a spatial register of $n_x = 3$ qubits. We see that for this time step, there is a gap for any number of time qubits, and that it decreases with the number of qubits for a constant dt as the number of eigenstates of the system increases. Our simulations with the modified Hamiltonian (A15) for $dt = 0.2$ show a constantly decreasing energy, without converging to a state that satisfies the eigenvalue condition (A12). For the states $|\xi\rangle$ which we found on the way, $\hat{H}(|\xi\rangle)$ presents a degenerate ground state. This suggests that there is indeed a closing gap for our modified Feynman-Kitaev Hamiltonian once we violate the stability condition.

APPENDIX B: EVALUATION OF THE COST FUNCTION

To measure the expectation value $\langle\hat{H}\rangle$ on a quantum computer, it needs to be expressed using unitary circuits. We have produced the numerical results in this paper by decomposing the cost function into Pauli strings. This is not a scalable approach because the number of these strings, in general, scales exponentially with the number of qubits [83]. When using a Pauli string decomposition for nonlinear problems, it is also necessary to decompose the cost function repeatedly during the optimization process to update the linearized function. In this section, we show that there is a decomposition into a constant number of circuits using QNPUs [49], if one can

efficiently create the initial state $|\psi_0\rangle$ on a quantum computer. We show that the number of two- or three-qubit gates in these circuits scales linearly with the number of qubits. This is important for the scalability of the algorithm.

The QNPU approach allows us to separate space and time in the cost function, as the propagator $\hat{T}$ is independent of time. The time part of the $\hat{C}_1$ term (A7b) is then given by the sums over $|i\rangle\langle i|$ and $|i+1\rangle\langle i+1|$. In general, the sum over all projectors $|i\rangle\langle i|$ is an identity. However, as these sums run until $N_t - 1$, they represent an identity with one missing entry, given by

$$\sum_{i=0}^{N_t-1} |i\rangle\langle i| = \hat{I} - |N_t\rangle\langle N_t|, \tag{B1a}$$

$$\sum_{i=0}^{N_t-1} |i+1\rangle\langle i+1| = \hat{I} - |0\rangle\langle 0|. \tag{B1b}$$

We can then rewrite the $\hat{C}_1$ term of the cost function as

$$\hat{C}_1 = \hat{I} \otimes (\hat{I} - |N_t\rangle\langle N_t|) + \hat{T}^\dagger\hat{T} \otimes (\hat{I} - |0\rangle\langle 0|). \tag{B2}$$

This reduces the complexity of the time part of $\hat{C}_1$ to that of a single projector for both terms. The first term is determined in one circuit by sampling the ansatz and counting the results where the time is not $|N_t\rangle$. The second term can be implemented efficiently by using an ancilla qubit to indicate whether the time was $|0\rangle$ before performing a Hadamard test to measure $\hat{T}^\dagger\hat{T}$, then postselecting the results for times other than $|0\rangle$. In general, the probability of this approaches 1 as n_t is increased. This circuit is shown in Fig. 2(d) in the main text.
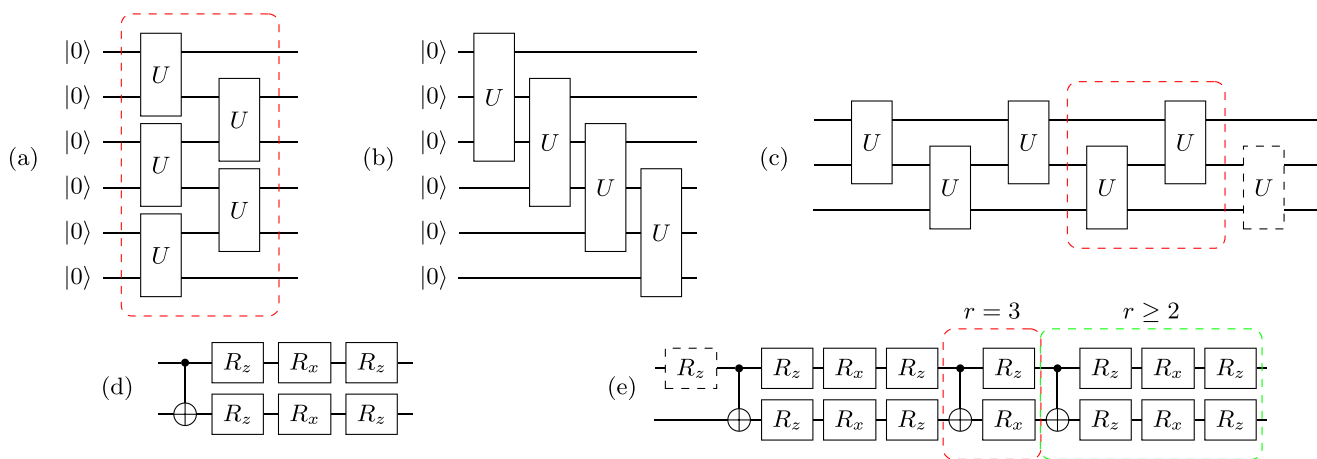
FIG. 11. Schematics of a brick-wall ansatz with a single layer and a generic MPS ansatz of bond dimension $\chi = 4$. (a) One layer of the brick-wall ansatz. (b) Generic quantum MPS with $\chi = 4$. (c) Three-qubit unitary, with the part taken out for the sparse $\chi = 4$ quantum MPS shown in red. The dashed unitary is left out if it overlaps with one of the next block. (d) Two-qubit unitary with one CNOT for the brick-wall ansatz. (e) Two-qubit unitary for the MPS ansatz for a number of CNOTs per block of $r = 1...3$. The dashed gate is left out if it overlaps with an R_z gate of the previous block.

We propose that one can combine it with a measurement of $\hat{C}_0$ (described below) as shown in Fig. 10(a).

The $\hat{C}_2$ term (A7c) concerns the shift operator $|i + 1\rangle\langle i|$ and its adjoint. This term can be encoded efficiently using an *adder* circuit, such as the one by Lubasch *et al.* [49], with a depth that scales linearly with the number of qubits. For small numbers of qubits, one could use the one by Sato *et al.* [44] instead. The shift operator in $\hat{C}_2$ is not periodic, thus it is necessary to filter out the $|N_t\rangle\langle 0|$ term. We achieve this by adding a qubit to our time register which we initialize to zero; instead of $|N_t\rangle\langle 0|$, there will be an $|N_t\rangle\langle N_t + 1|$ term, for which the Hadamard test will measure an overlap of zero. This can be seen in Fig. 10(b).

The spatial part of the $\hat{C}_1$ and $\hat{C}_2$ terms is given by the propagator $\hat{T}$, as well as $\hat{T}^\dagger\hat{T}$. For the second-order Taylor approximation used in our main results, $\hat{T}$ contains terms up to the square of the Laplace operator, and $\hat{T}^\dagger\hat{T}$ up to its fourth power. In general, one can formulate a QNPU as given by Lubasch *et al.* [49] to calculate any of the terms in the equation. In the propagator $\hat{T}$ to second order, one has the Laplacian $\frac{\partial^2}{\partial x^2}$, the nonlinear term $f\frac{\partial}{\partial x}$, as well as the squares of both and the product of the two with each other. For $\hat{T}^\dagger\hat{T}$, one also needs to implement the third and fourth powers. The first derivative $\frac{\partial}{\partial x}$ and the Laplacian $\frac{\partial^2}{\partial x^2}$ can be produced by applying an adder circuit in space, implementing the finite difference approximations of either derivative as $(\hat{A} - \hat{I})/\Delta x$ and $(\hat{A} - 2\hat{I} + \hat{A}^\dagger)/(\Delta x)^2$, respectively. For the nonlinearity, one can create a duplicate of the ansatz and do a pointwise multiplication. The differential operator $\mathcal{L}$ from Eq. (1) can then be calculated using

$$\mathcal{L} = \frac{1}{(\Delta x)^2}(\hat{A} - 2\hat{I} + \hat{A}^\dagger) - \beta\mathcal{M}\frac{1}{\Delta x}(\hat{A} - \hat{I})\hat{c}, \quad \text{(B3)}$$

where $\hat{A}$ represents the adder circuit and $\hat{c}$ the multiplication with a duplicate of the ansatz. One can see this as a diagonal operator with the coefficients of the state on the diagonal. $\mathcal{M}$ is a constant with which the nonlinear term needs to be multiplied to turn the amplitudes of the quantum state into the function values $f(x, t)$. The value of this can be determined from the norm $\mathcal{N}$ of the initial state divided by the probability to measure $|0\rangle$ time.

In this QNPU formulation, a complex-valued equation (A1) is being solved, instead of the real-valued Burgers equation, because the pointwise multiplication $\hat{c}$ multiplies with the complex amplitudes of the state. To reproduce the real-valued equation, the state $|\Psi\rangle$ has to be real-valued. This can be ensured using a real-valued ansatz [62] or by adding a penalty term to the cost function, such as $\hat{C}_3 = c_3[1 - \Re(\sum_i \psi_i^2)]$, with $\Re$ the real value, which is zero if all amplitudes are real. A circuit that can be used to measure $\Re(\sum_i \psi_i^2)$ up to a factor $\sqrt{N_t N_x}$ is shown in Fig. 10(d).

In the representation of Eq. (B3), the propagator $\hat{T}$ can be expressed up to first order in four circuits because one can take $\hat{A}$ and its adjoint together. The operator $\hat{T}^\dagger\hat{T}$ from the $\hat{C}_1$ can be expressed in 11 circuits up to first order. The term $\hat{I} \otimes (\hat{I} - |N_t\rangle\langle N_t|)$ from Eq. (B2) can be measured in a single circuit, and the same holds for ψ^2 from $\hat{C}_3$ when using a complex-valued ansatz. One additional circuit is necessary to determine the probability to measure the zero time state, which is needed for the normalization constant $\mathcal{M}$. This brings us to a total of 18 circuits for the Burgers equation with a first-order propagator.

The coefficients in front of all of these circuits depend on the Taylor series of the exponential of $\Delta t \mathcal{L}$ from Eq. (3), with $\mathcal{L}$ from Eq. (B3). For the diffusion term, this means that the coefficients scale with powers of $\Delta t/(\Delta x)^2$, while for the nonlinear term, they scale with $\Delta t/\Delta x$. It is important to make sure that these coefficients do not increase exponentially with the number of space or time qubits. A good scaling can be achieved when $\Delta t \propto \Delta x^2$, i.e., adding two time qubits to the system for every added space qubit. Interestingly, this coincides with the stability condition for a PDE that is first order in time and second order in space, which requires that $\Delta t/(\Delta x)^2$ is bounded [53].
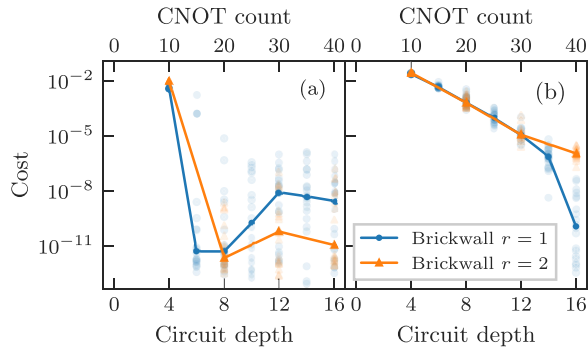
FIG. 12. Comparison of the cost $\langle \hat{H} \rangle$ achieved using the brick-wall ansatz with $r = 1$ or 2 CNOTs per block, with a reversed space entanglement structure, and a brick-wall ansatz with two to eight layers for $r = 1$ and one to four layers for $r = 2$. The lines represent the median of the cost resulting from 20 different random initializations, which are plotted as scattered points. (a) Diffusion. (b) Burgers equation.

1. Initial state

The time part of $\hat{C}_0$ is the projector $|0\rangle\langle 0|$, and the space part is an identity minus a projector $\hat{I} - |\psi_0\rangle\langle\psi_0|$. If one can prepare the initial state $|\psi_0\rangle$ as a quantum circuit, it is possible to do a swap test to measure the overlap between the ansatz and this initial state. This results in a linear number of controlled swap gates. One can make this more hardware efficient using a destructive swap test [66], which only requires CNOT gates instead of controlled swap gates. However, it is necessary to sample the time register and filter out the time states that are not $|0\rangle$. Because $\hat{C}_0$ only measures time $|0\rangle$, and the $\hat{T}^\dagger\hat{T}$ term in $\hat{C}_1$ only measures times other than $|0\rangle$, we propose a circuit where one conditionally measures $\hat{C}_0$ or $\hat{C}_1$, based on the state of the time qubits. This is shown in Fig. 10(a).

This measurement can be simplified if the initial state is a basis vector of the computational basis or if a basis transformation can be applied such that the initial state can be identified by a single bitstring. In this case, one can sample all qubits and determine the probability of finding time $|0\rangle$ and a spatial state that is not $|\psi_0\rangle$. In particular, the $\hat{C}_0$ term for a sine wave as the initial condition could be measured using a quantum Fourier transform on the space qubits instead of a swap test.

APPENDIX C: ADDITIONAL ANSATZ CIRCUITS

1. Brick-wall ansatz

The brick-wall ansatz [Fig. 11(a)] consists of layers of two-qubit unitaries. For the results in the main text, we used a two-qubit unitary consisting of six rotations and one CNOT gate, as given in Fig. 11(d). We have also run the simulations with two CNOTs per block by duplicating this two-qubit unitary, thus creating layers that are twice as deep. The achieved cost for both diffusion and the Burgers equation can be seen in Fig. 12. We see here that there is no notable disadvantage in most cases by using less, but deeper, layers.
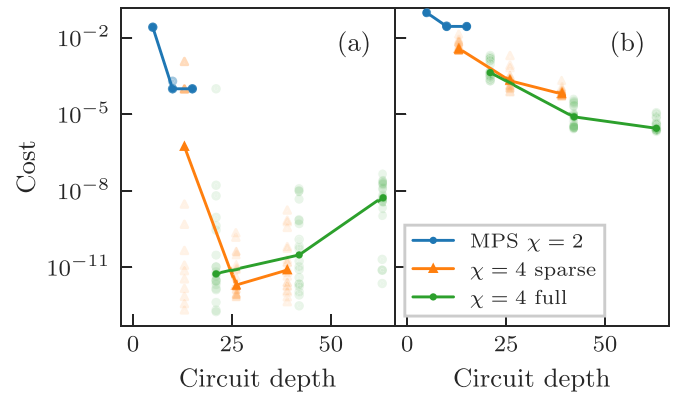

FIG. 13. Comparison of the cost $\langle \hat{H} \rangle$ achieved using different MPS ansatz circuits with $\chi = 2$ as well as $\chi = 4$, with three versions of the two-qubit unitary (three data points on every line) and two versions of the three-qubit unitary (orange and green lines), with a reversed space entanglement structure. The lines represent the median of the cost resulting from 20 different random initializations, which are plotted as scattered points. (a) Diffusion. (b) Burgers equation.

2. MPS ansatz

We consider an MPS ansatz of bond dimension $\chi = 4$ as given in Fig. 11(b). In general, each of the three-qubit unitaries in this ansatz circuit can be decomposed into six two-qubit unitaries [95]. Of these six two-qubit unitaries, the last one overlaps with the first one of the next three-qubit unitary so it is left out, except at the end of the circuit. This means that a structure with five two-qubit unitaries is obtained. To limit the circuit depth, we use a sparse MPS with only three two-qubit unitaries, as shown in Fig. 11(c). In this circuit, we also vary the complexity of the two-qubit unitary. A generic two-qubit unitary can be composed of 15 rotations and three CNOT gates [66,96]. We use three different depths of this circuit as shown in Fig. 11(e). In Fig. 13, we present the cost as a function of circuit depth and compare with the full $\chi = 4$ quantum MPS (with five two-qubit unitaries of varying depths), as well as a $\chi = 2$ quantum MPS, to verify that the
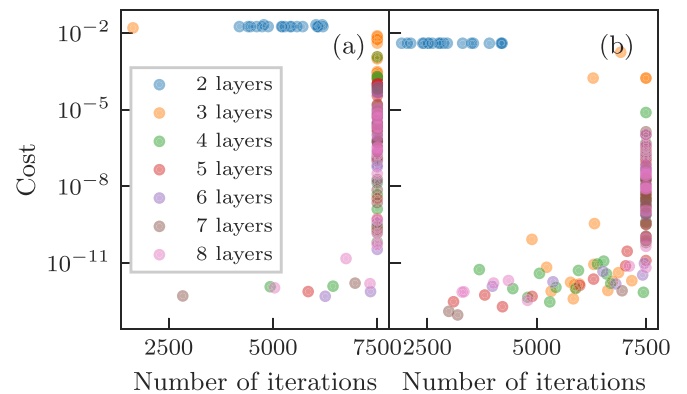

FIG. 14. Cost $\langle \hat{H} \rangle$ versus achieved number of L-BFGS-B iterations for all 20 random initial conditions for the diffusion problem using a brick-wall ansatz with two to eight layers. (a) Sequential entanglement structure. (b) Reversed space.
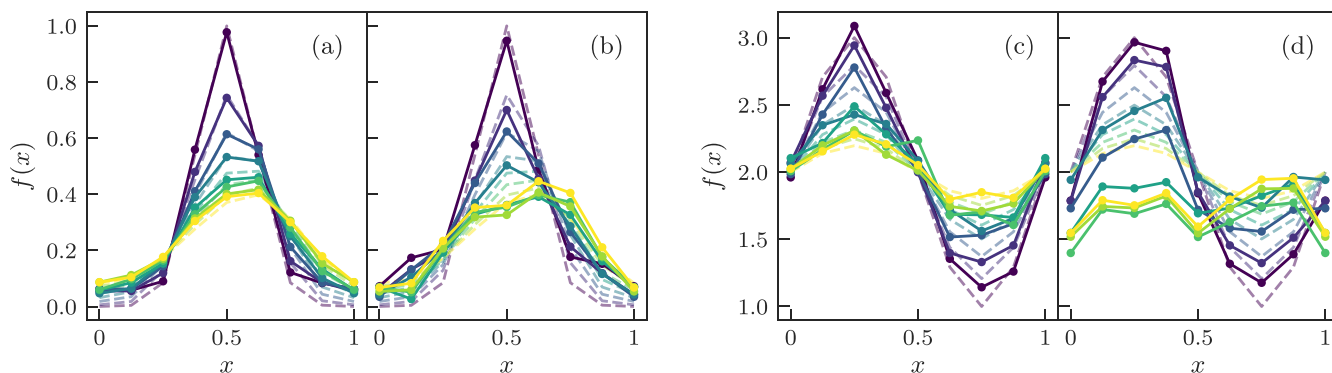
FIG. 15. Solutions to the Burgers and diffusion equation (12) sampled on NISQ machines. We use $3 + 3$ qubits ($2^3 = 8$ points in space and in time) and a reversed space quantum MPS ansatz (Sec. II C) with 26 CNOT gates (Burgers) or 13 CNOT gates (diffusion). (a) Burgers equation, Quantinuum H1-1E emulator. (b) Burgers equation, IBMQ Ehningen. (c) Diffusion equation, Quantinuum H1-1E emulator. (d) Diffusion equation, IBMQ Ehningen. The dashed lines correspond to a converged circuit from a noiseless simulator. This noiseless result overlaps with numerical results with an infidelity of 1.89×10^{-4} (Burgers) or 3.2×10^{-7} (diffusion). More details can be found in Sec. III A.

sparse $\chi = 4$ quantum MPS provides a good balance between circuit depth and expressibility.

APPENDIX D: OPTIMIZATION PROTOCOL

For both the diffusion and Burgers equation, we use an optimization procedure where we start with a configuration with a small D and $\beta = 0$ and then ramp up either D or β. This is equivalent to Barison *et al.* [59], who instead increased the time step in steps to avoid barren plateaus.

We optimize this initial configuration with 2500 steps of the Adam optimizer [75], and all following configurations with at most 2500 steps of the L-BFGS-B optimizer [76]. For the diffusion equation, we start with a diffusion coefficient of 1/8. After optimizing with Adam, we increase D in steps to 1/4, 1/2, and, finally, 1, optimizing with L-BFGS-B in between. For the Burgers equation (12), we start with a diffusion coefficient of $D = 0.05$ and a nonlinearity coefficient $\beta = 0$, so the optimization with Adam can be performed on a linear equation, which is less computationally expensive. We then increase β in steps to 1/8, 1/4, 1/2, and, finally, 1, optimizing

with L-BFGS-B in between. This means that we end with $\beta = 20D$, which is sufficient to observe nonlinear effects.

We chose these optimizers because the Adam optimizer allows us to avoid local minima that can be encountered when starting from a random initial state, whereas the L-BFGS-B optimizer is more local, and thus well suited in cases where the initial state is already reasonably close to the solution. For either optimizer, we use automatic differentiation [97] to determine the gradients of the cost function with respect to all ansatz parameters.

For the L-BFGS-B optimization, we consider the L-BFGS-B optimization converged if the difference in cost between two successive iterations is less than 10 times the machine precision. We choose a very low value here to be able to benchmark the cost function, which means that the maximum of 2500 iterations per value of D is exhausted for most runs of the linear problem.

We run the above protocols for both types of ansatz, with different depths for the brick-wall ansatz and different bond dimensions for the quantum MPS. We run each simulation 20 times with random initial conditions. We plot the cost for each random initialization as a function of the number of
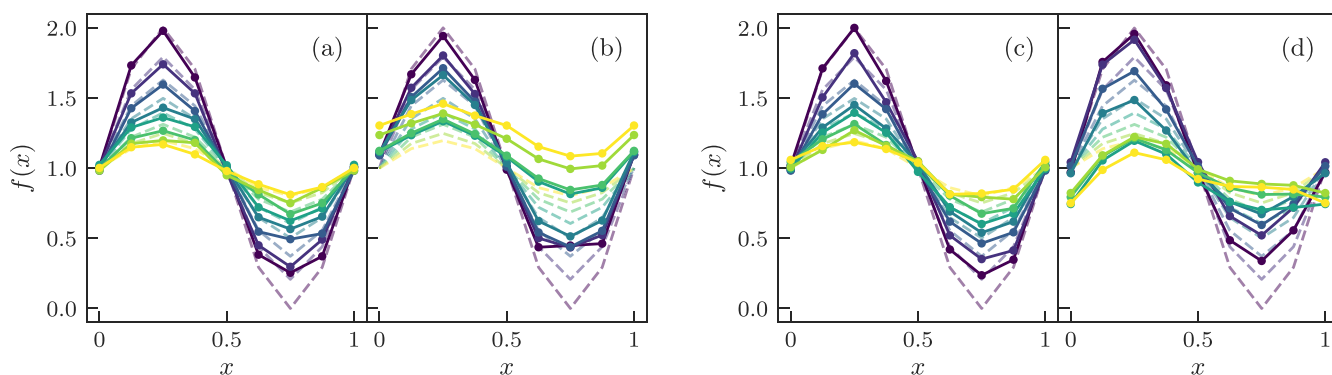

FIG. 16. Solutions to the diffusion equation [Eq. (12) with $\beta = 0$], without a shift in the y coordinate, sampled on NISQ machines. We use $3 + 3$ qubits using [(a), (b)] a brick-wall ansatz of three layers or [(c), (d)] a quantum MPS ansatz of 13 CNOT gates, and sample the result on [(a), (c)] Quantinuum H1-1E emulator and [(b), (d)] IBMQ Ehningen. The dashed lines correspond to the state-vector simulation, which corresponds to results from numerical integration with an infidelity of 1.09×10^{-6} for both cases.
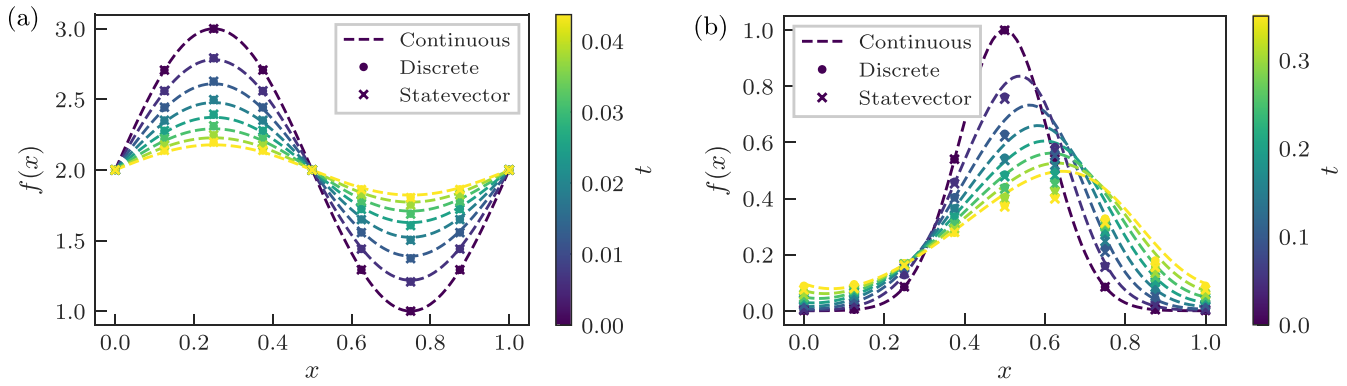
FIG. 17. Comparison of the solutions from a noiseless simulator with $3+3$ qubits ($2^3 = 8$ points in space and in time) to a classical simulation on a fine grid for (a) diffusion using a brick-wall ansatz with three layers and (b) Burgers equation using a brick-wall ansatz with four layers.

L-BFGS-B iterations in Fig. 14. We can see here that the simulations for shallow circuits converge around a cost of 10^{-2}, while the deeper circuits either exhaust the maximum number of iterations or converge at a cost below 10^{-10}, which is close to machine precision. Especially for the reversed space structure, there are many runs that converge at a low number of iterations for all numbers of layers $\geqslant 3$.

For shallow circuits, we find that the cost landscape is irregular, and thus a measurement protocol using different random initializations, then taking the one that leads to the lowest cost, is necessary. With deep circuits, we find that the optimization will eventually converge if the number of iterations is increased sufficiently, regardless of the initial condition. As we are focusing on shallow circuits for NISQ devices, we decided to present results from random initializations in this paper.

APPENDIX E: ADDITIONAL IBMQ AND QUANTINUUM RESULTS

In the main text, we have shown the final state resulting from noiseless optimization sampled on the IBMQ and Quantinuum machines in Fig. 3. In that figure, we used a brick-wall ansatz, and the initial condition for the diffusion

equation was given by $f(x, t_0) = 2 + \sin(2\pi x)$. Here we show similar results with a quantum MPS ansatz in Fig. 15. We also show results for the diffusion in Fig. 16 for both ansatz types without the upward shift in the y coordinates, using the initial condition $f(x, t_0) = 1 + \sin(2\pi x)$. We can see here that both quantum computers have trouble representing amplitudes close to zero. We see in these figures that the lowest amplitudes are better represented when using a shift, but the overall precision goes down for both quantum computers.

It should be noted that the IBMQ results have been produced using readout error mitigation from the M3 library [81], as well as dynamical decoupling [82] to avoid decoherence during the idle time of the qubits, which is especially relevant for the MPS ansatz. As these methods have not been applied to the results from the Quantinuum platform, there is still room for improvement here. In particular, one could use an ansatz that uses the ZZPhase two-qubit gate defined by $\text{ZZPhase}(\theta) = \exp\left(-i\frac{\pi\theta}{2}(Z \otimes Z)\right)$, which has an error depending on the phase angle, instead of the CNOT gate which fully entangles two qubits and is not a native gate of the Quantinuum platform.

In Fig. 17, we show the same state-vector result that was shown as a dashed line in Fig. 3, now combined with a numerical solution on a very fine grid. We see that the error
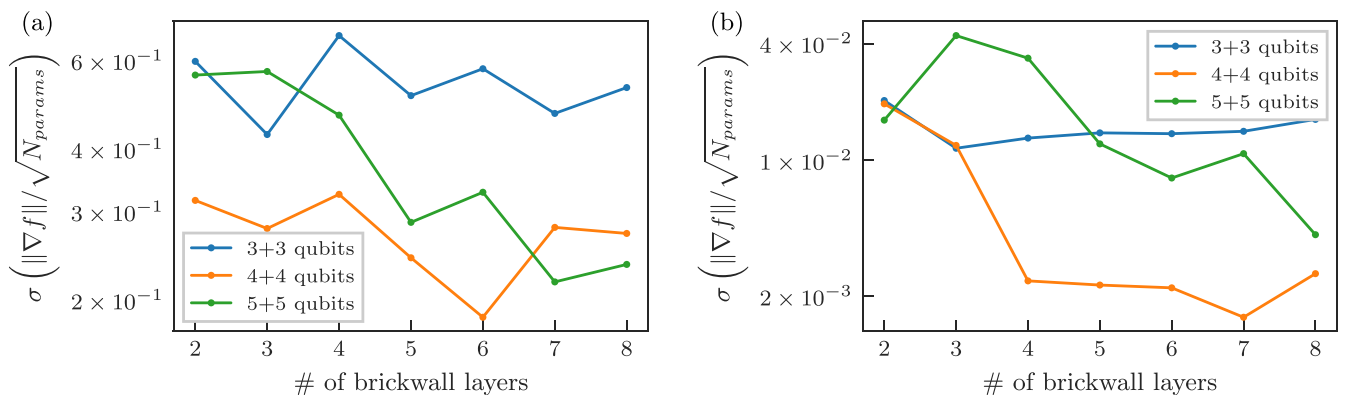

FIG. 18. Standard deviation of the gradient norm of the cost function for the diffusion equation with respect to the ansatz parameters, divided by the square root of the number of ansatz parameters, for 20 random initial conditions. (a) We consider the beginning of the first optimization step for $D = 0.125$ and random initialization. (b) We consider the beginning of the last optimization step for $D = 1$, right after the converged optimization for $D = 0.5$.
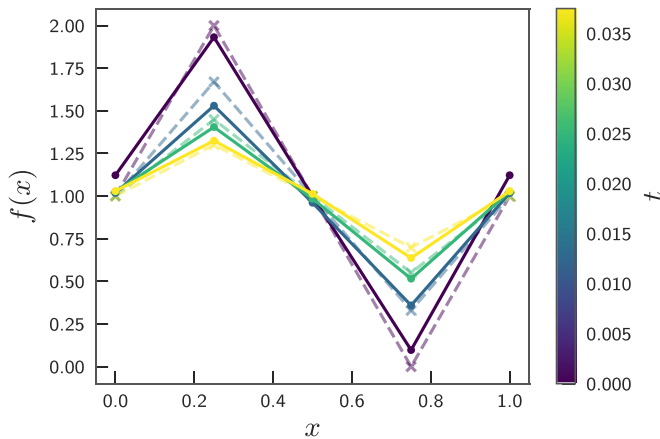
FIG. 19. Result of noisy optimization of the diffusion equation on $2 + 2$ qubits. The optimization was done using the IBMQ Montreal noise model and the result has been plotted without noise. The dashed lines correspond to the solution obtained via numerical integration.

caused by the quantum algorithm is significantly smaller than the discretization error.

APPENDIX F: BARREN PLATEAUS

The presence of barren plateaus is generally indicated by an exponential decrease of the variance of the gradient of the cost function, calculated for different random initializations, as a function of the number of qubits [63]. It has been shown that there is a relation between barren plateaus and ansatz depth [98] or ansatz expressibility [99] (and a

thorough understanding exists for quantum tensor networks [100,101] such as the quantum MPS considered in this paper), especially for local cost functions. Our cost function is not entirely local because the shift operators in space and time affect all space or time qubits, which means that one would expect barren plateaus for shallow (or less expressive) ansatz circuits as well. For this reason, we calculate the gradient of the cost function during the optimization process for the diffusion equation, using different numbers of qubits and different ansatz depths.

In Fig. 18, we show the results for two different stages of the optimization: after initializing randomly and after ramping up the diffusion coefficient from $D = 0.5$ to $D = 1$ (which is the last time that D is increased in the optimization procedure). Instead of taking the standard deviation of one component of the gradient, we have taken the standard deviation of the gradient norm divided by the square root of the number of parameters to be more rigorous. In this figure, there is no indication of an exponential decrease with the number of qubits.

APPENDIX G: NOISY VQE

We perform a preliminary study of the effect of noise on the VQE optimization procedure by optimizing a $2 + 2$-qubit system on a noisy simulator with the IBMQ Montreal noise model (another quantum computer with the same 27-qubit chip as IBMQ Ehningen.) We obtained the result from Fig. 19 after less than 300 iterations. To highlight the effect of noise on the optimization procedure, we have sampled the result state without noise. We see that, despite the noise, optimization is still successful and a state close to the desired solution is found.

[1] M. Mani and A. J. Dorgan, A perspective on the state of aerospace computational fluid dynamics technology, Annu. Rev. Fluid Mech. **55**, 431 (2023).

[2] F. Porté-Agel, M. Bastankhah, and S. Shamsoddin, Wind-turbine and wind-farm flows: A review, Boundary-Layer Meteorol. **174**, 1 (2020).

[3] M. Schammer, B. Horstmann, and A. Latz, Theory of transport in highly concentrated electrolytes, J. Electrochem. Soc. **168**, 026511 (2021).

[4] L. J. Bolay, T. Schmitt, S. Hein, O. S. Mendoza-Hernandez, E. Hosono, D. Asakura, K. Kinoshita, H. Matsuda, M. Umeda, Y. Sone *et al.*, Microstructure-resolved degradation simulation of lithium-ion batteries in space applications, J. Power Sources Adv. **14**, 100083 (2022).

[5] E. Tadmor, A review of numerical methods for nonlinear partial differential equations, Bull. Amer. Math. Soc. **49**, 507 (2012).

[6] P. Givi, A. J. Daley, D. Mavriplis, and M. Malik, Quantum speedup for aeroscience and engineering, AIAA J. **58**, 3715 (2020).

[7] Y. Cao, A. Papageorgiou, I. Petras, J. Traub, and S. Kais, Quantum algorithm and circuit design solving the Poisson equation, New J. Phys. **15**, 013021 (2013).

[8] D. W. Berry, High-order quantum algorithm for solving linear differential equations, J. Phys. A **47**, 105301 (2014).

[9] A. Montanaro and S. Pallister, Quantum algorithms and the finite element method, Phys. Rev. A **93**, 032324 (2016).

[10] D. W. Berry, A. M. Childs, A. Ostrander, and G. Wang, Quantum algorithm for linear differential equations with exponentially improved dependence on precision, Commun. Math. Phys. **356**, 1057 (2017).

[11] P. C. S. Costa, S. Jordan, and A. Ostrander, Quantum algorithm for simulating the wave equation, Phys. Rev. A **99**, 012323 (2019).

[12] A. M. Childs, J.-P. Liu, and A. Ostrander, High-precision quantum algorithms for partial differential equations, Quantum **5**, 574 (2021).

[13] K. K. Saha, W. Robson, C. Howington, I.-S. Suh, Z. Wang, and J. Nabrzyski, Advancing algorithm to scale and accurately solve quantum Poisson equation on near-term quantum hardware, arXiv:2210.16668.

[14] L. Cui, Z. Wu, and H. Xiang, Quantum radial basis function method for the Poisson equation, J. Phys. A **56**, 225303 (2023).

[15] S. S. Bharadwaj and K. R. Sreenivasan, Hybrid quantum algorithms for flow problems, Proc. Natl. Acad. Sci. USA **120**, e2311014120 (2023).

[16] J. Ingelmann, S. S. Bharadwaj, P. Pfeffer, K. R. Sreenivasan, and J. Schumacher, Two quantum algorithms for solving the one-dimensional advection-diffusion equation, Comput. Fluids **281**, 106369 (2024).

[17] S. Lloyd, G. De Palma, C. Gokler, B. Kiani, Z.-W. Liu, M. Marvian, F. Tennie, and T. Palmer, Quantum algorithm for nonlinear differential equations, arXiv:2011.06571.

[18] J.-P. Liu, H. Ø. Kolden, H. K. Krovi, N. F. Loureiro, K. Trivisa, and A. M. Childs, Efficient quantum algorithm for dissipative nonlinear differential equations, Proc. Natl. Acad. Sci. USA **118**, e2026805118 (2021).

[19] S. Jin and N. Liu, Quantum algorithms for computing observables of nonlinear partial differential equations, arXiv:2202.07834.

[20] L. Lapworth, A hybrid quantum-classical CFD methodology with benchmark HHL solutions, arXiv:2206.00419.

[21] S. Jin, N. Liu, and Y. Yu, Time complexity analysis of quantum algorithms via linear representations for nonlinear ordinary and partial differential equations, J. Comput. Phys. **487**, 112149 (2023).

[22] S. Wiesner, Simulations of many-body quantum systems by a quantum computer, arXiv:quant-ph/9603028.

[23] C. Zalka, Simulating quantum systems on a quantum computer, Proc. R. Soc. Lond. A **454**, 313 (1998).

[24] J. M. Arrazola, T. Kalajdzievski, C. Weedbrook, and S. Lloyd, Quantum algorithm for nonhomogeneous linear partial differential equations, Phys. Rev. A **100**, 032306 (2019).

[25] S. Jin, N. Liu, and Y. Yu, Quantum simulation of partial differential equations via Schrödingerisation, arXiv:2212.13969.

[26] S. Jin, N. Liu, and Y. Yu, Quantum simulation of partial differential equations via Schrödingerisation: Technical details, arXiv:2212.14703.

[27] L. Wright, C. Mc Keever, J. T. First, R. Johnston, J. Tillay, S. Chaney, M. Rosenkranz, and M. Lubasch, Noisy intermediate-scale quantum simulation of the one-dimensional wave equation, arXiv:2402.19247.

[28] G. Xu, A. J. Daley, P. Givi, and R. D. Somma, Turbulent mixing simulation via a quantum algorithm, AIAA J. **56**, 687 (2018).

[29] G. Xu, A. J. Daley, P. Givi, and R. D. Somma, Quantum algorithm for the computation of the reactant conversion rate in homogeneous turbulence, Combust. Theory Model. **23**, 1090 (2019).

[30] F. Gaitan, Finding flows of a Navier–Stokes fluid through quantum computing, npj Quantum Inf. **6**, 61 (2020).

[31] F. Gaitan, Finding solutions of the Navier–Stokes equations through quantum computing—recent progress, a generalization, and next steps forward, Adv. Quantum Technol. **4**, 2100055 (2021).

[32] F. Oz, R. K. S. S. Vuppala, K. Kara, and F. Gaitan, Solving Burgers' equation with quantum computing, Quant. Inf. Process. **21**, 30 (2022).

[33] Y. Suzuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto, Amplitude estimation without phase estimation, Quant. Inf. Process. **19**, 75 (2020).

[34] P. Rao, K. Yu, H. Lim, D. Jin, and D. Choi, Quantum amplitude estimation algorithms on IBM quantum devices, in *Quantum Communications and Quantum Imaging XVIII* (SPIE, Bellingham, WA, USA, 2020), Vol. 11507, pp. 49–60.

[35] K. Plekhanov, M. Rosenkranz, M. Fiorentini, and M. Lubasch, Variational quantum amplitude estimation, Quantum **6**, 670 (2022).

[36] S. Aaronson, Read the fine print, Nat. Phys. **11**, 291 (2015).

[37] N. Linden, A. Montanaro, and C. Shao, Quantum vs. classical algorithms for solving the heat equation, Commun. Math. Phys. **395**, 601 (2022).

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

[39] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[40] Y. Y. Liu, Z. Chen, C. Shu, S.-C. Chew, B. C. Khoo, X. Zhao, and Y. D. Cui, Application of a variational hybrid quantum-classical algorithm to heat conduction equation and analysis of time complexity, Phys. Fluids **34**, 117121 (2022).

[41] R. Demirdjian, D. Gunlycke, C. A. Reynolds, J. D. Doyle, and S. Tafur, Variational quantum solutions to the advection–diffusion equation for applications in fluid dynamics, Quant. Inf. Process. **21**, 322 (2022).

[42] Y. Y. Liu, Z. Chen, C. Shu, P. Rebentrost, Y. G. Liu, S. C. Chew, B. C. Khoo, and Y. D. Cui, A variational quantum algorithm-based numerical method for solving potential and Stokes flows, Ocean Eng. **292**, 116494 (2024).

[43] H.-L. Liu, Y.-S. Wu, L.-C. Wan, S.-J. Pan, S.-J. Qin, F. Gao, and Q.-Y. Wen, Variational quantum algorithm for the Poisson equation, Phys. Rev. A **104**, 022418 (2021).

[44] Y. Sato, R. Kondo, S. Koide, H. Takamatsu, and N. Imoto, Variational quantum algorithm based on the minimum potential energy for solving the Poisson equation, Phys. Rev. A **104**, 052409 (2021).

[45] J. Joo and H. Moon, Quantum variational PDE solver with machine learning, arXiv:2109.09216.

[46] N. M. Guseynov, A. A. Zhukov, W. V. Pogosov, and A. V. Lebedev, Depth analysis of variational quantum algorithms for the heat equation, Phys. Rev. A **107**, 052422 (2023).

[47] H.-M. Li, Z.-X. Wang, and S.-M. Fei, Variational quantum algorithms for Poisson equations based on the decomposition of sparse Hamiltonians, Phys. Rev. A **108**, 032418 (2023).

[48] A. Mezzacapo, M. Sanz, L. Lamata, I. L. Egusquiza, S. Succi, and E. Solano, Quantum simulator for transport phenomena in fluid flows, Sci. Rep. **5**, 13153 (2015).

[49] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, Phys. Rev. A **101**, 010301(R) (2020).

[50] L. Budinski, Quantum algorithm for the advection–diffusion equation simulated with the lattice Boltzmann method, Quant. Inf. Process. **20**, 57 (2021).

[51] O. Kyriienko, A. E. Paine, and V. E. Elfving, Solving nonlinear differential equations with differentiable quantum circuits, Phys. Rev. A **103**, 052416 (2021).

[52] A. E. Paine, V. E. Elfving, and O. Kyriienko, Quantum kernel methods for solving regression problems and differential equations, Phys. Rev. A **107**, 032428 (2023).

[53] Z. Li, Z. Qiao, and T. Tang, *Numerical Solution of Differential Equations: Introduction to Finite Difference and Finite Element Methods* (Cambridge University Press, Cambridge, 2017).

[54] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[55] R. P. Feynman, Quantum mechanical computers, Opt. News **11**, 11 (1985).

[56] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics (American Mathematical Society, Providence, RI, USA, 2002), Vol. 47.

[57] J. McClean, J. A. Parkhill, and A. Aspuru-Guzik, Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics, Proc. Natl. Acad. Sci. USA **110**, 3901 (2013).

[58] D. G. Tempel and A. Aspuru-Guzik, The Kitaev–Feynman clock for open quantum systems, New J. Phys. **16**, 113066 (2014).

[59] S. Barison, F. Vicentini, I. Cirac, and G. Carleo, Variational dynamics as a ground-state problem on a quantum computer, Phys. Rev. Res. **4**, 043161 (2022).

[60] A. J. Pool, A. D. Somoza, M. Lubasch, and B. Horstmann, Solving partial differential equations using a quantum computer, in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Broomfield, CO, USA, 2022), pp. 864–866.

[61] D. Jaksch, P. Givi, A. J. Daley, and T. Rung, Variational quantum algorithms for computational fluid dynamics, AIAA J. **61**, 1885 (2023).

[62] A. Sarma, T. W. Watts, M. Moosa, Y. Liu, and P. L. McMahon, Quantum variational solving of nonlinear and multidimensional partial differential equations, Phys. Rev. A **109**, 062616 (2024).

[63] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).

[64] M. Lubasch, P. Moinier, and D. Jaksch, Multigrid renormalization, J. Comput. Phys. **372**, 587 (2018).

[65] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Quantum fingerprinting, Phys. Rev. Lett. **87**, 167902 (2001).

[66] M. Benedetti, M. Fiorentini, and M. Lubasch, Hardware-efficient variational quantum algorithms for time evolution, Phys. Rev. Res. **3**, 033083 (2021).

[67] F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum computation and quantum-state engineering driven by dissipation, Nat. Phys. **5**, 633 (2009).

[68] B. Horstmann, J. I. Cirac, and G. Giedke, Noise-driven dynamics and phase transitions in fermionic systems, Phys. Rev. A **87**, 012108 (2013).

[69] M. S. J. Tepaske, D. Hahn, and D. J. Luitz, Optimal compression of quantum many-body time evolution operators into brickwall circuits, SciPost Phys. **14**, 073 (2023).

[70] A. V. Uvarov, A. S. Kardashin, and J. D. Biamonte, Machine learning phase transitions with a quantum processor, Phys. Rev. A **102**, 012415 (2020).

[71] C. Mc Keever and M. Lubasch, Classically optimized Hamiltonian simulation, Phys. Rev. Res. **5**, 023146 (2023).

[72] N. Gourianov, M. Lubasch, S. Dolgov, Q. Y. van den Berg, H. Babaee, P. Givi, M. Kiffner, and D. Jaksch, A quantum-inspired approach to exploit turbulence structures, Nat. Comput. Sci. **2**, 30 (2022).

[73] N. Gourianov, Exploiting the structure of turbulence with tensor networks, Ph.D. thesis, University of Oxford, 2022.

[74] R. Haghshenas, J. Gray, A. C. Potter, and G. K.-L. Chan, Variational power of quantum circuit tensor networks, Phys. Rev. X **12**, 011047 (2022).

[75] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[76] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput. **16**, 1190 (1995).

[77] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Günlük, T. Itoko, N. Kanazawa *et al.*, Demonstration of quantum volume 64 on a superconducting quantum computing system, Quantum Sci. Technol. **6**, 025020 (2021).

[78] System Model H1 Product Data Sheet, https://www.quantinuum.com/hardware/h1 (2023).

[79] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer *et al.*, Demonstration of the trapped-ion quantum CCD computer architecture, Nature (London) **592**, 209 (2021).

[80] Qiskit contributors, Qiskit: An open-source framework for quantum computing, Version 0.44.0, Zenodo, https://doi.org/10.5281/zenodo.8190968 (2023).

[81] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, Scalable mitigation of measurement errors on quantum computers, PRX Quantum **2**, 040326 (2021).

[82] A. M. Souza, G. A. Álvarez, and D. Suter, Robust dynamical decoupling, Philos. Trans. R. Soc. A **370**, 4748 (2012).

[83] F. Motzoi, M. P. Kaicher, and F. K. Wilhelm, Linear and logarithmic time compositions of quantum many-body operators, Phys. Rev. Lett. **119**, 160503 (2017).

[84] H. C. Fogedby, Solitons and diffusive modes in the noiseless Burgers equation: Stability analysis, Phys. Rev. E **57**, 2331 (1998).

[85] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).

[86] J. Kim, J. Kim, and D. Rosa, Universal effectiveness of high-depth circuits in variational eigenproblems, Phys. Rev. Res. **3**, 023203 (2021).

[87] M. Larocca, N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo, Theory of overparametrization in quantum neural networks, Nat. Comput. Sci. **3**, 542 (2023).

[88] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandao, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[89] H. Alghassi, A. Deshmukh, N. Ibrahim, N. Robles, S. Woerner, and C. Zoufal, A variational quantum algorithm for the Feynman–Kac formula, Quantum **6**, 730 (2022).

[90] T. Albash and D. A. Lidar, Adiabatic quantum computation, Rev. Mod. Phys. **90**, 015002 (2018).

[91] N. N. Hegade, K. Paul, Y. Ding, M. Sanz, F. Albarrán-Arriagada, E. Solano, and X. Chen, Shortcuts to adiabaticity in digitized adiabatic quantum computing, Phys. Rev. Appl. **15**, 024038 (2021).

[92] C. Mc Keever and M. Lubasch, Towards adiabatic quantum computing using compressed quantum circuits, PRX Quantum **5**, 020362 (2024).

[93] M. Kiffner and D. Jaksch, Tensor network reduced order models for wall-bounded flows, Phys. Rev. Fluids **8**, 124101 (2023).

[94] D. Suess and M. Holzaepfel, mpnum: A matrix product representation library for Python, J. Open Source Softw. **2**, 465 (2017).

[95] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Elementary gates for quantum computation, Phys. Rev. A **52**, 3457 (1995).

[96] V. V. Shende, I. L. Markov, and S. S. Bullock, Minimal universal two-qubit controlled-NOT-based circuits, Phys. Rev. A **69**, 062321 (2004).

[97] T. Jones and J. Gacon, Efficient calculation of gradients in classical simulations of variational quantum algorithms, arXiv:2009.02823.

[98] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[99] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).

[100] C. Zhao and X.-S. Gao, Analyzing the barren plateau phenomenon in training quantum neural networks with the ZX-calculus, Quantum **5**, 466 (2021).

[101] E. Cervero Martín, K. Plekhanov, and M. Lubasch, Barren plateaus in quantum tensor network optimization, Quantum **7**, 974 (2023).