Variational quantum algorithms for nonlinear problems

Michael Lubasch[1], Jaewoo Joo[1], Pierre Moinier[2], Martin Kiffner[3,1], and Dieter Jaksch[1,3]

Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom[1]
BAE Systems, Computational Engineering, Buckingham House,
FPC 267 PO Box 5, Filton, Bristol BS34 7QW, United Kingdom[2] and
Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543[3]

We show that nonlinear problems including nonlinear partial differential equations can be efficiently solved by variational quantum computing. We achieve this by utilizing multiple copies of variational quantum states to treat nonlinearities efficiently and by introducing tensor networks as a programming paradigm. The key concepts of the algorithm are demonstrated for the nonlinear Schrödinger equation as a canonical example. We numerically show that the variational quantum ansatz can be exponentially more efficient than matrix product states and present experimental proof-of-principle results obtained on an IBM Q device.

Nonlinear problems are ubiquitous in all fields of science and engineering and often appear in the form of nonlinear partial differential equations (PDEs). Standard numerical approaches seek solutions to PDEs on discrete grids. However, many problems of interest require extremely large grid sizes for achieving accurate results, in particular in the presence of unstable or chaotic behaviour that is typical for nonlinear problems [1–3]. Examples include large-scale simulations for reliable weather forecasts [4–6] and computational fluid dynamics [7–9].

Quantum computers promise to solve problems that are intractable on conventional, i.e. standard classical, computers through their quantum-enhanced capabilities. In the context of PDEs, it has been realized that quantum computers can solve the Schrödinger equation faster than conventional computers [10–12], and these ideas have been generalized recently to other linear PDEs [13–18]. However, nonlinear problems are intrinsically difficult to solve on a quantum computer due to the linear nature of the underlying framework of quantum mechanics.

Recently, the concept of variational quantum computing (VQC) attracted considerable interest [19–32] for solving optimization problems. VQC is a quantum-classical hybrid approach where the evaluation of the cost function $\mathcal{C}(\boldsymbol{\lambda})$ is delegated to a quantum computer, while the optimization of variational parameters $\boldsymbol{\lambda}$ is performed on a conventional classical computer. The concept of VQC has been applied, e.g., to simulating the dynamics of strongly correlated electrons through non-equilibrium dynamical mean field theory [22, 23, 33, 34], and quantum chemistry calculations were successfully carried out on existing noisy superconducting [21, 25, 26] and ion quantum computers [28].

We extend and adapt the concept of VQC to solving nonlinear problems efficiently on a quantum computer by virtue of two key concepts. First, we introduce a quantum nonlinear processing unit (QNPU) that efficiently calculates nonlinear functions of the form $F = f^{(1)^*} \prod_{j=1}^{r}(O_j f^{(j)})$ for VQC. Measuring the ancilla qubit connected to the QNPU as shown in Fig. 1(a) directly yields the sum of all function values $\sum_k \Re\{F_k\}$,
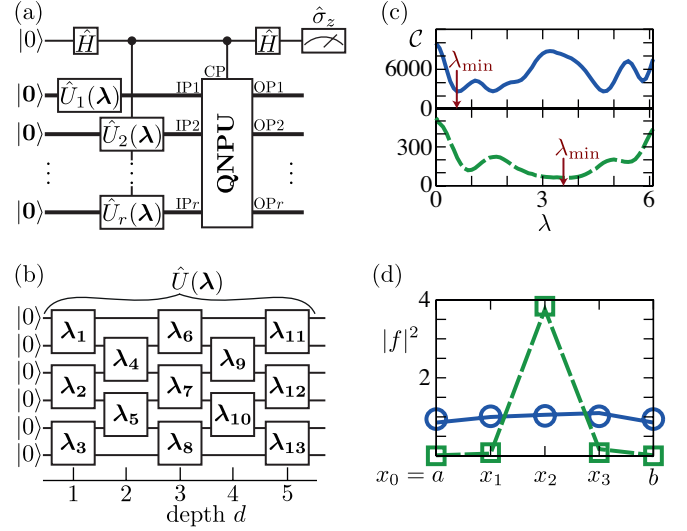

FIG. 1. (a) Quantum network for summing a nonlinear function of the form $F = f^{(1)^*} \prod_{j=1}^{r}(O_j f^{(j)})$. The ancilla qubit on the top line undergoes Hadamard gates $\hat{H}$ and controls operations of the rest of the network via the control port (CP) and is measured in the computational basis. Starting from product states $|\mathbf{0}\rangle$ of n qubits shown as thick lines, variational states $|\psi(\boldsymbol{\lambda})_j\rangle = \hat{U}_j(\boldsymbol{\lambda})|\mathbf{0}\rangle$ representing the functions $f^{(j)}$ are created and fed to the QNPU through input ports IP. The QNPU contains problem specific quantum networks defining the linear operators O_j and has the output ports OP. (b) Network $\hat{U}(\boldsymbol{\lambda})$ of depth $d = 5$ with $n = 6$. The values $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots\}$ determine the form of the two-qubit gates. (c) Cost function $\mathcal{C}(\lambda)$ for the nonlinear Schrödinger equation Eq. (1) for a harmonic potential V, a single variational parameter λ, and $g = 10^4$ (solid line) and $g = 10$ (dashed line). The arrows indicate optimal values $\lambda = \lambda_{\min}$. (d) Solutions $|f(x)|^2$ for $\lambda = \lambda_{\min}$ with $N = 4$ grid points and periodic boundary conditions $f(b) = f(a)$ for $g = 10^4$ (solid line) and $g = 10$ (dashed line). The circles and squares are numerically exact results. Experimental data in (c) and (d) was obtained – from a modified circuit – on IBM Q devices (see [35] for details).

where $\Re\{\cdot\}$ denotes the real part. The functions $f^{(j)}$ are encoded in variational n-qubit states $|\psi(\boldsymbol{\lambda})_j\rangle = \hat{U}_j(\boldsymbol{\lambda})|\mathbf{0}\rangle$

created by networks of the form shown in Fig. 1(b). The same function $f^{(i)} = f^{(j)}$ may appear multiple times by choosing $\hat{U}_i(\boldsymbol{\lambda}) = \hat{U}_j(\boldsymbol{\lambda})$. Second, we use tensor networks as a programming paradigm for QNPUs to create optimized circuits that efficiently calculate linear operators O_j acting on functions $f^{(j)}$. In this way all quantum resources of nonlinear VQC scale polynomially with the number of qubits, which represents an exponential reduction compared with some conventional algorithms.

The variational states $|\psi(\boldsymbol{\lambda})_j\rangle$ represent $N = 2^n$ values of the functions $f^{(j)}$ which form a trial solution to the problem of interest. The cost function $\mathcal{C}(\boldsymbol{\lambda})$ for nonlinear VQC is built up from outputs of different QNPUs that are then processed classically to iteratively determine the optimal set $\boldsymbol{\lambda}$. Large grid sizes that are intractable on a conventional computer require only $n \gtrsim 20$ qubits which is within reach of noisy intermediate-scale quantum (NISQ) devices. In addition, the scheme is applicable to other types of nonlinear problems that can be solved via the minimization of a cost function $\mathcal{C}(\boldsymbol{\lambda})$ [36].

We demonstrate the concept and performance of nonlinear VQC by emulating it classically for the canonical example of the time-independent one-dimensional nonlinear Schrödinger equation

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) + g|f(x)|^2 \right] f(x) = E f(x), \quad (1)$$

where V is an external potential and g denotes the strength of the nonlinearity. We also implement nonlinear VQC for Eq. (1) on IBM quantum computers to establish its feasibility on current NISQ devices. Proof-of-principle results are shown in Figs. 1(c) and 1(d) demonstrating excellent agreement with numerically exact solutions. The nonlinear Schrödinger equation and its generalizations to higher dimensions describe various physical phenomena ranging from Bose-Einstein condensation to light propagation in nonlinear media [37–42]. In particular, below we consider Eq. (1) with quasi-periodic potentials V that are in the focus of current cold atom experiments [43] and that make Eq. (1) challenging to solve numerically. The methods used for this equation here are straightforwardly modified to handle other nonlinear terms and time-dependent problems as illustrated in [35] for the Burgers equation appearing in fluid dynamics.

The ground state of Eq. (1) can be found by minimizing the cost function

$$\mathcal{C} = \langle\!\langle K \rangle\!\rangle_c + \langle\!\langle P \rangle\!\rangle_c + \langle\!\langle I \rangle\!\rangle_c, \quad (2)$$

where $\langle\!\langle K \rangle\!\rangle_c$, $\langle\!\langle P \rangle\!\rangle_c$ and $\langle\!\langle I \rangle\!\rangle_c$ are the mean kinetic, potential and interaction energies, respectively. In Eq. (2) $\langle\!\langle \cdot \rangle\!\rangle_c$ denotes averages with respect to a single real-valued function $f^{(1)} \equiv f$ on the interval $[a, b]$ satisfying the normalization condition $\int_a^b |f(x)|^2 \mathrm{d}x = 1$.

In line with standard numerical approaches [44–46] we apply the finite difference method (FDM) to Eq. (1) and discretize the interval $[a, b]$ into N equidistant grid points

$x_k = a + h_N k$, where $h_N = \ell/N$ is the grid spacing, $\ell = b - a$ is the length of the interval and $k \in \{0, \dots, N-1\}$. Each grid point is associated with a variational parameter f_k that approximates the continuous solution $f(x_k)$ at x_k. Furthermore, we impose periodic boundary conditions (i.e., $f_N = f_0$) and the normalization condition imposed on the continuous functions f translates to

$$1 = h_N \sum_{k=0}^{N-1} |f_k|^2 = \sum_{k=0}^{N-1} |\psi_k|^2, \quad (3)$$

where $\psi_k = \sqrt{h_N} f_k$. Note that the condition on the set of parameters $\{\psi_k\}$ is independent of the grid spacing, and in the following we consider optimizing the cost function with respect to them.

All averages $\langle\!\langle \cdot \rangle\!\rangle_c$ in Eq. (2) can be approximated by corresponding expressions of the discrete problem $\langle\!\langle \cdot \rangle\!\rangle$. We find [44–46] $\langle\!\langle \cdot \rangle\!\rangle_c = \langle\!\langle \cdot \rangle\!\rangle + \mathcal{E}_{\mathrm{grid}}$, where $\mathcal{E}_{\mathrm{grid}} \propto 1/N^2$ is the error associated with the trapezoidal rule when transforming integrals into sums, and

$$\langle\!\langle K \rangle\!\rangle = -\frac{1}{2}\frac{1}{h_N^2} \sum_{k=0}^{N-1} \psi_k^* \left(\psi_{k+1} - 2\psi_k + \psi_{k-1} \right), \quad (4\text{a})$$

$$\langle\!\langle P \rangle\!\rangle = \sum_{k=0}^{N-1} \left[\psi_k^* V(x_k) \psi_k \right], \quad (4\text{b})$$

$$\langle\!\langle I \rangle\!\rangle = \frac{1}{2}\frac{g}{h_N} \sum_{k=0}^{N-1} |\psi_k|^4. \quad (4\text{c})$$

Note that $\langle\!\langle K \rangle\!\rangle$ in Eq. (4a) uses a FDM representation of the second-order derivative in Eq. (1).

For evaluating the terms in Eq. (4) on a quantum computer we consider quantum registers with n qubits and basis states $|\boldsymbol{q}\rangle = |q_1 \dots q_n\rangle = |q_1\rangle \otimes \dots \otimes |q_n\rangle$, where $q_j \in \{0, 1\}$ denotes the computational states of qubit j. Regarding the sequence $q_1 \dots q_n = \mathrm{binary}(k)$ as the binary representation of the integer $k = \sum_{j=1}^n q_j 2^{n-j}$, we encode all $N = 2^n$ amplitudes ψ_k in the normalized state

$$|\psi\rangle = \sum_{k=0}^{N-1} \psi_k |\mathrm{binary}(k)\rangle. \quad (5)$$

We prepare the quantum register in a variational state $|\psi(\boldsymbol{\lambda})\rangle$ via the quantum circuit $\hat{U}(\boldsymbol{\lambda})$ of depth d shown in Fig. 1(b). We consider depths $d \propto \mathrm{poly}(n)$ such that the quantum ansatz requires exponentially fewer parameters $\mathcal{N}$ than standard classical schemes with N parameters. Note that the number of variational parameters scales like $\mathcal{N} \propto nd$ for our quantum ansatz [35]. The power of this ansatz is rooted in the fact that it encompasses all matrix product states (MPS) [47–50] with bond dimension $\chi \sim \mathrm{poly}(n)$ [35]. Since polynomials and Fourier series [51, 52] can be efficiently represented by MPS, the quantum ansatz simultaneously contains universal basis functions that are capable of approximating a large class of solutions to nonlinear problems efficiently. Furthermore, we show below that the quantum ansatz is capable of storing solutions with exponentially fewer resources
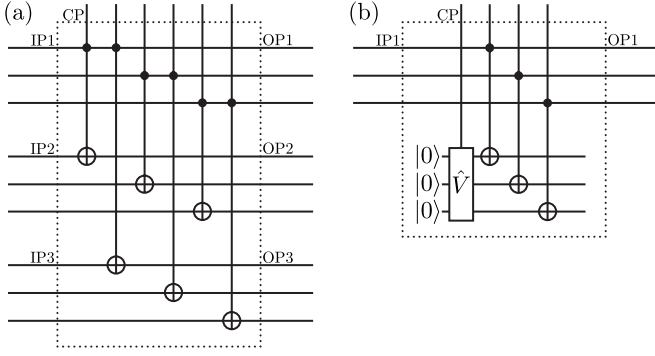
FIG. 2. (a) QNPU circuit calculating the nonlinear term $|\psi|^4$. The networks are shown for $n = 3$ and all input ports IP are fed the same variational quantum states created by $\hat{U}(\boldsymbol{\lambda})$. IP2 is fed $\hat{U}(\boldsymbol{\lambda})$ and IP3 is fed $\hat{U}^*(\boldsymbol{\lambda})$. (b) QNPU circuit for working out the potential energy term $\tilde{V}|\psi|^2$.

than the classically optimized MPS ansatz. Note that the number of variational parameters scales like $\mathcal{N} \propto n\chi^2$ for MPS of bond dimension χ [35].

Figure 2(a) demonstrates the basic working principle of the QNPU for the nonlinear term $\langle\!\langle I \rangle\!\rangle$. The effect of the controlled NOT operations between pairs of qubits is to provide a point-wise multiplication with the ancilla thus measuring $\sum_k |\psi_k|^4$. In Fig. 2(b) we show the circuit for measuring $\langle\!\langle P \rangle\!\rangle$. The unitary $\hat{V} \hat{=} O_1$ encodes function values of the external potential V. A copy of ψ is effectively multiplied point-wise with the external potential by controlled NOT gates to give $\sum_k \tilde{V}_k|\psi_k|^2$. Similarly, multiplying ψ with their shifted versions using adder circuits (see [35] for details) allows evaluating the kinetic energy term.

The measured expectation value of the ancilla qubit is directly related to the desired quantities as $\langle\!\langle I \rangle\!\rangle = g\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^I/2h_N$ for the nonlinear term, $\langle\!\langle P \rangle\!\rangle = \alpha\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^P$ for the potential energy and $\langle\!\langle K \rangle\!\rangle = \left(1 - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K\right)/h_N^2$ for the kinetic energy [53, 54]. Furthermore, derivatives of the cost function, as required by some minimization algorithms [36, 55, 56], can be evaluated by combining the ideas presented here with the quantum circuits discussed in [24, 30, 57, 58].

The unitary network $\hat{V}$ represents scaled function values $\tilde{V}_k$ of the external potential where $\sum_{k=0}^{N-1} |\tilde{V}_k|^2 = 1$, and $\alpha > 0$ a scaling parameter such that $V_k = \alpha\tilde{V}_k$. Efficient quantum circuits $\hat{V}$ for measuring $\langle\!\langle P \rangle\!\rangle$ can be systematically obtained by establishing tensor networks as a programming paradigm. To this end we expand the external potential in polynomials or Fourier series, $\tilde{V}(x) \approx \sum_j^J c_j b_j(x)$, where $b_j(x)$ are basis functions and c_j are expansion coefficients [35]. In the case of Fourier series of order J, the approximate potential is represented by an MPS of bond dimension $\chi = J$ [51, 52]. Next we write the MPS in terms of $n - \lceil \log \chi \rceil$ unitaries [59–61], where $\lceil \cdot \rceil$ is the ceiling function. Each of these unitaries acts on 2χ qubits and can be decomposed in terms of elementary two-qubit gates [35, 62–65]. An upper

bound for the depth of the resulting quantum circuit is $d_> \leq 9n[(23/48)(2\chi)^2 + 4/3]$ [64]. The depth thus scales polynomially with the number of qubits n and with χ, and many problems of interest show an even more advantageous scaling. For example, in the following we consider the potential

$$V(x) = s_1 \sin(\kappa_1 x) + s_2 \sin(\kappa_2 x) \tag{6}$$

and set $\kappa_2 = 2\kappa_1/(1 + \sqrt{5})$. This potential realizes an incommensurate bichromatic lattice where the ratio s_1/s_2 determines the amount of disorder in the lattice [66]. The trap potential $V(x)$ in Eq. (6) is exactly represented by an MPS of bond dimension $\chi = 4$. The depth of the corresponding quantum circuit $d = 5(n - 2) + 1 \ll d_>$ is much smaller than the upper bound [35].

Next we analyze the Monte Carlo sampling error [44] associated with the measurement of the ancilla qubit. We denote the absolute sampling error associated with quantity X by $\mathcal{E}_{\mathrm{MC}}^X$, and the corresponding relative error is [35]

$$\epsilon_{\mathrm{MC}}^P = \frac{\mathcal{E}_{\mathrm{MC}}^P}{\langle\!\langle P \rangle\!\rangle} = C_P \frac{1}{\sqrt{M}}, \tag{7a}$$

$$\epsilon_{\mathrm{MC}}^K = \frac{\mathcal{E}_{\mathrm{MC}}^K}{\langle\!\langle K \rangle\!\rangle} \approx C_K \frac{N}{N_{\min}} \frac{1}{\sqrt{M}}, \tag{7b}$$

$$\epsilon_{\mathrm{MC}}^I = \frac{\mathcal{E}_{\mathrm{MC}}^I}{\langle\!\langle I \rangle\!\rangle} \approx C_I \frac{N}{N_{\min}} \frac{1}{\sqrt{M}}. \tag{7c}$$

In this equation, we assume $N \geq N_{\min}$ and $N_{\min} = \ell/\ell_{\min}$ is the minimal number of grid points for resolving the smallest length scale $\ell_{\min}$ of the problem. The parameters C_X in Eq. (7) are of the order of unity [35], and all sampling errors decrease with the number of samples M as $1/\sqrt{M}$. While the relative error associated with the potential term in Eq. (7a) is independent of the number of grid points, ϵ_{MC}^K and ϵ_{MC}^I increase linearly with $N/N_{\min}$. It follows that increasing the grid size requires larger values of M in order to keep the sampling error small. However, the grid error scales like $\mathcal{E}_{\mathrm{grid}} \propto 1/N^2 = 2^{-2n}$ for $N \geq N_{\min}$ [35]. We thus conclude that only moderate ratios $N/N_{\min} > 1$ and therefore relatively small values of M are needed in order to achieve accurate solutions with small grid errors.

The quantum ansatz in Fig. 1(b) is inspired from tensor network theory and can be regarded, for example, as the Trotter decomposition of the time t evolution operator $\exp(-iH(t)t/\hbar)$ of a time-dependent spin Hamiltonian $H(t)$ with arbitrary short-range interactions acting on the initial state $|\mathbf{0}\rangle$ [67]. Similarly to the coupled cluster ansatz in quantum chemistry VQC calculations [19], there is currently no known efficient classical ansatz for this state [68, 69]. From the VQC perspective, our quantum ansatz in Fig. 1(b) is composed of generic two-qubit gates that in an experiment are accurately approximated by short sequences of gates if a sufficiently tunable or universal gate set is experimentally available [70]. We envisage that this quantum ansatz is more efficient than
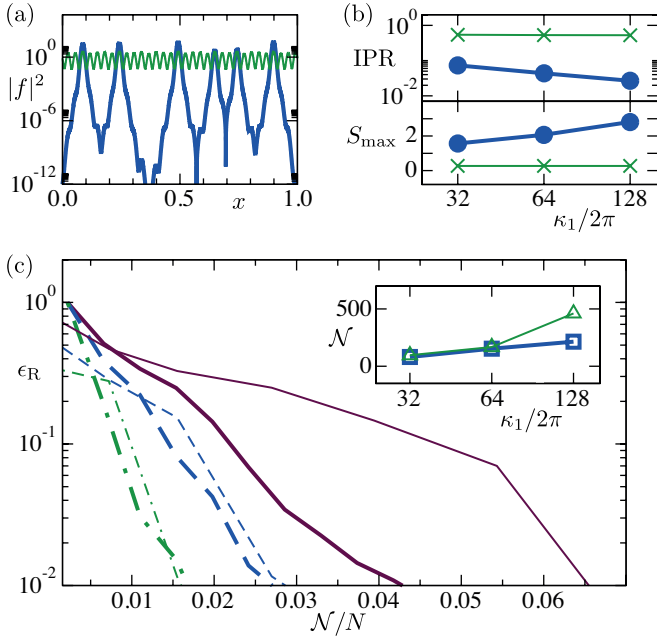
FIG. 3. (a) Numerically exact solution $|f(x)|^2$ of Eq. (1) on a logarithmic scale, $V(x)$ in Eq. (6) with $s_1 = 2 \times 10^4$ and $\kappa_1 = 2\pi \times 32$. The green thin line (blue normal line) is for $s_1/s_2 = 200$ ($s_1/s_2 = 2$). (b) Log-log plot of the IPR (top panel) and lin-log plot of $S_{\max}$ of the exact solution $|\psi^{\text{exact}}\rangle$ (bottom panel) as a function of κ_1. Green crosses (blue dots) correspond to $s_1/s_2 = 200$ ($s_1/s_2 = 2$). (c) Representation error ϵ_R of the exact solution for the quantum ansatz (thick lines) and the MPS ansatz (thin lines) as a function of $\mathcal{N}/N$. All curves are for $s_1/s_2 = 2$ and correspond to $s_1 = 2 \times 10^4$ and $\kappa_1 = 2\pi \times 32$ (green dash-dotted), $s_1 = 8 \times 10^4$ and $\kappa_1 = 2\pi \times 64$ (blue dashed), and $s_1 = 3.2 \times 10^5$ and $\kappa_1 = 2\pi \times 128$ (purple solid), respectively. The inset shows $\mathcal{N}$ as a function of κ_1 for $\epsilon_R = 0.05$ and $s_1/s_2 = 2$. Blue squares (green triangles) correspond to the quantum ansatz (MPS ansatz). All curves in (a)-(c) are for $N = 2^{13} = 8192$ grid points and $g = 50$.

methods based on an MPS ansatz on a classical computer like the multigrid renormalization (MGR) method in [56]. This MGR method is the most efficient and accurate classical algorithm known to us for the problem considered here, for which it can already be exponentially faster than standard classical algorithms. A comparison with this powerful classical method, which is based on variational classical MPS, will allow us to validate the superior variational power of the quantum ansatz in Fig. 1(b) on a quantum computer. Note that the difficulty of solving a classical optimization problem in many variables does not go away by using the quantum ansatz, as the actual optimization is classical and there is no quantum advantage there. The quantum advantage stems solely from the faster evaluation of the cost function for our quantum ansatz in Fig. 1(b) on a quantum computer.

To provide numerical evidence for this we first obtain the numerically exact solution of Eq. (1) on the interval $[0, 1]$ via the MGR algorithm [56] and by allowing for the maximal bond dimension χ of the MPS

ansatz [35]. In this case the numerically exact solution is described by $N = 2^n$ parameters like in other conventional algorithms. The results are shown in Fig. 3 (a) for two different values of s_1/s_2. In the weakly disordered regime $s_1/s_2 \gg 1$, $|f(x)|^2$ varies on the length scale set by $1/\kappa_1$. On the contrary, the strongly disordered regime $s_1/s_2 \approx 1$ is characterized by strongly localized solutions in space. The localization of the wavefunction can be quantified using the inverse participation ratio (IPR) [71], IPR $= \left(N \sum_{k=0}^{N-1} |\psi_k|^4\right)^{-1}$. We show the IPR in Fig. 3(b) as a function of κ_1 (top panel) and find that it stays constant for $s_1/s_2 \gg 1$. On the other hand, the IPR decreases according to a power law with κ_1 for $s_1/s_2 \approx 1$, showing that the localized character of the wavefunction increases dramatically with κ_1.

Next we encode the function values of the numerically exact solution in the state $|\psi^{\text{exact}}\rangle$ via Eq. (5) and calculate the maximum bipartite entanglement entropy $S_{\max}$ of all possible bipartitions of the n-qubit wave function $|\psi^{\text{exact}}\rangle$. The quantity $S_{\max}$ is a measure of the entanglement of $|\psi^{\text{exact}}\rangle$ and is shown in the bottom panel of Fig. 3(b). The value of $S_{\max}$ is small and stays constant with κ_1 for $s_1/s_2 \gg 1$ in the weakly disordered regime. Contrary to this, for $s_1/s_2 \approx 1$ in the strongly disordered regime we observe that $S_{\max} \propto \log(\kappa_1)$.

The entanglement measure $S_{\max}$ provides a useful necessary criterion for efficient MPS approximations [72]. A MPS of bond dimension χ can at most contain an amount of entanglement $S_{\max}[\text{MPS}] \leq \log(\chi)$ [47, 50]. For MPS to be efficient, we require that χ scales at most polynomially with n, i.e. $\chi = \text{poly}(n)$, so that $S_{\max}[\text{MPS}] \leq \log(\text{poly}(n))$ and therefore MPS can only capture small amounts of entanglement efficiently. The small values of $S_{\max}$ for $s_1/s_2 \gg 1$ suggest that MPS work well in the weakly disordered regime and indeed we have confirmed numerically that this is true. Therefore, in the following we focus on the strongly disordered regime $s_1/s_2 \approx 1$. In this regime MPS cannot be an efficient approximation for large values of κ_1: The total number of variational parameters of MPS $\mathcal{N}[\text{MPS}]$ depends quadratically on χ [35], i.e. $\mathcal{N}[\text{MPS}] \propto \chi^2$, and χ needs to grow polynomially with κ_1, i.e. $\chi \propto \text{poly}(\kappa_1)$, to satisfy the observed entanglement scaling, such that $\mathcal{N}[\text{MPS}] \propto \text{poly}(\kappa_1)$. Our quantum ansatz (QA) of Fig. 1(b) can capture much larger amounts of entanglement $S_{\max}[\text{QA}] \lesssim d$ efficiently [69] and therefore this ansatz can be an efficient approximation for large values of κ_1: Because the total number of variational parameters $\mathcal{N}$ depends linearly on d for the quantum ansatz [35], i.e. $\mathcal{N}[\text{QA}] \propto d$, and d just needs to grow logarithmically, i.e. $d \propto \log(\kappa_1)$, for the observed entanglement requirements, we conclude that $\mathcal{N}[\text{QA}] \propto \log(\kappa_1)$. These entanglement considerations show that the quantum ansatz has the potential to be exponentially more efficient than MPS in the strongly disordered regime for increasing values of κ_1.

To quantitatively analyze and demonstrate the efficiency of the quantum ansatz in this regime, we ob-

tain the set of parameters $\boldsymbol{\lambda}$ that maximize the fidelity $\mathcal{F} = |\langle\psi^{\text{exact}}|\psi(\boldsymbol{\lambda})\rangle|$ for different depths d [35]. The infidelity $\epsilon_{\text{R}} = 1 - \mathcal{F}$ is thus a measure of the error when approximating the exact solution by this ansatz, and in the following we refer to ϵ_{R} as the representation error. As shown in Fig. 3(c), the representation error decreases exponentially as a function of $\mathcal{N}$ for all values of κ_1 and therefore we obtain accurate solutions for $\mathcal{N}/N \ll 1$. Even for the largest value of $\kappa_1 = 2\pi \times 128$ and $\epsilon_{\text{R}} \approx 10^{-2}$, we find $\mathcal{N}/N \approx 0.04$ so that we only require 4% of the full number of parameters needed in conventional algorithms. Most importantly, the inset of Fig. 3(c) shows that, to obtain a fixed representation error of $\epsilon_{\text{R}} = 0.05$, the number of parameters of our quantum ansatz needs to grow as $\mathcal{N}[\text{QA}] \propto \log(\kappa_1)$. This numerical analysis therefore confirms our expectation, from the entanglement arguments in the previous paragraph, that the quantum ansatz efficiently approximates solutions in the strongly disordered regime even for large values of κ_1. This is possible because the quantum ansatz captures the required entanglement $S_{\max} \propto \log(\kappa_1)$ by means of just the small number of parameters $\mathcal{N}[\text{QA}] \propto \log(\kappa_1)$. The entanglement capabilities of MPS imply that $\mathcal{N}[\text{MPS}] \propto \text{poly}(\kappa_1)$ has to be fulfilled. Therefore we conclude that our quantum ansatz is exponentially more efficient than MPS in the strongly disordered regime for growing values of κ_1. This key finding shows that nonlinear VQC can be exponentially more efficient than optimized, classical variational schemes that are based on the MPS ansatz.

We now propose a calculation that illustrates the exponential advantage of our quantum ansatz on a quantum computer particularly clearly. We consider the ground state problem of Eq. (1) with the quasi-random external potential of Eq. (6) on the interval $[0, 1]$ discretized by $N = 2^n$ equidistant grid points and use the FDM from above that leads to Eqs. (4a), (4b), and (4c). We choose $\kappa_1 = 2\pi \times 2^{n/2}$ and assume that $n = 4, 6, 8, \ldots$ qubits store the variational quantum ansatz. The smallest wave length present in our external potential of Eq. (6) is determined by κ_1 and reads $2\pi/\kappa_1 = 2^{-n/2}$. The interval $[0, 1]$ accomodates $1/2^{-n/2} = 2^{n/2}$ of such wave lengths and our grid of $N = 2^n$ equidistant points resolves each such wave length using $N/2^{n/2} = 2^{n/2}$ grid points. Therefore the randomness of our quasi-random external potential of Eq. (6) as well as its FDM resolution grow exponentially with the number of qubits n. At the same time, all FDM errors (using a FDM representation of the Laplace operator, the trapezoidal rule for integration, and so on) decrease polynomially with the grid spacing $1/N = 2^{-n}$ and thus exponentially with n. Therefore, with growing values of n, our FDM representation of Eq. (1) converges exponentially fast to the continuous problem and

the accuracy of randomness in Eq. (6) grows exponentially. Our proposal is to solve this problem for the strongly disordered regime $s_1/s_2 \approx 1$ in Eq. (6), i.e. compute the corresponding ground states, using an increasing number $n = 4, 6, 8, \ldots$. Based on the numerical analysis of Fig. 3, we conjecture that MPS require resources $\mathcal{N}[\text{MPS}] \propto \text{poly}(\kappa_1) = \text{poly}(2\pi \times 2^{n/2}) = \exp(n)$ growing exponentially with n, whereas the quantum ansatz of Fig. 1(b) just needs resources $\mathcal{N}[\text{QA}] \propto \log(\kappa_1) = \log(2\pi \times 2^{n/2}) = \text{poly}(n)$ increasing polynomially with n. This calculation can be used to demonstrate quantum supremacy as, by successively increasing n, the limit of what is computationally possible on a classical computer (using MPS or exact classical methods) is reached quickly and our quantum ansatz on a quantum computer remains efficient far beyond this classical limit.

The quantum hardware requirements for this quantum supremacy calculation go beyond the current capabilities of available NISQ devices [73]. Nevertheless the superior performance of our quantum ansatz is relevant for current NISQ devices as only the most efficient variational states can succeed in the presence of the current experimental errors. We test the feasibility of nonlinear VQC on NISQ devices by calculating the ground state of Eq. (1) for a simple harmonic potential and a single variational parameter on an IBM Q device [74] utilizing further network optimizations (see [35] for details). The experimental implementation of the nonlinear VQC algorithm was able to identify the optimal variational parameter with an error of less than 10% leading to excellent agreement of the ground state solutions with exact numerical solutions [c.f. Figs. 1(c)(d)].

The methods presented here are readily modified to two- and three-dimensional problems with an overhead scaling linearly in the number of dimensions, and can be applied to a broad range of nonlinear terms and differential operators. An exciting prospect for future work will be to utilize intermediate-scale quantum computers for solving non-linear problems on grid sizes beyond the scope of conventional computers.

[1] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer, New York, USA, 2003).

[2] L. A. Aguirre and C. Letellier, "Modeling Nonlinear Dynamics and Chaos: A Review," Mathematical Problems in Engineering **2009**, 238960 (2009).

[3] S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Westview Press, Boulder, USA, 2015).

[4] P. Lynch, *The Emergence of Numerical Weather Prediction* (Cambridge University Press, Cambridge, UK, 2006).

[5] T. T. Warner, *Numerical Weather and Climate Prediction* (Cambridge University Press, Cambridge, UK, 2011).

[6] R. A. Pielke Sr., *Mesoscale Meteorological Modeling* (Academic Press, Waltham, USA, 2013).

[7] L. M. Milne-Thomson, *Theoretical Aerodynamics* (Dover Publications, New York, USA, 1973).

[8] J. H. Ferziger and M. Perić, *Computational Methods for Fluid Dynamics* (Springer, Berlin, Germany, 2002).

[9] H. K. Versteeg and W. Malalasekera, *An Introduction to Computational Fluid Dynamics* (Pearson Education Limited, Harlow, UK, 2007).

[10] S. Lloyd, "Universal Quantum Simulators," Science **273**, 1073 (1996).

[11] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, "Simulated Quantum Computation of Molecular Energies," Science **309**, 1704 (2005).

[12] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, "Polynomial-time quantum algorithm for the simulation of chemical dynamics," Proc. Natl. Acad. Sci. **105**, 18681 (2008).

[13] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum Algorithm for Linear Systems of Equations," Phys. Rev. Lett. **103**, 150502 (2009).

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

[15] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, "Simulating Hamiltonian Dynamics with a Truncated Taylor Series," Phys. Rev. Lett. **114**, 090502 (2015).

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

[17] I. D. Kivlichan, N. Wiebe, R. Babbush, and A. Aspuru-Guzik, "Bounding the costs of quantum simulation of many-body physics in real space," J. Phys. A: Math. Theor. **50**, 305301 (2017).

[18] A. M. Childs, R. Kothari, and R. D. Somma, "Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision," SIAM J. Comput. **46**, 1920 (2017).

[19] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nat. Commun. **5**, 4213 (2014).

[20] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms," New Journal of Physics **18**, 023023 (2016).

[21] P. J. J. OMalley, R. Babbush, I. D. Kivlichan, J. Romero, R. Barends J. R. McClean, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, "Scalable Quantum Simulation of Molecular Energies," Phys. Rev. X **6**, 031007 (2016).

[22] J. M. Kreula, S. R. Clark, and D. Jaksch, "Non-linear quantum-classical scheme to simulate non-equilibrium strongly correlated fermionic many-body dynamics," Scientific Reports **6**, 32940 (2016).

[23] J. M. Kreula, L. García-Álvarez, L. Lamata, S. R. Clark, E. Solano, and D. Jaksch, "Few-qubit quantum-classical simulation of strongly correlated lattice fermions," EPJ Quantum Technology **3**, 11 (2016).

[24] Y. Li and S. C. Benjamin, "Efficient Variational Quantum Simulator Incorporating Active Error Minimization," Phys. Rev. X **7**, 021050 (2017).

[25] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," Nature **549**, 242 (2017).

[26] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, "Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm," Phys. Rev. X **8**, 011021 (2018).

[27] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, "Cloud Quantum Computing of an Atomic Nucleus," Phys. Rev. Lett. **120**, 210501 (2018).

[28] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, "Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator," Phys. Rev. X **8**, 031022 (2018).

[29] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz," Quantum Sci. Technol. **4**, 014008 (2018).

[30] S. McArdle, T. Jones, S. Endo, Y. Li, S. Benjamin, and X. Yuan, "Variational quantum simulation of imaginary time evolution," arXiv preprint, arXiv:1804.03023v3 (2018).

[31] S. Endo, Y. Li, S. Benjamin, and X. Yuan, "Variational quantum simulation of general processes," arXiv preprint, arXiv:1812.08778v1 (2018).

[32] M.-C. Chen, M. Gong, X.-S. Xu, X. Yuan, J.-W. Wang, C. Wang, C. Ying, J. Lin, Y. Xu, Y. Wu, S. Wang, H. Deng, F. Liang, C.-Z. Peng, S. C. Benjamin, X. Zhu, C.-Y. Lu, and J.-W. Pan, "Demonstration of Adiabatic Variational Quantum Computing with a Superconducting Quantum Coprocessor," arXiv preprint, arXiv:1905.03150v1 (2019).

[33] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, "Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions," Rev. Mod. Phys. **68**, 13 (1996).

[34] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, "Electronic structure calculations with dynamical mean-field theory," Rev. Mod. Phys. **78**, 865 (2006).

[35] See Supplemental Material at http://link.aps.org/supplemental/xxx for details about the IBM quantum computations, adder circuit for the kinetic energy, nonlinear VQC for the Burgers equation, MPS ansatz, sampling error, fidelity $\mathcal{F}$ optimization and number of variational parameters.

[36] I. Griva, S. G. Nash, and A. Sofer, *Linear and Nonlinear Optimization* (SIAM, Philadelphia, USA, 2009).

[37] E. P. Gross, "Structure of a quantized vortex in boson systems," Il Nuovo Cimento **20**, 454 (1961).

[38] L. P. Pitaevskii, "Vortex lines in an imperfect Bose gas," JETP **13**, 451 (1961).

[39] A. J. Leggett, "Bose-Einstein condensation in the alkali gases: Some fundamental concepts," Rev. Mod. Phys. **73**, 307 (2001).

[40] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford, UK, 2003).

[41] A. Scott, *Encyclopedia of Nonlinear Science* (Routledge, New York, USA, 2005).

[42] G. Agrawal, *Nonlinear Fiber Optics* (Academic Press, Waltham, USA, 2013).

[43] K. Viebahn, M. Sbroscia, E. Carter, J.-C. Yu, and U. Schneider, "Matter-Wave Diffraction from a Quasicrystalline Optical Lattice," Phys. Rev. Lett. **122**, 110404 (2019).

[44] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, USA, 1992).

[45] C. Grossmann, H.-G. Roos, and M. Stynes, *Numerical Treatment of Partial Differential Equations* (Springer, Berlin, Germany, 2007).

[46] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations* (Cambridge University Press, Cambridge, UK, 2009).

[47] F. Verstraete, V. Murg, and J. I. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," Adv. Phys. **57**, 143 (2008).

[48] I. V. Oseledets, "Approximation of $2^d \times 2^d$ Matrices Using Tensor Decomposition," SIAM J. Matrix Anal. Appl. **31**, 2130 (2010).

[49] I. V. Oseledets, "Tensor-Train Decomposition," SIAM J. Sci. Comput. **33**, 2295 (2011).

[50] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," Ann. Phys. **349**, 117 (2014).

[51] B. N. Khoromskij, "$O(d \log N)$-Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling," Constr. Approx. **34**, 257 (2011).

[52] I. V. Oseledets, "Constructive Representation of Functions in Low-Rank Tensor Formats," Constr. Approx. **37**, 1 (2013).

[53] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, "Direct Estimations of Linear and Nonlinear Functionals of a Quantum State," Phys. Rev. Lett. **88**, 217901 (2002).

[54] C. M. Alves, P. Horodecki, D. K. L. Oi, L. C. Kwek, and A. K. Ekert, "Direct estimation of functionals of density operators by local operations and classical communication," Phys. Rev. A **68**, 032306 (2003).

[55] J. M. Spall, "Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation," IEE Trans. Automat. Contr. **37**, 332 (1992).

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

[57] T. E. O'Brien, B. Senjean, R. Sagastizabal, X. Bonet-Monroig, A. Dutkiewicz, F. Buda, L. DiCarlo, and L. Visscher, "Calculating energy derivatives for quantum chemistry on a quantum computer," arXiv preprint, arXiv:1905.03742v3 (2019).

[58] K. Mitarai, Y. O. Nakagawa, and W. Mizukami, "Theory of analytical energy derivatives for the variational quantum eigensolver," arXiv preprint, arXiv:1905.04054v2 (2019).

[59] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, "Sequential Generation of Entangled Multiqubit States," Phys. Rev. Lett. **95**, 110503 (2005).

[60] C. Schön, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano, "Sequential generation of matrix-product states in cavity QED," Phys. Rev. A **75**, 032311 (2007).

[61] M. C. Bañuls, D. Pérez-García, M. M. Wolf, F. Verstraete, and J. I. Cirac, "Sequentially generated states for the study of two-dimensional systems," Phys. Rev. A **77**, 052306 (2008).

[62] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental realization of any discrete unitary operator," Phys. Rev. Lett. **73**, 58 (1994).

[63] 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).

[64] V. V. Shende, S. S. Bullock, and I. L. Markov, "Synthesis of quantum-logic circuits," IEEE Trans. on Computer-Aided Design **25**, 1000 (2006).

[65] R. Iten, R. Colbeck, I. Kukuljan, J. Home, and M. Christandl, "Quantum circuits for isometries," Phys. Rev. A **93**, 032318 (2016).

[66] B. Deissler, M. Zaccanti, G. Roati, C. DErrico, M. Fattori, M. Modugno, G. Modugno, and M. Inguscio, "Delocalization of a disordered bosonic system by repulsive interactions," Nature Phys. **6**, 354 (2010).

[67] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, "Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces," J. Stat. Mech. **4**, P04005 (2004).

[68] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, "Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas," Nature Phys. **8**, 325 (2012).

[69] J. Schachenmayer, B. P. Lanyon, C. F. Roos, and A. J. Daley, "Entanglement Growth in Quench Dynamics with Variable Range Interactions," Phys. Rev. X **3**, 031015 (2013).

[70] M. A Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[71] B. Kramer and A. MacKinnon, "Localization: theory and experiment," Rep. Prog. Phys. **56**, 1469 (1993).

[72] G. Vidal, "Efficient Classical Simulation of Slightly Entangled Quantum Computations," Phys. Rev. Lett. **91**, 147902 (2003).

[73] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, "Validating quantum computers using

randomized model circuits," Phys. Rev. A **100**, 032328 (2019).

[74] We acknowledge use of the IBM Q for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Q team.

Supplemental Material:
Variational quantum algorithms for nonlinear problems

Michael Lubasch[1], Jaewoo Joo[1], Pierre Moinier[2], Martin Kiffner[3,1], and Dieter Jaksch[1,3]

Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom[1]
BAE Systems, Computational Engineering, Buckingham House,
FPC 267 PO Box 5, Filton, Bristol BS34 7QW, United Kingdom[2] and
Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543[3]

First we provide the details for the IBM quantum computer experiments in Sec. I. Then Sec. II contains a description of the adder circuit used in the main text for computing the kinetic energy. Section III presents how to do time evolution with the Burgers equation using the variational framework. We explain the matrix product state (MPS) ansatz in Sec. IV and derive its quantum circuit representation. The Monte Carlo sampling error is explained in Sec. V. In Sec. VI we describe the fidelity optimization algorithm and provide the number of variational parameters in the quantum and MPS ansatz. We use the same notation and definitions as in the main text.

I. IBM QUANTUM COMPUTATIONS

To run on IBM's noisy intermediate-term quantum computers, we need to reduce the hardware requirements and optimize the circuits shown in the main text. These optimizations lead to deviations from the generic nonlinear VQC circuit structure discussed in the main text. For computing the expectation value of an operator $\hat{O}$ that is diagonal, $\hat{O} = \hat{D}_O$, or can be diagonalized easily, i.e. $\hat{O} = \hat{D}_O = \sum_k o_k |k\rangle\langle k|$, we use a quantum circuit that does not require ancilla qubits. The expectation value of such an operator reads $\langle\psi|\hat{D}_O|\psi\rangle = \sum_k o_k \langle\psi|k\rangle\langle k|\psi\rangle = \sum_k |\langle k|\psi\rangle|^2 o_k$. This expectation value is obtained by measuring all qubits and computing the mean value after M experiments as $\langle\psi|\hat{D}_O|\psi\rangle \approx (1/M) \sum_{m=1}^{M} o_{k^{(m)}}$, as shown in Fig. S1 (a). The following hardware-reduced quantum algorithms were originally inspired by the standard methods that were used in previous experiments for the experimental computation of expectation values of quantum chemistry Hamiltonians, e.g. [1, 2].

Each measurement m gives a multi-index binary$(k)^{(m)}$ from which we obtain $k^{(m)}$. Then the expectation values read $\langle\psi|\hat{D}_{\tilde{V}}|\psi\rangle \approx (1/M) \sum_{m=1}^{M} \tilde{V}_{k^{(m)}}$ and $\langle\psi|\hat{D}_{|\psi|^2}|\psi\rangle \approx (1/M) \sum_{m=1}^{M} |\psi_{k^{(m)}}|^2$. We compute $|\psi_k|^2$ for a particular value of k by counting how often that value of k appears in M experiments and then dividing that number by M.

Figure S1 (b) shows the corresponding quantum circuit for the Laplace operator Δ. The Laplace operator is diagonalized by the Quantum Fourier Transform (QFT) and has eigenvalues $\Delta_k = 2N^2 \left(\cos(2\pi k/N) - 1\right)$ where $N = 2^n$ for n qubits. Therefore $\langle\psi|\hat{\Delta}|\psi\rangle = \langle\psi|(\hat{\text{QFT}})^\dagger \hat{D}_\Delta (\hat{\text{QFT}})|\psi\rangle = \sum_k |\psi_k|^2 \Delta_k$ where ψ_k is obtained from the QFT of $|\psi\rangle$, i.e. $(\hat{\text{QFT}})|\psi\rangle$. After we
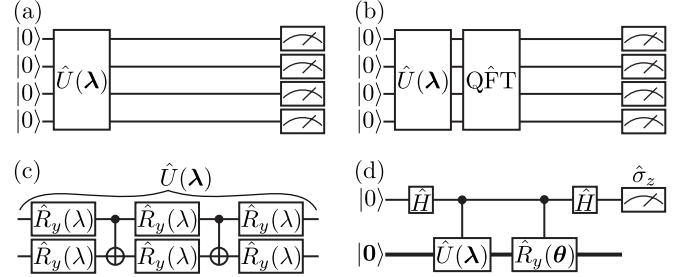

FIG. S1. (a) Quantum circuits for expectation value computation of the potential $\langle\psi|\hat{D}_{\tilde{V}}|\psi\rangle = \sum_k |\psi_k|^2 \tilde{V}_k$ or nonlinear term $\langle\psi|\hat{D}_{|\psi|^2}|\psi\rangle = \sum_k |\psi_k|^2 |\psi_k|^2$ in the nonlinear Schrödinger equation, where $|\psi\rangle = \hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$. (b) Quantum circuit for expectation value computation of the Laplace operator $\langle\psi|\hat{\Delta}|\psi\rangle$. (c) The quantum ansatz that we have used as variational state on the IBM quantum computer. Here $\hat{R}_y(\lambda) = \exp(-i\lambda\hat{\sigma}_y/2)$ is the standard y rotation gate. The quantum ansatz has just one variational parameter, i.e. $\boldsymbol{\lambda} = \lambda$. In the IBM experiment we are restricting λ to be a real parameter and therefore the quantum ansatz has only real entries. (d) Quantum circuit for evaluating specific or averaged function values $\Re\{\psi_k\}$ where $\Re\{\cdot\}$ denotes the real part. The gates $\hat{R}_y(\boldsymbol{\theta})$ are individually controlled rotations around the y-axis and depending on $\boldsymbol{\theta}$ allow reading out function values and averages of $\Re\{\psi_k\}$ by measuring the ancilla qubit in the computational basis [3, 4]. For example, individual function values can be measured efficiently at an arbitrary grid point k by choosing rotation angles $\boldsymbol{\theta} = \pi \text{ binary}(k)$. Averages over the least significant qubits can be formed by choosing the corresponding rotation angles $\pi/2$ when reading the function on a coarse grid. We obtain the imaginary part $\Im\{\psi_k\}$ by including an additional phase shift gate $R_{-\pi/2} = \text{diag}(1, \exp(-i\pi/2))$ directly before the second Hadamard gate. Function values of more general functions $F = \prod_{j=1}^{r} (O_j f^{(j)})$ are evaluated by combining the circuit presented here with the QNPU from Fig. 1 in the main text.

have applied the QFT to $|\psi\rangle = \hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$, we compute the expectation value as in Fig. S1 (a).

To illustrate that this quantum algorithm works, we chose the smallest non-trivial system size of $n = 2$ qubits and the variational ansatz of Fig. S1 (c) which is parametrized by one real parameter λ. To make use of the best available hardware, we chose to work on the 20 qubit IBM devices Tokyo and Poughkeepsie. On the Tokyo device, the best T1 is $148.5\mu s$, the best T2 is $78.4\mu s$, the best two-qubit error rate is 1.47%, and the best single-qubit error rate is 0.064% [5]. On the

Poughkeepsie device, the best T1 is $123.3\mu s$, the best T2 is $123.6\mu s$, the best two-qubit error rate is 1.11%, and the best single-qubit error rate is 0.052% [5]. These experimental errors are just a little bit too large for our quantum algorithm presented in the main text but are good enough for our hardware-efficient quantum algorithm presented here (see also [6]). We considered the nonlinear Schrödinger equation with a harmonic trap potential $V(x) = 2000(x - 0.5)^2$ and $x \in [0, 1)$. Using the variational ansatz, we evaluated the quantum circuits of Fig. S1 (a) and (b) for all values of $\lambda \in [0, 2\pi)$ in steps of 0.1. For each value of λ we computed the average of the particular quantity of interest over 24000 experimental shots. This allowed us to compute the cost function (total energy in the nonlinear Schrödinger equation) of Fig. 1 in the main text and determine its minimum as well as the corresponding solution function. Fig. 1 in the main text shows the squared absolute value of the solution functions, which is also referred to as the ground state solution (density). This is compared to the numerically exact solution obtained from imaginary time evolution [7].

Figure S1 (d) shows the general quantum circuit for evaluating specific or averaged function values. We did not use this circuit in the IBM experiment where, instead, we used the circuit of Fig. S1 (a) to compute $|\psi_0|^2$, $|\psi_1|^2$, $|\psi_2|^2$, and $|\psi_3|^2$.

Note that the QFT required here could be realized using just single-qubit gates with classical control [8], which would simplify the hardware requirements significantly. We did not exploit this simplification here since the IBM quantum computers do not offer classical control yet. Therefore we used the complete QFT quantum circuit shown in Fig. S1 (b).

II. ADDER CIRCUIT FOR KINETIC ENERGY

The quantum circuit presented in Fig. S2 is the QNPU for computing the kinetic energy in the main text via the adder operator $\hat{A}$. This operator increments the index of every wave function coefficient by one and can therefore be seen as a special case of the general quantum circuits for arithmetic operations discussed in Refs. [9, 10]. Compared with the more general circuits of Refs. [9, 10], this adder circuit requires fewer ancilla qubits and gates. We deduce from Fig. S2 that for $n > 2$ qubits, the adder circuit is composed of $n - 2$ ancilla qubits, $n - 2$ CNOT and $2n - 2$ Toffoli gates. For $n = 2$ the adder circuit requires one CNOT and one Toffoli gate, and for $n = 1$ one CNOT gate suffices.

III. NONLINEAR VQC FOR THE BURGERS EQUATION

To illustrate how a time-dependent nonlinear partial differential equation is solved, we develop a quantum al-
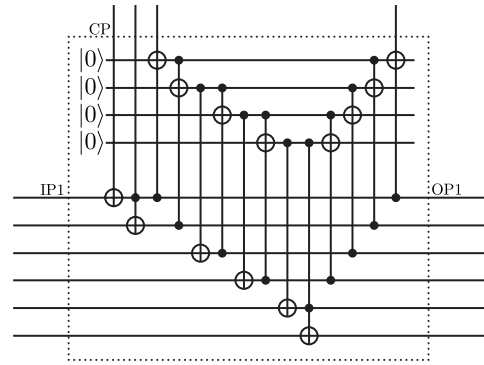

FIG. S2. QNPU for computing the kinetic energy in the main text via the adder operator $\hat{A}$. This operator acts on a state $|\psi\rangle = \sum_{k=0}^{N-1} \psi_k |\text{binary}(k)\rangle$ in such a way that $\hat{A} \sum_{k=0}^{N-1} \psi_k |\text{binary}(k)\rangle = \sum_{k=0}^{N-1} \psi_{k+1} |\text{binary}(k)\rangle$. We consider periodic boundary conditions where $\psi_N = \psi_0$. Inserting this QNPU in Fig. 1 (a) in the main text gives the circuit for determining the kinetic energy $\langle\!\langle K \rangle\!\rangle$. Note that $\hat{U}(\lambda)$ is fed on input port IP1 and in $|\psi\rangle = \hat{U}(\lambda)|0\rangle$, the lowermost qubit is the most significant qubit and the uppermost qubit is the least significant.

gorithm that approximates time evolution of the Burgers equation

$$\frac{\partial}{\partial t} f(x, t) = \nu \frac{\partial^2}{\partial x^2} f(x, t) - f(x, t) \frac{\partial}{\partial x} f(x, t). \quad (S1)$$

Here, ν is the so-called coefficient of kinematic viscosity and Eq. (S1) gives rise to turbulence when $1/\nu$ becomes large [11, 12]. We discretize the spatial coordinate x (as done in the main text of this article for the nonlinear Schrödinger equation) and we represent the resulting function values f_k at time t by a vector $|f\rangle$. Then Eq. (S1) takes on the form:

$$\frac{\partial}{\partial t} |f\rangle = \nu \Delta |f\rangle - D_f \nabla |f\rangle, \quad (S2)$$

where Δ and ∇ are the discretized Laplace and Nabla matrix, respectively, and D_f is a diagonal matrix with the values f_k on its diagonal.

For the classical simulation of Eq. (S2), we use the time-dependent variational principle algorithm for MPS [13, 14]. This algorithm requires that Δ and $|f\rangle\nabla$ are written as matrix product operators and we show how this is done in Ref. [15].

For the quantum algorithm, it is important to note that Eq. (S2) does not conserve the norm of $|f\rangle$. Therefore we choose the variational ansatz to be $|f\rangle = \lambda_0 |\psi(\lambda)\rangle$, where λ_0 is a new variational parameter and $\lambda = (\lambda_1, \lambda_2, \dots)$ as before (see discussion in main text). The introduction of λ_0 allows us to handle arbitrary norms $\langle f|f\rangle$ of the variational solution $|f\rangle$, while the wave function on the quantum computer always fulfills $\langle\psi(\lambda)|\psi(\lambda)\rangle = 1$.

To keep this discussion as simple as possible, we just consider the Euler method for time evolution [16] here.

The Euler method computes the time-evolved solution after time step τ, $|f(t+\tau)\rangle$, by using the previous solution, $|f(t)\rangle$, and the right-hand side of Eq. S2 at time t. If we summarize the right-hand side of Eq. S2 by $O|f\rangle$, then the Euler method identifies $|f(t+\tau)\rangle = (\mathbb{1} + \tau O(t))|f(t)\rangle$. In the variational setting, instead of directly computing $|f(t+\tau)\rangle$ via this formula, it is more efficient to define the cost function

$$\mathcal{C}(|f(t+\tau)\rangle) = \||f(t+\tau)\rangle - (\mathbb{1} + \tau O(t))|f(t)\rangle\|^2 \quad (S3)$$

and minimize this cost function via the variational parameters of $|f(t+\tau)\rangle$.

For the quantum algorithm we define $|f(t+\tau)\rangle = \lambda_0|\psi(\boldsymbol{\lambda})\rangle = \lambda_0\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$ and $|f(t)\rangle = \tilde{\lambda}_0|\tilde{\psi}\rangle = \tilde{\lambda}_0\hat{\tilde{U}}|\mathbf{0}\rangle$ for every value of t. Then for each time step τ, the following cost function needs to be minimized:

$$\mathcal{C}(\lambda_0,\boldsymbol{\lambda}) = \|\lambda_0|\psi(\boldsymbol{\lambda})\rangle - (\mathbb{1} + \tau\hat{O})\tilde{\lambda}_0|\tilde{\psi}\rangle\|^2 \quad (S4)$$
$$= |\lambda_0|^2 - 2\Re\{\lambda_0\tilde{\lambda}_0^*\langle\tilde{\psi}|(\mathbb{1} + \tau\hat{O})|\psi(\boldsymbol{\lambda})\rangle\} + \text{const.}$$
$$= |\lambda_0|^2 - 2\Re\{\lambda_0\tilde{\lambda}_0^*\langle\mathbf{0}|\hat{\tilde{U}}^\dagger(\mathbb{1} + \tau\hat{O})\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle\} + \text{const.},$$

where $\hat{O} = \nu\Delta - \tilde{\lambda}_0\hat{D}_{\tilde{\psi}}\nabla$. For each time step, Eq. (S4) is minimized via the procedures discussed in Refs. [2, 17]. Figure S3 shows the quantum circuits that are required for the computation of this cost function. Note that the QNPUs need to be modified after every time step, as $\hat{\tilde{U}}$ changes after every time step.

More sophisticated time evolution algorithms can readily be realized for nonlinear partial differential equations on a quantum computer. For example, a time-dependent variational principle can be formulated similarly to Refs. [18, 19] where the required new ingredients (for nonlinear terms) are obtained from the quantum circuits presented here and in the main text.

IV. MPS ANSATZ

In this section, we describe the general procedure for transforming MPS into quantum circuits. We also discuss how plane wave, cosine, and sine functions are represented on a quantum computer.

A. General Procedure

An MPS of bond dimension χ,

$$|\psi^{\text{MPS}}\rangle = \sum_{q_1,\dots,q_n} B[1]^{q_1}\dots B[n]^{q_n}|q_1,\dots,q_n\rangle, \quad (S5)$$

is defined in terms of n tensors $B[j]^{q_j}_{\alpha_{j-1},\alpha_j}$, where the index $q_j \in \{0,1\}$ and all indices α_j run from 1 to χ except α_0 and α_n which both take on only the value 1. The MPS tensor entries are real or complex numbers and constitute the variational parameters in the MPS ansatz. We
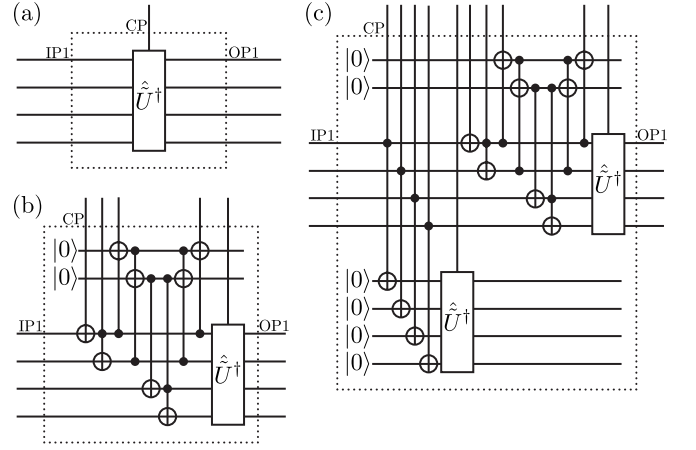

FIG. S3. QNPUs for evaluating the cost function Eq. (S4) of the Burgers equation for an example with $n = 4$. These QNPUs need to be inserted in the circuit of Fig. 1 (a) from the main text with a small modification: in this circuit the unitary on input port IP1 needs to be controlled by the ancilla (as are all unitaries on the other input ports IP2 to IPr). (a) $\langle\tilde{\psi}|\psi\rangle = \langle\mathbf{0}|\hat{\tilde{U}}^\dagger\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$. (b) $\langle\tilde{\psi}|\hat{A}|\psi\rangle = \langle\mathbf{0}|\hat{\tilde{U}}^\dagger\hat{A}\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$. (c) $\langle\tilde{\psi}|\hat{A}\hat{D}^\dagger_{\tilde{\psi}}|\psi\rangle = \langle\mathbf{0}|\hat{\tilde{U}}^\dagger\hat{A}\hat{D}^\dagger_{\tilde{\psi}}\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$. Here $\hat{\tilde{U}}^\dagger$ is the adjoint of $\hat{\tilde{U}}$ which defines $|\tilde{\psi}\rangle = \hat{\tilde{U}}|\mathbf{0}\rangle$, i.e. the function $|f(t)\rangle = \tilde{\lambda}_0|\tilde{\psi}\rangle$ at time t. Note that $\hat{\tilde{U}}^\dagger$ is fixed and has no variational parameters. The variational parameters reside outside the QNPUs in $\hat{U}(\boldsymbol{\lambda})$ which defines $|f(t+\tau)\rangle = \lambda_0|\psi\rangle = \lambda_0\hat{U}(\boldsymbol{\lambda})|\mathbf{0}\rangle$. Furthermore, in the notation here, $\hat{A}$ is the adder circuit defined in the caption of Fig. S2, and $\hat{D}_{\tilde{\psi}}$ denotes the diagonal operator that has the values of $\tilde{\psi}_k$ on its diagonal.

observe that there are $\mathcal{O}(n\chi^2) = \mathcal{O}(\text{poly}(n))$ variational parameters, bipartite entanglement entropy is limited to values $\lesssim \log(\chi)$, and most MPS algorithms have a computational cost $\mathcal{O}(\text{poly}(\chi))$ [20, 21].

Figure S4 shows the required steps for transforming a MPS of bond dimension χ into a quantum circuit. The MPS quantum circuit consists of $n - s$ unitary matrices each of dimension $2^{s+1} \times 2^{s+1} = 2\tilde{\chi} \times 2\tilde{\chi}$. Each of these unitary matrices can be further decomposed into a product of $\mathcal{O}(\tilde{\chi}^2)$ generic two-qubit unitaries [22]. We can embed these resulting two-qubit unitaries in the quantum ansatz of Fig. 1 (b) in the main text. Therefore any MPS of bond dimension $\chi \sim \text{poly}(n)$ is efficiently contained in the quantum ansatz of depth $d \sim \text{poly}(n)$. Figure S5 illustrates how the resulting quantum circuits look like for an MPS of bond dimensions $\chi = 2$ and 4. For the depths d of these quantum circuits we obtain $d = n - 1$ for $\chi = 2$ and $d = 5(n-2) + 1$ for $\chi = 4$.

B. Plane Wave, Sine, and Cosine Function Representations on a Quantum Computer

MPS contain polynomials as well as Fourier series [24, 25] and so these are also contained in the quantum ansatz
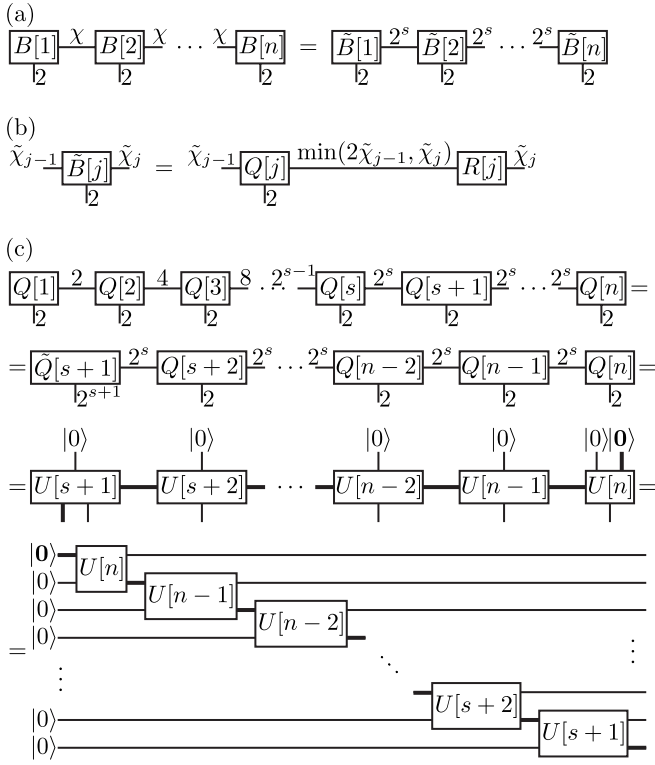
FIG. S4. Transformation from MPS into quantum circuit. (a) We first compute the value $s = \lceil \log_2(\chi) \rceil$, where $\lceil \cdot \rceil$ is the ceiling function. This allows us to rewrite the MPS with bond dimension $\tilde{\chi} = 2^s$. If $\tilde{\chi} > \chi$ then we achieve this by embedding each tensor $B[j]$ in a new larger tensor $\tilde{B}[j]$ where the additional new elements are set to zero. (b) Then successive QR decompositions turn this MPS into a product of isometric matrices (and possibly a normalization factor). We start with the QR decomposition of $\tilde{B}[1]$ and move towards and end with $\tilde{B}[n]$. After the QR decomposition of $B[j]$ we multiply $R[j]$ with $\tilde{B}[j+1]$ before performing the QR decomposition of $\tilde{B}[j+1]$. The resulting tensor network consists of isometric matrices only. (c) We multiply all tensors left of $Q[s+1]$ with $Q[s+1]$ to obtain $\tilde{Q}[s+1]$. The resulting isometric matrices are then embedded in unitary matrices $U[j]$. Rearranging the final network gives the MPS quantum circuit.

FIG. S5. Example MPS quantum circuits. (a) MPS of bond dimension $\chi = 2$ represented in terms of $n-1$ generic two-qubit unitaries. (b) MPS of bond dimension $\chi = 4$ represented in terms of $n-2$ generic three-qubit unitaries. (c) MPS of bond dimension $\chi = 4$ represented in terms of $5(n-2)+1$ generic two-qubit unitaries using the decomposition of three-qubit unitaries of Ref. [23].

FIG. S6. (a) This quantum circuit realizes the wave function $|\psi\rangle = (1/\sqrt{2^n}) \sum_{q_1,q_2,\ldots,q_n} \exp(i\kappa \sum_{j=1}^{n} q_j 2^{-j})|q_1, q_2, \ldots, q_n\rangle$ which represents a plane wave of wave vector κ. $R_\varphi = \mathrm{diag}(1, \exp(i\varphi))$ is a phase shift gate of phase shift φ. (b) Quantum circuit for the probabilistic generation of $\psi(x) \propto \cos(\kappa x)$ and $\psi(x) \propto \sin(\kappa x)$ with probability 0.5, respectively. When 0 is measured in the uppermost qubit, then $|\psi\rangle = (1/\sqrt{2^n}) \sum_{q_1,q_2,\ldots,q_n} \exp(i\kappa \sum_{j=1}^{n} q_j 2^{-j}) + \exp(-i\kappa \sum_{j=1}^{n} q_j 2^{-j})|q_1, q_2, \ldots, q_n\rangle$, and the wave function coefficients are proportional to the cosine function. When 1 is measured in the uppermost qubit, then $|\psi\rangle = (1/\sqrt{2^n}) \sum_{q_1,q_2,\ldots,q_n} \exp(i\kappa \sum_{j=1}^{n} q_j 2^{-j}) - \exp(-i\kappa \sum_{j=1}^{n} q_j 2^{-j})|q_1, q_2, \ldots, q_n\rangle$, and the wave function coefficients are proportional to the sine function.

presented in Fig. 1 (b) in the main text. A plane wave function $\exp(i\kappa x)$ can be written as an MPS of bond dimension $\chi = 1$ [24] and its quantum circuit representation requires only single-qubit gates, as can be seen in Fig. S6 (a). We write the function $\sin(\kappa x)$ as an MPS of bond dimension $\chi = 2$ [24], which is a quantum circuit of the form shown in Fig. S5 (a). We write the sum of two different sine functions as an MPS of bond dimension $\chi = 4$, which is a quantum circuit of the form shown in Fig. S5 (b) and (c). This MPS of bond dimension $\chi = 4$ is used in the main text for representing the potential $V(x) = s_1 \sin(\kappa_1 x) + s_2 \sin(\kappa_2 x)$.

Depending on the hardware capabilities, there exist several alternatives for realizing $V(x) = s_1 \sin(\kappa_1 x) + s_2 \sin(\kappa_2 x)$ on a quantum computer. Note that $\sin(\kappa x)$
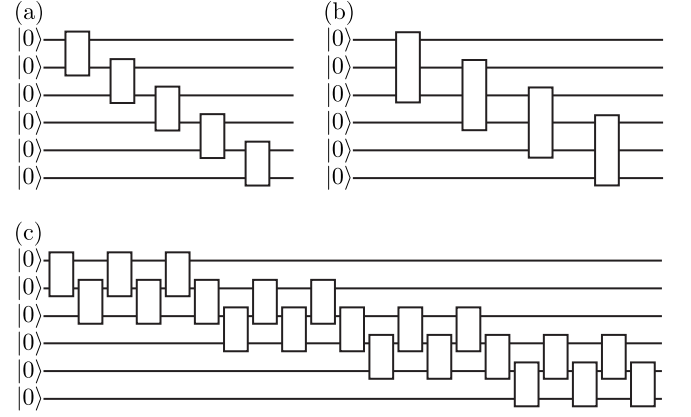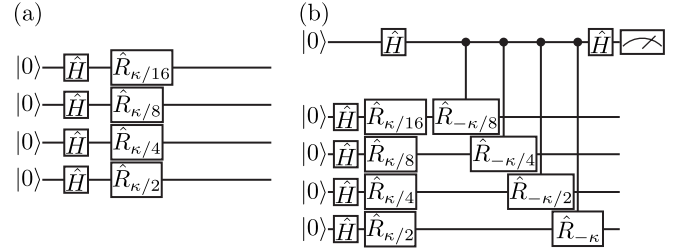
as well as $\cos(\kappa x)$ can be written as a sum of two plane waves, respectively. Therefore two quantum circuits of the form of Fig. S6 (a) suffice to represent a single function $\sin(\kappa x)$. Four quantum circuits of the form of Fig. S6 (a) represent the potential $V(x) = s_1 \sin(\kappa_1 x) + s_2 \sin(\kappa_2 x)$. Therefore, instead of a single MPS quantum circuit, we can also use several simpler circuits (composed of single-qubit unitaries only) for the potential representation on a quantum computer. Sine and cosine functions are also generated probabilistically with the quantum circuit of Fig. S6 (b).

V. SAMPLING ERROR

Here we analyze the grid error and the Monte Carlo sampling error. We first investigate the convergence of the numerically exact solution with the number of grid points. These results facilitate the analysis of the Monte Carlo sampling error.

A. Grid Error

We numerically calculate the expectation values $\langle\!\langle K\rangle\!\rangle$, $\langle\!\langle P\rangle\!\rangle$ and $\langle\!\langle I\rangle\!\rangle$ in the main text for different grid sizes $N = 2^n$. The results are presented in Fig. S7 and show that all expectation values increase with an increasing number of grid points until they level off and converge at a critical number of grid points $N_{\min} = 2^{n_{\min}}$. This critical number depends only on the value of the wavenumber κ_1 which controls the shape of the external potential. More specifically, $N_{\min}$ approximately coincides with the number of grid points required to resolve the potential

$$V(x) = s_1 \sin(\kappa_1 x) + s_2 \sin(\kappa_2 x). \qquad (S6)$$

By taking into account that one needs at least four grid points to resolve a single wave length, we obtain $n_{\min} = 6$ for $\kappa_1 = 2\pi \times 16$, $n_{\min} = 7$ for $\kappa_1 = 2\pi \times 32$, $n_{\min} = 8$ for $\kappa_1 = 2\pi \times 64$, and $n_{\min} = 9$ for $\kappa_1 = 2\pi \times 128$. These simple estimates are in very good agreement with the results in Fig. S7. More generally, we define

$$N_{\min} = \frac{\ell}{\ell_{\min}}, \qquad (S7)$$

where $\ell_{\min}$ is the smallest lengthscale of the problem. The expectation values $\langle\!\langle \cdot \rangle\!\rangle$ converge if the grid size obeys $N > N_{\min}$, and only in this regime the grid error scales like $\mathcal{E}_{\mathrm{grid}} \propto 1/N^2$.

The expectation values $\langle\!\langle X\rangle\!\rangle$ are related in the main text to expectation values of an ancilla qubit $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^X$ when measured on a quantum computer. The results for $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^X$ are also shown in Fig. S7 and illustrate how the values $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^X$ converge for $N > N_{\min}$. We observe that $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^P$ converges to a constant value and

$$1 - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K \propto 2^{-2n}, \qquad (S8a)$$

$$\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^I \propto 2^{-n}, \qquad (S8b)$$

for $n > n_{\min}$. In particular, we thus conclude $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K \to 1$ and $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^I \to 0$ for $n \to \infty$. Note that the expected scaling in Eq. (S8) can be used in practice to determine the minimal grid size $N_{\min}$. This is important for finding the optimal grid size that results in converged results with reasonable Monte Carlo sampling sizes, as discussed in the next section.

B. Monte Carlo Sampling Error

The expectation values $\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}$ in Eqs. (6) and (7) of the main text represent exact quantum mechanical ex-

pectation values. Measuring these quantities on a quantum computer involves averaging over M Monte Carlo samples,

$$\langle\overline{\hat{\sigma}_z}\rangle_{\mathrm{anc}} = \frac{1}{M} \sum_{m=1}^{M} \langle\hat{\sigma}_z\rangle_m, \qquad (S9)$$

where $\langle\hat{\sigma}_z\rangle_m$ denotes the eigenvalue of $\hat{\sigma}_z$ obtained for the mth sample. In general, the Monte Carlo sampling error $\mathcal{E}_{\mathrm{MC}} = \langle\overline{\hat{\sigma}_z}\rangle_{\mathrm{anc}} - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}$ associated with drawing M random samples is [16]

$$\mathcal{E}_{\mathrm{MC}} = \frac{\sqrt{\langle\hat{\sigma}_z^2\rangle_{\mathrm{anc}} - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^2}}{\sqrt{M}}, \qquad (S10a)$$

$$= \frac{\sqrt{1 - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^2}}{\sqrt{M}}, \qquad (S10b)$$

where we used $\hat{\sigma}_z^2 = \mathbb{1}$ in the second line.

With the help of these definitions we obtain

$$\mathcal{E}_{\mathrm{MC}}^P = \alpha \frac{\sqrt{1 - [\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^P]^2}}{\sqrt{M}}, \qquad (S11a)$$

$$\mathcal{E}_{\mathrm{MC}}^K = \frac{N^2}{\ell^2} \frac{\sqrt{1 - [\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K]^2}}{\sqrt{M}}, \qquad (S11b)$$

$$\mathcal{E}_{\mathrm{MC}}^I = \frac{1}{2} g \frac{N}{\ell} \frac{\sqrt{1 - [\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^I]^2}}{\sqrt{M}}, \qquad (S11c)$$

where $\mathcal{E}_{\mathrm{MC}}^X$ is the absolute Monte Carlo error associated with the average $\langle\!\langle X\rangle\!\rangle$. The expressions in Eq. (S11) can be further simplified, and we begin with the error associated with the potential energy $\langle\!\langle P\rangle\!\rangle$. With the help of Eq. (7)(a) in the main text, the relative Monte Carlo sampling error can be written as

$$\epsilon_{\mathrm{MC}}^P = \frac{\mathcal{E}_{\mathrm{MC}}^P}{\langle\!\langle P\rangle\!\rangle} = C_P \frac{1}{\sqrt{M}}, \qquad (S12)$$

where

$$C_P = \sqrt{\frac{\alpha^2}{\langle\!\langle P\rangle\!\rangle^2} - 1}. \qquad (S13)$$

Since $\alpha^2/\langle\!\langle P\rangle\!\rangle^2 \geq 1$ by virtue of the definition of α in the main text, C_P is always real. For the parameters of Fig. S7 C_P takes on values between 64.3 and 67.2.

Next we investigate the error associated with the kinetic energy term $\langle\!\langle K\rangle\!\rangle$ and note that

$$1 - [\langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K]^2 = \left(1 - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K\right)\left(1 + \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K\right), \qquad (S14a)$$

$$1 - \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K = h_N^2 \langle\!\langle K\rangle\!\rangle. \qquad (S14b)$$

Combining the equations above we obtain the relative sampling error associated with $\langle\!\langle K\rangle\!\rangle$,

$$\epsilon_{\mathrm{MC}}^K = \frac{\mathcal{E}_{\mathrm{MC}}^K}{\langle\!\langle K\rangle\!\rangle} = \frac{N}{\ell} \frac{\sqrt{1 + \langle\hat{\sigma}_z\rangle_{\mathrm{anc}}^K}}{\sqrt{\langle\!\langle K\rangle\!\rangle}} \frac{1}{\sqrt{M}}. \qquad (S15)$$
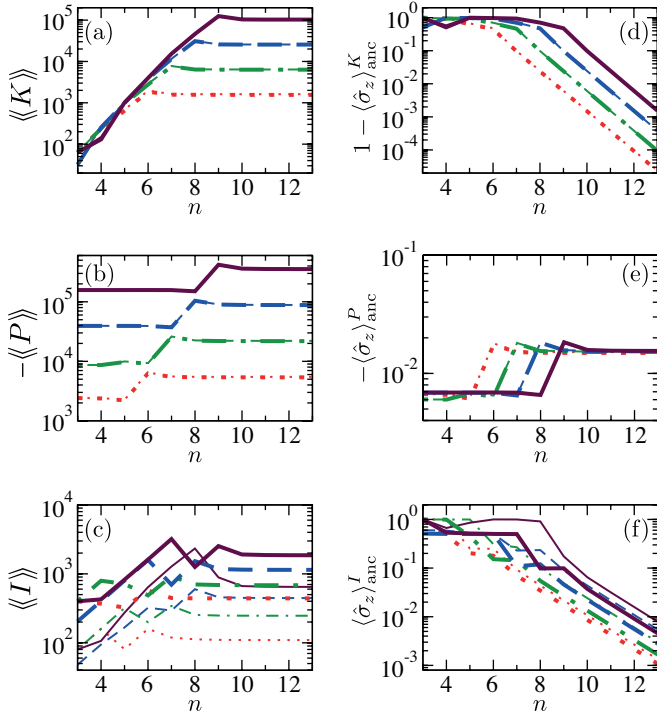
FIG. S7. Energy expectation values $\langle\langle K\rangle\rangle = 4^n\Re\{\langle\psi|(\mathbb{1} - \hat{A})|\psi\rangle\}$ (a), $\langle\langle P\rangle\rangle = \Re\{\langle\psi|\hat{V}|\psi\rangle\}$ (b), $\langle\langle I\rangle\rangle = g2^n\Re\{\langle\psi|\hat{D}_{|\psi|^2}|\psi\rangle\}$ (c), and the corresponding values related to the ancilla qubit measurement: $1 - \langle\hat{\sigma}_z\rangle^K_{\rm anc} = \langle\langle K\rangle\rangle/4^n$ (d), $\langle\hat{\sigma}_z\rangle^P_{\rm anc} = \langle\langle P\rangle\rangle/\alpha$ (e), and $\langle\hat{\sigma}_z\rangle^I_{\rm anc} = \langle\langle I\rangle\rangle/(g2^n)$ (f). Here $\hat{A}$ is the adder operator defined in the caption of Fig. S2, $\hat{V}$ represents the potential $V(x) = s_1\sin(\kappa_1 x) + s_2\sin(\kappa_2 x)$, and $\hat{D}_{|\psi|^2}$ is diagonal with the values of $|\psi_k|^2$ on its diagonal. We consider the strongly disordered regime of Fig. 3 in the main text (where $s_1/s_2 = 2$ and $\kappa_2 = 2\kappa_1/(1+\sqrt{5})$) with $s_1 = 5\times 10^3$ and $\kappa_1 = 2\pi\times 16$ (red dotted), $s_1 = 2\times 10^4$ and $\kappa_1 = 2\pi\times 32$ (green dash-dotted), $s_1 = 8\times 10^4$ and $\kappa_1 = 2\pi\times 64$ (blue dashed), $s_1 = 3.2\times 10^5$ and $\kappa_1 = 2\pi\times 128$ (purple solid), and we compare $g = 10$ (thin lines) with 50 (normal lines). All results are for $N = 2^{13}$ grid points.

To further simplify this we assume that $N \geq N_{\rm min}$ such that $\langle\hat{\sigma}_z\rangle^K_{\rm anc} \approx 1$ [see Eq. (S8)] and utilize Eq. (S7),

$$\epsilon^K_{\rm MC} \approx C_K \frac{N}{N_{\rm min}} \frac{1}{\sqrt{M}}, \qquad (\text{S16})$$

where

$$C_K = \frac{\sqrt{2}}{l_{\rm min}\sqrt{\langle\langle K\rangle\rangle}}. \qquad (\text{S17})$$

The constant C_K is expected to be of the order of unity since $\langle\langle K\rangle\rangle \approx 1/\ell^2_{\rm min}$. For the parameters of Fig. S7 C_K takes on values between 2.24 and 2.30.

Finally, we consider the nonlinear term and point out that $\langle\hat{\sigma}_z\rangle^I_{\rm anc} \to 0$ for $N \to \infty$, see Eq. (S8). The leading

term in Eq. (S11c) is thus $\mathcal{E}^I_{\rm MC} \approx gN/(2\ell\sqrt{M})$, and the relative Monte Carlo sampling error is

$$\epsilon^I_{\rm MC} = \frac{\mathcal{E}^I_{\rm MC}}{\langle\langle I\rangle\rangle} \approx C_I \frac{N}{N_{\rm min}} \frac{1}{\sqrt{M}}, \qquad (\text{S18})$$

where

$$C_I \approx \frac{gN_{\rm min}}{2\ell\langle\langle I\rangle\rangle}. \qquad (\text{S19})$$

For the parameters of Fig. S7 C_I takes on values between 2.58 and 6.88.

The preceeding analysis concludes the proof of Eq. (8) in the main text. For $N < N_{\rm min}$ all observables require of the order of $M \approx \epsilon^{-2}_{\rm MC}$ Monte Carlo samples in order to be accurately measured. The same result holds for the potential term even if $N \geq N_{\rm min}$. On the other hand, $\langle\langle K\rangle\rangle$ and $\langle\langle I\rangle\rangle$ require $M \approx 4^n\epsilon^{-2}_{\rm MC}$ experiments and thus M increases exponentially with n. The grid error also tends to zero exponentially fast $\propto 4^{-n}$ for $n > n_{\rm min}$, and thus n only needs to be slightly larger than $n_{\rm min}$ in order to achieve accurate solutions. This avoids exponentially many Monte Carlo samples.

Note that the minimal grid size $N_{\rm min}$ can be determined by gradually increasing N until the measured expectation values $\langle\hat{\sigma}_z\rangle_{\rm anc}$ show the expected scaling behaviour in Eq. (S8). In this way optimal grid sizes resulting in accurate solutions with moderate Monte Carlo sampling sizes can be chosen.

VI. FIDELITY $\mathcal{F}$ OPTIMIZATION AND NUMBER OF VARIATIONAL PARAMETERS $\mathcal{N}$

In the main text of this article we determine optimal approximations in terms of MPS $|\psi^{\rm MPS}\rangle$ and in terms of our quantum ansatz $|\psi^{\rm QA}\rangle$ for numerically exact solutions $|\psi^{\rm exact}\rangle$. We achieve this by maximizing $\mathcal{F} = \langle\psi^{\rm exact}|\psi^{\rm MPS}\rangle$ as well as $\mathcal{F} = \langle\psi^{\rm exact}|\psi^{\rm QA}\rangle$ using standard tensor network optimization techniques [20, 21]. Our algorithm optimizes the variational unitaries one after another and for each unitary U_j determines the new unitary $\tilde{U}_j$ that maximizes $\mathcal{F}$ under the assumption that all other unitaries are fixed. Because $\mathcal{F} = {\rm tr}(U_j R_j)$ we obtain the optimal $\tilde{U}_j$ from a singular value decomposition of $R_j = U_{R_j}\Sigma_{R_j}V_{R_j}$ as $\tilde{U}_j = V^\dagger_{R_j}U^\dagger_{R_j}$.

For real-valued unitary gates (as considered throughout our analysis), our quantum ansatz of Fig. 1 (b) in the main text has 3 variational parameters per two-qubit unitary in the first column and then 6 variational parameters per two-qubit unitary in all the other columns. A MPS tensor $B[j]^{q_j}_{\alpha_{j-1},\alpha_j}$ has $\dim(\alpha_{j-1})\dim(\alpha_j)\dim(q_j) - \dim(\alpha_j)(\dim(\alpha_j) + 1)/2$ variational parameters. This is derived by reshaping $B[j]^{q_j}_{\alpha_{j-1},\alpha_j} = B[j]_{(\alpha_{j-1},q_j),\alpha_j}$ as an isometric matrix with row-muliindex (α_{j-1},q_j) and column-index α_j. This reshaped form allows us to count the number of free parameters of such an isometry.

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nat. Commun. **5**, 4213 (2014).

[2] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," Nature **549**, 242 (2017).

[3] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, "Direct Estimations of Linear and Nonlinear Functionals of a Quantum State," Phys. Rev. Lett. **88**, 217901 (2002).

[4] C. M. Alves, P. Horodecki, D. K. L. Oi, L. C. Kwek, and A. K. Ekert, "Direct estimation of functionals of density operators by local operations and classical communication," Phys. Rev. A **68**, 032306 (2003).

[5] Https://www.ibm.com/blogs/research/2019/03/power-quantum-device/.

[6] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, "Validating quantum computers using randomized model circuits," Phys. Rev. A **100**, 032328 (2019).

[7] E. H. Lieb and J. Yngvason, "Ground State Energy of the Low Density Bose Gas," Phys. Rev. Lett. **80**, 2504 (1998).

[8] R. B. Griffiths and C.-S. Niu, "Semiclassical Fourier Transform for Quantum Computation," Phys. Rev. Lett. **76**, 3228 (1996).

[9] V. Vedral, A. Barenco, and A. Ekert, "Quantum networks for elementary arithmetic operations," Phys. Rev. A **54**, 147 (1996).

[10] M. A Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[11] J. Bec and K. Khanin, "Burgers turbulence," Physics Rep. **447**, 1 (2007).

[12] G. B. Whitham, *Linear and Nonlinear Waves* (John Wiley and Sons, New York, USA, 2011).

[13] C. Lubich, I. V. Oseledets, and B. Vandereycken, "Time Integration of Tensor Trains," SIAM J. Numer. Anal. **53**, 917 (2015).

[14] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, "Unifying time evolution and optimization with matrix product states," Phys. Rev. B **94**, 165116 (2016).

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

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, USA, 1992).

[17] I. H. Kim and B. Swingle, "Robust entanglement renormalization on a noisy quantum computer," arXiv preprint, arXiv:1711.07500v1 (2017).

[18] Y. Li and S. C. Benjamin, "Efficient Variational Quantum Simulator Incorporating Active Error Minimization," Phys. Rev. X **7**, 021050 (2017).

[19] S. McArdle, T. Jones, S. Endo, Y. Li, S. Benjamin, and X. Yuan, "Variational quantum simulation of imaginary time evolution," arXiv preprint, arXiv:1804.03023v3 (2018).

[20] F. Verstraete, V. Murg, and J. I. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," Adv. Phys. **57**, 143 (2008).

[21] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," Ann. Phys. **349**, 117 (2014).

[22] V. V. Shende, S. S. Bullock, and I. L. Markov, "Synthesis of quantum-logic circuits," IEEE Trans. on Computer-Aided Design **25**, 1000 (2006).

[23] 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).

[24] B. N. Khoromskij, "$O(d \log N)$-Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling," Constr. Approx. **34**, 257 (2011).

[25] I. V. Oseledets, "Constructive Representation of Functions in Low-Rank Tensor Formats," Constr. Approx. **37**, 1 (2013).