

Entanglement scaling in matrix product state representation of smooth functions and their shallow quantum circuit approximations

Vladyslav Bohun^{1,*}, Illia Lukin^{1,2,*}, Mykola Luhanko^{1,3}, Georgios Korpas^{4,5,6}, Philippe J. S. De Brouwer,⁷
Mykola Maksymenko¹, and Maciej Koch-Janusz^{1,8,†}

¹*Haiqu, Inc., 95 Third Street, San Francisco, California 94103, USA*

²*Akhiezer Institute for Theoretical Physics, NSC KIPT, Akademichna 1, 61108 Kharkiv, Ukraine*

³*V. N. Karazin Kharkiv National University, Kharkiv, Ukraine*

⁴*HSBC Lab, Emerging Technologies, Innovation & Ventures, Singapore, Singapore*

⁵*Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic*

⁶*Archimedes Research Unit on AI, Data Science and Algorithms, Athena Research Center, Marousi, Greece*

⁷*HSBC Service Delivery Sp. z o.o., 30-347 Kraków, Poland*

⁸*Department of Physics, University of Zürich, 8057 Zürich, Switzerland*



(Received 1 April 2025; accepted 2 March 2026; published 20 April 2026)

Encoding classical data in a quantum state is a key prerequisite of many quantum algorithms. Recently, matrix product state (MPS) methods emerged as the most promising approach for constructing shallow quantum circuits approximating input functions, including probability distributions, with only linear number of gates. We derive rigorous asymptotic expansions for the decay of entanglement across bonds in the MPS representation depending on the smoothness of the input function, real or complex. We also consider the dependence of the entanglement on localization properties and function support. Based on these analytical results, we construct an improved MPS-based algorithm yielding shallow and accurate encoding quantum circuits. By using tensor cross interpolation, we are able to construct utility-scale quantum circuits in a compute- and memory-efficient way. We validate our methods by loading heavy-tailed distributions, including Lévy, important in finance, but they apply to any smooth function inputs. We test the performance of the resulting quantum circuits by executing and sampling from them on IBM quantum devices, for up to 156 qubits.

DOI: [10.1103/yyvr-dtsb](https://doi.org/10.1103/yyvr-dtsb)

I. INTRODUCTION

The first step of many quantum algorithms is state preparation, encoding the classical input data in, e.g., the amplitudes of a quantum state. For a general input, the depth of the quantum circuit preparing a corresponding quantum state may scale exponentially [1], making this key step intractable on real quantum devices. To alleviate this problem, different strategies have been proposed, including general methods based on quantum generative adversarial networks [2], quantum Fourier transform [3–5], or handcrafted approaches [6], but all of these have shortcomings in terms of scaling, circuit complexity, trainability, or generality, limiting their usefulness in practice. Fortunately, for important classes of data accurate approximate and only linearly deep circuits can be systematically constructed. This is based on two properties: First, the representability of the input data as a matrix product

state (MPS) [7,8] (more generally: a tensor network) of *small bond dimension*. Second, the ability to convert a small bond dimension MPS into a quantum circuit with gate count *linear* in the number of qubits [9–14].

We focus on the general problem of representing one-dimensional real or complex input functions with a quantum circuit. These may represent, e.g., the initial state for a quantum partial differential equation (PDE) solver [15], or a probability density function (PDF). The two key questions we address are how the analytical properties of the input influence the entanglement structure in the intermediate MPS representation, and how this structure can be used to yield even shallower approximating quantum circuits.

For the first question, it is known that while a generic function may require an MPS approximation of an exponential rank (bond dimension), polynomials are represented *exactly* with small bond MPS [16]. Further, results for ranks and error bounds of low-rank MPS approximations exist for smooth (infinitely differentiable) functions [3,17–20], some of which were used in the previous works on distribution encoding [21–26]. Here, we significantly strengthen these analyses by deriving results on the precise structure of the entanglement in the MPS approximation in terms of entanglement spectra and purities. In particular, we show universal scaling relations of these entanglement measures for smooth functions, and their breakdown in the absence of smoothness.

*These authors contributed equally to this work.

†Contact author: maciej@haiqu.ai

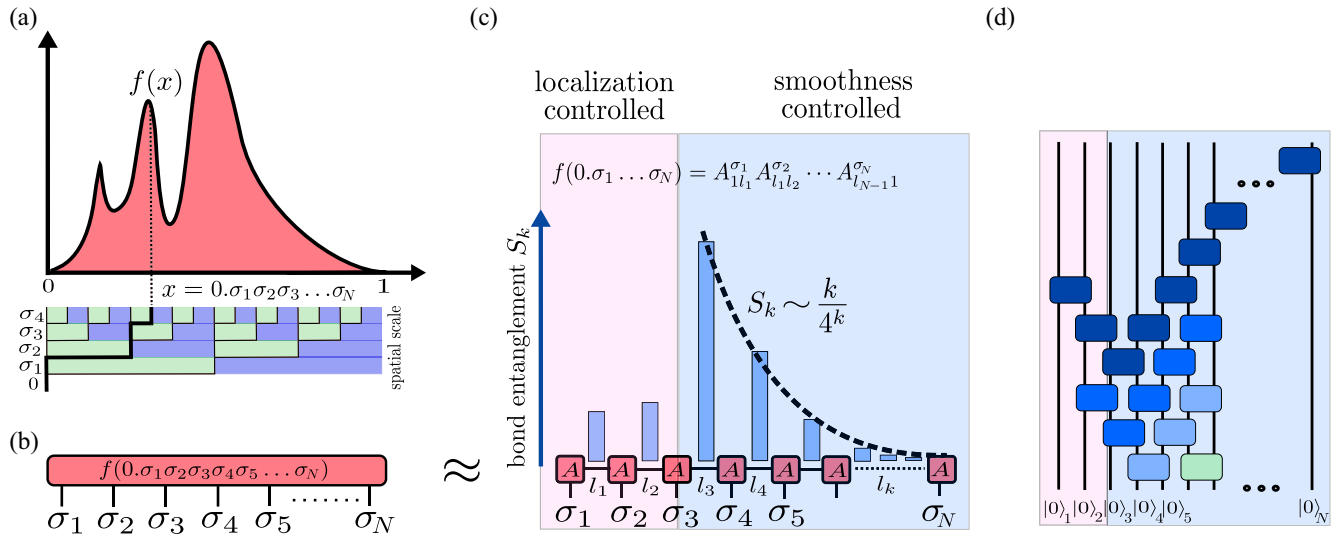


FIG. 1. (a) A real or complex function $f(x)$ on $[0, 1]$ is discretized on a binary grid $x = \sum_i \sigma_i 2^{-i}$. Successive digits σ_i of the binary expansion describe decreasing spatial scales: Parts of the interval $[0, 1]$ sharing the same value of σ_i are shown in green and blue. (b) Discretized function can be represented as a rank- N tensor $f_{\sigma_1 \sigma_2 \dots \sigma_N}$. (c) An MPS can approximate $f_{\sigma_1 \sigma_2 \dots \sigma_N}$, but the bond dimensions l_k , and thus the entanglement across bond k , may be exponential in general case. We show that for *smooth functions*, the entanglement *decays exponentially* with k , effectively decoupling fine spatial scales. For the very largest scales, the behavior is nonuniversal and depends on the localization properties of the function. (d) We use these analytical results to construct optimized quantum circuits with $O(N)$ gates generating the state encoding the function f . The algorithm minimizes the number of two-qubit gates based on the entanglement present at different scales.

These allow us to derive analytically functionals controlling, e.g., the fidelity of a single-layer approximate encoding circuits.

Armed with these results on entanglement generated by approximate function representations, we tackle the second question. We introduce an improved algorithm for generating encoding quantum circuits with reduced depth compared to previous MPS-based proposals. We also introduce a number of additional optimizations based on, e.g., the compilation of isometries, further reducing the gate count. This is summarized in Fig. 1.

Finally, we deploy and test our algorithm in practice on the example of encoding probability distributions. This important case is motivated by their role in quantum finance routines [27,28], particularly of the α -stable ones, i.e., the Lévy distribution, which are used to effectively model the fat-tailed or skewed data in risk assessment, portfolio management, and derivative pricing [29–35]. We test our methods by executing the circuits on real and simulated IBM quantum devices on up to 156 qubits. In particular, we execute and sample from the quantum circuits generated by our method on the *ibm_torino* quantum processing unit (QPU) on up to 25 qubits, with the generated distributions passing standard statistical tests (e.g., Kolmogorov-Smirnov). We also execute utility-scale circuits for up to 156 qubits on the *ibm_marrakesh* and *ibm_kingston* QPUs. These we generate using a scalable compute- and memory-efficient tensor cross interpolation (TCI) [36] extension of our algorithm, obtaining qualitative agreement with the exact distributions even for these large utility-scale problems.

The paper is structured as follows: In Sec. II, we introduce the problem and the necessary background. In Sec. III, we state our theoretical results, numerical experiments, and discuss their implications. In Sec. IV, we give an improved

circuit construction algorithm based on them. In Sec. V, we benchmark the results on real and simulated quantum hardware, and in Sec. VI we discuss future directions. Appendixes give mathematical details of the proofs and further algorithmic details.

II. PRELIMINARIES: MPS APPROXIMATION OF FUNCTIONS

A. The MPS representation

Our goal is to encode a function $f(x) \rightarrow \mathbb{C}$ with $x \in [0, 1]$, normalized in L_2 norm: $\int_0^1 |f(x)|^2 dx = 1$. To this end, $f(x)$ is discretized on an N bit binary grid [see Fig. 1(a)]:

$$x_\sigma = \sum_{i=1}^N \frac{\sigma_i}{2^i} \equiv 0.\sigma_1\sigma_2\dots\sigma_N, \quad (1)$$

where $\sigma = (\sigma_1, \dots, \sigma_N)$ with $\sigma_i \in \{0, 1\}$. Evaluating $f(x)$ on this grid yields a rank- N tensor $f_{\sigma_1 \sigma_2 \dots \sigma_N}$, with an $O(1/2^N)$ discretization error. These 2^N values will be (approximately) encoded, using a quantum circuit, in an N -qubit quantum state $|f\rangle$ [see Fig. 1(b)]:

$$|f\rangle = \sum_{\sigma_1, \dots, \sigma_N=0}^1 \frac{f_{\sigma_1 \sigma_2 \dots \sigma_N}}{\sqrt{2^N}} |\sigma_1 \sigma_2 \dots \sigma_N\rangle, \quad (2)$$

where $\sqrt{2^N}$ ensures the proper normalization. Note that this encoding is hierarchical: The largest indices (qubits) $\dots, \sigma_{N-1}, \sigma_N$ describe the function on the smallest scales of order $1/2^N$, while the smallest indices $\sigma_1, \sigma_2, \dots$ account for the largest scales.

The exponentially sized tensor f can be rewritten and decomposed into an MPS [7,8], also known as quantized tensor

train [37] [see Fig. 1(c)]:

$$|f\rangle \approx \sum_{\sigma_1, \dots, \sigma_N=0}^1 c_{\sigma_1 \sigma_2 \dots \sigma_N} |\sigma_1 \sigma_2 \dots \sigma_N\rangle, \quad (3)$$

where

$$c_{\sigma_1 \sigma_2 \dots \sigma_N} = A_{1,l_1}^{\sigma_1} A_{2,l_1 l_2}^{\sigma_2} \dots A_{N-1,l_{N-2} l_{N-1}}^{\sigma_{N-1}} A_{N,l_{N-1}}^{\sigma_N}, \quad (4)$$

and where $A_{k,l_{k-1} l_k}^{\sigma_k}$ are rank-3 tensors with a “physical” index σ_k and “internal” indices l_{k-1}, l_k . The summation range of the contracted internal indices (we use the Einstein summation convention), also known as bonds, is called the bond dimension χ_k . Equivalently, the above tensor decomposition can be written in the so-called canonical form as

$$\begin{aligned} c_{\sigma_1 \sigma_2 \dots \sigma_N} &= \Gamma_{1,l_1}^{\sigma_1} \Lambda_{1,l_1} \Gamma_{2,l_1 l_2}^{\sigma_2} \Lambda_{2,l_2} \dots \Gamma_{N-1,l_{N-2} l_{N-1}}^{\sigma_{N-1}} \Lambda_{N-1,l_{N-1}} \Gamma_{N,l_{N-1}}^{\sigma_N}, \end{aligned} \quad (5)$$

where Γ_k are also rank-3 tensors, while Λ_k are positive diagonal matrices, which obey certain isometrical conditions (for a thorough introduction to MPS, see Refs. [7,38–40]). These canonical form conditions have the form

$$\sum_{l_k, \sigma_k} \Gamma_{k,l_{k-1} l_k}^{\sigma_k} \Lambda_{k,l_k}^2 \Gamma_{k,l_{k-1} l_k}^{\dagger \sigma_k} = \delta_{l_{k-1}, l_{k-1}}, \quad (6)$$

$$\sum_{l_{k-1}, \sigma_k} \Gamma_{k,l_{k-1} l_k}^{\sigma_k} \Lambda_{k-1,l_{k-1}}^2 \Gamma_{k,l_{k-1} l_k}^{\dagger \sigma_k} = \delta_{l_k, l_k}. \quad (7)$$

Importantly, the entries of the diagonal matrices Λ in the canonical form encode the MPS entanglement spectra. To wit, the entanglement entropy for the bipartition on the bond k is given by

$$S_k = - \sum_{i=0}^{\chi_k-1} \Lambda_{k,i}^2 \log \Lambda_{k,i}^2, \quad (8)$$

where the summation range over virtual bond dimension index i starts from 0. We will also use a different entanglement measure, the purity, which is easier to compute for an arbitrary function. The purity on bond k is given by $p_k = \text{Tr} \rho_k^2$, where ρ_k is the reduced density matrix for the subsystem $\sigma_1 \sigma_2 \dots \sigma_k$. It characterizes how close the state is to the product state. In particular, if the initial wave function factorizes with respect to the bond k , then the purity is equal to one. The purities p_k are given by

$$p_k = \sum_{i=0}^{\chi_k-1} \Lambda_{k,i}^4. \quad (9)$$

In what follows, the MPS approximation of the function will be converted to a shallow quantum circuit preparing the encoded function on a QPU [Fig. 1(d)] using our improved algorithm (see Sec. IV), which extends the methods of Refs. [21–23,25,26] (see also Ref. [41]).

We will also find it useful in Sec. IV, the center canonical form (also known as mixed canonical form; see Ref. [7]). In this form, some bond (and the corresponding Λ matrix) is chosen as central. Then, all the Λ matrices on the other bonds are absorbed into their neighboring Γ tensors, so that resulting

combined tensors become isometries. The MPS thus obtained (here with the k th bond being central) has the form

$$A_{L,1,l_1}^{\sigma_1} \dots A_{L,k,l_{k-1} l_k}^{\sigma_k} \Lambda_{k,l_k} A_{R,k+1,l_k l_{k+1}}^{\sigma_{k+1}} \dots A_{R,N,l_{N-1}}^{\sigma_N}, \quad (10)$$

where the tensors A_L and A_R obey the left isometrical and right isometrical conditions, respectively:

$$\sum_{l_k, \sigma_k} A_{R,k,l_{k-1} l_k}^{\sigma_k} A_{R,k,l_{k-1} l_k}^{\dagger \sigma_k} = \delta_{l_{k-1}, l_{k-1}}, \quad (11)$$

$$\sum_{l_{k-1}, \sigma_k} A_{L,k,l_{k-1} l_k}^{\sigma_k} A_{L,k,l_{k-1} l_k}^{\dagger \sigma_k} = \delta_{l_k, l_k}. \quad (12)$$

The position of the center in the center canonical form can be shifted by absorbing the central bond Λ matrix into the neighboring A_L or A_R tensors and proceeding with singular value decomposition (SVD) of the resulting tensor. To wit, to shift the center position from bond k to bond $k-1$, we first define a new tensor $A_{C,k,l_{k-1} l_k}^{\sigma_k} = A_{L,k,l_{k-1} l_k}^{\sigma_k} \Lambda_{k,l_k}$, which we then decompose with SVD: $A_{C,k,l_{k-1} l_k}^{\sigma_k} = U_{l_{k-1} l_{k-1}}^{\sigma_k} \Lambda_{k-1,l_{k-1}}^{\sigma_k} A_{R,k,l_{k-1} l_k}^{\sigma_k}$, where Λ is the new central bond Λ matrix, $A_{R,k}$ is the new right-isometric tensor on the k th qubit, and l_{k-1}' is the new index on $(k-1)$ th bond. Finally, the matrix $U_{l_{k-1} l_{k-1}}^{\sigma_k}$ is absorbed into the tensor $A_{L,k-1}$: $A_{L,k-1,l_{k-2} l_{k-1}}^{\sigma_{k-1}} \rightarrow A_{L,k-1,l_{k-2} l_{k-1}}^{\sigma_{k-1}} U_{l_{k-1} l_{k-1}}^{\sigma_k} = A_{L,k-1,l_{k-2} l_{k-1}}^{\sigma_{k-1}}$.

B. Constructing the MPS

While the MPS decomposition can be exact, for arbitrary tensors the MPS ranks (bond dimensions) may be exponential in the number of qubits. Fortunately, in many important cases this is not the case, as, e.g., when representing the ground states of quantum Hamiltonians [7]. Thus, if the entanglement in $|f\rangle$ in Eq. (2) is not large, the state can be effectively MPS approximated with polynomial resources by truncating to small bond dimensions in Eqs. (3)–(5). Later, we show that this holds for smooth functions.

The MPS decomposition itself is carried out either with SVD or approximate methods [20,42–45]. The SVD-based method works as follows: The tensor $c_{\sigma_1 \sigma_2 \dots \sigma_N}$ is interpreted as a matrix $c^{\sigma_1 \rho}$ whose first index is σ_1 , and the second one comprises all of the remaining ones: $\rho = (\sigma_2 \dots \sigma_N)$. SVD decomposes this matrix into the form

$$c^{\sigma_1 \rho} = A_{1,l_1}^{\sigma_1} s_{l_1} V_{l_1}^{\rho} = A_{1,l_1}^{\sigma_1} \Lambda_{1,l_1} V_{l_1}^{\sigma_2 \sigma_3 \dots \sigma_N}.$$

The tensor $\Lambda_{1,l_1} V_{l_1}^{\sigma_2 \sigma_3 \dots \sigma_N}$ can again be interpreted as a matrix with indices σ_2 and $(\sigma_3 \dots \sigma_N)$ and be decomposed with SVD once more: $\Lambda_{1,l_1} V_{l_1}^{\sigma_2 \sigma_3 \dots \sigma_N} = A_{2,l_1 l_2}^{\sigma_2} \tilde{V}_{l_2}^{\sigma_3 \dots \sigma_N}$. This is iterated until the original tensor c is fully factorized into the product of tensors $A_{k,l_{k-1} l_k}^{\sigma_k}$ and a residual tensor $A_{N,l_{N-1}}^{\sigma_N} = \Lambda_{N-1,l_{N-1}} V_{l_{N-1}}^{\sigma_N}$. This factorized form can be exact, but in an MPS approximation SVD spectra Λ_l are truncated below a small threshold, which we set to 10^{-12} . This is how Eq. (4) is obtained.

The MPS in a canonical form [Eq. (5)] is obtained by inserting $1 = \Lambda_{\alpha} \cdot \Lambda_{\alpha}^{-1}$ on the bonds, and contracting the inverse Λ_{α}^{-1} with the following A tensor, obtaining the Γ tensors. Due to the SVD spectra truncation, the canonical conditions [Eqs. (6) and (7)] may not be satisfied exactly. To correct this, one more MPS sweep is performed.

The main limitation of the SVD method is that the original tensor $c_{\sigma_1\sigma_2\dots\sigma_N}$ has to be stored explicitly, implying an $O(2^N)$ memory scaling. This can only be done for moderate numbers of qubits (~ 30). For larger problems, the MPS has to be constructed differently. We use (TCI) [36,45], which allows an MPS approximation to be found provided only access to a callable function (i.e., the ability to evaluate the function at different points) without requiring to store the complete set of function values. TCI constructs the MPS representation of the function by sampling of the function's values, which it uses for interpolation. As these are the only values stored, it drastically reduces the memory requirements, allowing to construct MPS representation for large qubit numbers. The number of function evaluations in TCI scales with the number of qubits as $O(N\chi^2)$, where χ is the maximal MPS bond dimension. This is in contrast to the standard SVD approach, which involves the calculation of a whole state vector with 2^N function evaluations. As its name implies, TCI is a generalization of matrix cross interpolation [46], which constructs a low-rank approximation of a general matrix from a sparse set of its rows and columns. In TCI, this is applied iteratively bond by bond, as in the SVD case.

In what follows, we will see that smooth functions do not require large MPS bonds to be represented accurately. One further advantage of the TCI method is that it does not require explicit representation of the function being sampled; it can be provided as a black box. For a pedagogical introduction and technical details of the TCI algorithm, we refer the reader to Refs. [36,45].

III. ENTANGLEMENT IN TENSORIAL APPROXIMANTS OF SMOOTH FUNCTIONS

A. General results

Here, we summarize our theoretical results on the behavior of entanglement in tensor approximations of smooth functions. Particularly, we derive universal scaling relations for purities and entanglement spectra for the qubits describing the smallest scales. For the qubits describing the largest scales, outside the universal regime, we note the dependence of the entanglement spectra on function concentration. Taken together, these results enable valuable insights into the construction and precision of the quantum circuit encoding of the function. Proofs of the theorems and corollaries are provided in Appendix A.

Theorem 1. For a smooth (having infinitely many continuous derivatives) and normalized function $f(x)$ on the interval $[0, 1]$, the purities p_k of the tensorial representation $f_{\sigma_1\sigma_2\dots\sigma_N}$ decay exponentially with the bond number k [i.e., with decreasing spatial scale of function variations, via Eq. (1)]. More precisely, as $k \rightarrow \infty$ we have

$$p_k = 1 - g_1(f)/(6 \times 4^k) + O(1/8^k), \quad (13)$$

where

$$g_1(f) = \int_0^1 f'(u)(f^*)'(u)du - \int_0^1 f'(u)f^*(u)du \times \int_0^1 f(u)(f^*)'(u)du, \quad (14)$$

and where f' and f^* denote the derivative and complex conjugate of f , respectively.

In Appendix A, we show that $g_1(f)$ vanishes precisely for functions that are expressible as trivially factorizable MPS of bond dimension one. A key consequence of Theorem 1 is that a state encoding a smooth function approaches a factorized state with increasing k .

We further study the reduced density matrices obtained from a bipartition of the qubit system. Let us thus consider the reduced density matrix of the first k qubits ρ_k . We have the following.

Theorem 2. Under the same assumptions as in Theorem 1, to leading order in $1/2^k$, the eigenvalues of the reduced density matrix of the function encoding on the first k qubits scale as $\rho_{k,n} = O(1/4^{kn})$, where $n \in \mathbb{N}_0$ corresponds to n th eigenvalue and $k \rightarrow \infty$.

Further, from the above results on the scaling of purities and eigenvalues, we immediately obtain corollaries about the behavior of entanglement spectra and entanglement entropy. To wit, the relation between purities and entanglement spectra is given by Eq. (9). As the purity approaches unity, the leading coefficient $\Lambda_{k,0}$ dominates, and others vanish. More specifically as follows.

Corollary 1. For a smooth and normalized on $[0, 1]$ function $f(x)$, the subleading coefficient $\Lambda_{k,1}$ scales as

$$\Lambda_{k,1} \sim \frac{1}{2^k} \sqrt{\frac{g_1(f)}{12}} \quad (15)$$

as $k \rightarrow \infty$, where g_1 is defined in Theorem 1. The leading entanglement spectrum coefficient scales as $\Lambda_{k,0} \sim 1 - g_1(f)/(24 \times 4^k)$.

Corollary 2. The entanglement entropy S_k decays exponentially with the bond index k :

$$S_k = O\left(\frac{k}{4^k}\right) \text{ as } k \rightarrow \infty. \quad (16)$$

In particular, from Theorem 2 for the eigenvalue $\rho_{k,1}$ we obtain the asymptotic $\rho_{k,1} \sim g_1(f)/(12 \times 4^k)$, which agrees with Corollary 1. The next eigenvalue $\rho_{k,2}$ can also be relatively simply computed, which we do in Appendix B. We will also use this value in further numerical analysis. In Appendix B, we also provide several analytical examples to illustrate this scaling. Additionally, in Appendix A4 we observe that if the function is not smooth, but has absolutely continuous derivatives up to r th order, then our previous results hold for entanglement spectra up to order $r+1$, while the $r+2$ and higher orders are nonuniversal.

The above results are based on function smoothness and characterize the asymptotic behavior for qubits describing the smallest spatial scales of function's variations [as per Eq. (1)]. The entanglement spectra for the leading qubits, corresponding to the large scales, are nonuniversal, but we observe a strong dependence on the localization properties of the function. For example, a maximally localized delta-peaked distribution is described by an entirely unentangled product state. Similarly, as described in detail below, exponentially decaying functions are less entangled than the power-law-tailed ones.

Let us compare our results with those previously obtained in the literature. In Ref. [3], a bound for the purity was obtained in the following form: $p_k \geq 1 - 4 \max |f(x)f'(x)|/2^k$ for a positive f . In contrast, our analytic expansions for purities show that the first nonvanishing term scales $1/4^k$ for a general f , i.e., an exponential improvement in the bound. In Ref. [17], an Ansatz for entanglement spectra of the form $\sqrt{\rho_{kn}} = \Lambda_{k,n} = \alpha_k \exp(-\beta_k n)$ is used, and is fit to empirical data. While this captures the exponential dependence on n , and is thus consistent with our results, we emphasize that in our case, this scaling dependence is analytically derived, and not empirically postulated. Finally, Ref. [20] studied approximability of functions with matrix product states. Assuming extendibility to the complex plane, they bounded the maximal deviation between the function and its approximation with MPS with k th bond dimension truncated to χ by $O(2^{-k\chi})$. While Ref. [20] does not study the entanglement properties of the encoding, this result agrees with our predictions, as the maximal deviation should be of the same order as the maximal truncated entanglement spectra, which in this case is precisely of the order $O(2^{-k\chi})$. Moreover, our results allow to compute the exact coefficients in the obtained expansions and bounds, which we demonstrate for probability distributions.

B. Encoding probability distributions

While our results hold for a more general class of functions (e.g., complex valued), we now focus on the important problem of encoding probability distributions $p(x)$. This is special case of function encoding with $f(x) = \sqrt{p(x)}$, where the normalization of the probability corresponds to L_2 normalization of f . These functions, particularly modeling heavy-tailed distributions, have practical applications in areas such as portfolio optimization and derivative pricing [30,47] and risk management [29,30].

We consider the following examples of smooth distributions on the interval $[0, 1]$: the (truncated) normal, log-normal, and Lévy distributions:

$$p(x)_{\text{normal}} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (17)$$

$$p(x)_{\text{log-normal}} = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad (18)$$

$$p(x)_{\text{Levy}} = \sqrt{\frac{c}{2\pi}} \frac{\exp\left[-\frac{c}{2x}\right]}{x^{3/2}}. \quad (19)$$

We generally set location parameters μ to $1/2$ so that the distributions are mostly located in $[0, 1]$ interval, and vary the scale parameters σ or c . We note that an additional (often negligible) normalization is required due to truncation of the support to a finite interval. The functionals $g_1(f)$ from Eq. (14), as well as quantities related to the third eigenvalue $\Lambda_{k,2}$, used below, are computed explicitly in Appendix B.

In Fig. 2(a), we compare the subleading entanglement eigenvalue, which, via Eq. (15), is given by $\Lambda_{k,1} \sim \sqrt{g_1(f)}$, for the Lévy distribution truncated to the interval $[0, 1]$, to the values obtained from performing SVD on the input tensor $f_{\sigma_1\sigma_2\ldots\sigma_N}$ of the discretized function values. For large k , we find the precise agreement between the analytical and numerical results, as expected from Theorem 1 and Corollary 1, while

for the leading bonds, i.e., large scales, the behavior of entanglement is nonuniversal, as discussed before.

In Fig. 2(b), we compare the same property for the normal distribution (with $\sigma = 1/2$ and $\sigma = 1/8$). We obtain a nearly perfect match between the analytical prediction and numerics. The small discrepancy on the last bond is due to the truncation of the function's discretization to the finite scale $1/2^N$, while the discrepancy on the first bonds is caused by the subleading corrections in $1/2^k$.

Finally, in Figs. 2(c) and 2(d) we compute and compare the third entanglement eigenvalue $\Lambda_{k,2}$ for Lévy and normal distributions, respectively. As expected, our analytical result matches the numerical data, which indeed behaves as $O(1/4^k)$. As in the previous cases, a discrepancy with analytic results appears only on the first bonds corresponding to the large scales.

C. Encoding dependence on the localization and support of the function

Previously, we studied the entanglement on the smallest scales and showed it exhibits a universal exponential decay. While no such universality holds true for the largest scales (see Appendix C for details), we can nevertheless understand the entanglement at large scales better for a specific class of functions, namely, ones exhibiting *localization*. This class contains, in particular, the probability distributions we were concerned with in the preceding section.

For this discussion, it will be useful to bring back the explicit dependence on the support size. Consider a function $f(x)$ defined on $[0, \infty)$. Numerically, to represent the function its support will be truncated to a chosen finite interval $[0, L]$, the function then normalized in L_2 norm, and discretized at a fixed resolution $L/2^N$ (using N qubits). This is of course equivalent to encoding $\sqrt{L}f(Lx)$, with $x \in [0, 1]$. To iteratively increase the support size at a fixed resolution, we may simply double the interval to $[0, 2L]$, while simultaneously adding another qubit (adjusting the normalization, as required).

We now assume that $f(x)$ is localized, exponentially or polynomially, without loss of generality on the subinterval $[0, L]$ of $[0, \infty)$. By this, we mean that the mass is concentrated on a finite interval $[0, L]$ and is vanishing in the distribution tail. More formally, for all n the integral of $|f|^2$ on the subinterval $[0, 2^{n-1}L]$ is close to 1, while it is exponentially (resp. polynomially) small in the subinterval size on $[2^{n-1}L, 2^nL]$ as we increase n .

We can now examine the entanglement in the highest scale qubits, as we successively increase the support size at fixed resolution, using the doubling procedure described above. As discussed in more detail in Appendix C, the $|f|^2$ integrals over $[0, 2^{n-1}L]$ are in fact equal to the matrix elements of the reduced density matrix obtained by tracing out all but the highest scale qubit. The off-diagonal elements, and thus also the entanglement, can easily be estimated to be of order $\exp(-\lambda 2^{n-1}L)$ for some $\lambda > 0$ in the exponential localization case, and of order $(2^{n-1}L)^{-\alpha+1}$ in the polynomial case, where λ and α characterize the exponential/polynomial tail. Thus, as we add qubits to represent the increasing spatial scales (and increase the size of the support to which the function is

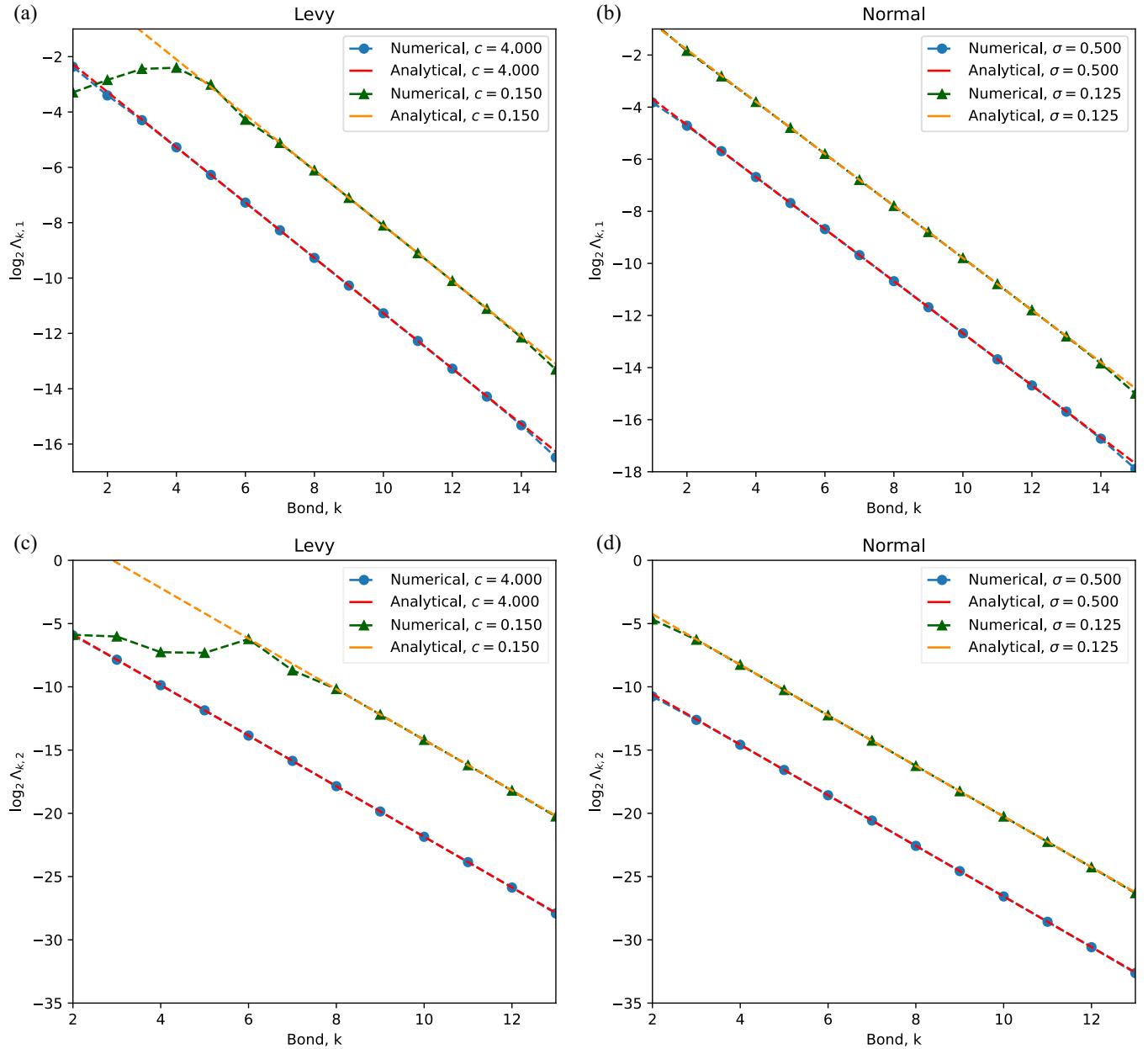


FIG. 2. Comparison of the analytical (see Theorem 1 and Corollary 1) and numerical SVD results on the dependence of the entanglement spectra $\Lambda_{k,1}$ and $\Lambda_{k,2}$ on bond position k . (a) $\Lambda_{k,1}$ for the Lévy distribution with $c = 0.15$ and $c = 4$. (b) $\Lambda_{k,1}$ for the normal distribution with $\mu = 0.5$ for $\sigma = 0.125$ and $\sigma = 0.5$. (c) $\Lambda_{k,2}$ for the Lévy distribution with $c = 0.15$ and $c = 4$. (d) $\Lambda_{k,2}$ for the normal distribution with $\mu = 0.5$ for $\sigma = 0.125$ and $\sigma = 0.5$. Note the excellent agreement and universal behavior for large k , and nonuniversal features for small k , as discussed in the text.

truncated), the entanglement of the leading qubits with those qubits representing lower scales decreases superexponentially in n for an exponentially localized function, and at least exponentially in n for a polynomially localized function. For large enough n , the qubits representing the large scales of the localized function will be almost fully disentangled from the qubits representing the lowest scales. This argument allows us to understand how large an interval is needed to faithfully represent a function.

The above argument is borne out by numerical experiments. In Figs. 3(a) and 3(b), we plot the logarithm of the second entanglement eigenvalue $\Lambda_{k,1}$ on bond k between

qubits k and $k + 1$, for the normal and Lévy distributions, respectively. The large- k tails are universal and exponentially decaying, as derived in Sec. III A. The low- k behavior, however, demonstrates a nonuniversal decay of entanglement toward large scales. Moreover, there is a clear distinction between the exponential localization of the normal distribution and the weaker polynomial one of the Lévy distribution. For the former, in Fig. 3(a), the large-scale (small k) drop of entanglement is almost instantaneous (superexponential), and also the universal asymptotic regime of exponential decay for lower scales (large k) is visible already for a support truncated to a small interval $[0, 2^4]$. Note that the maximum is

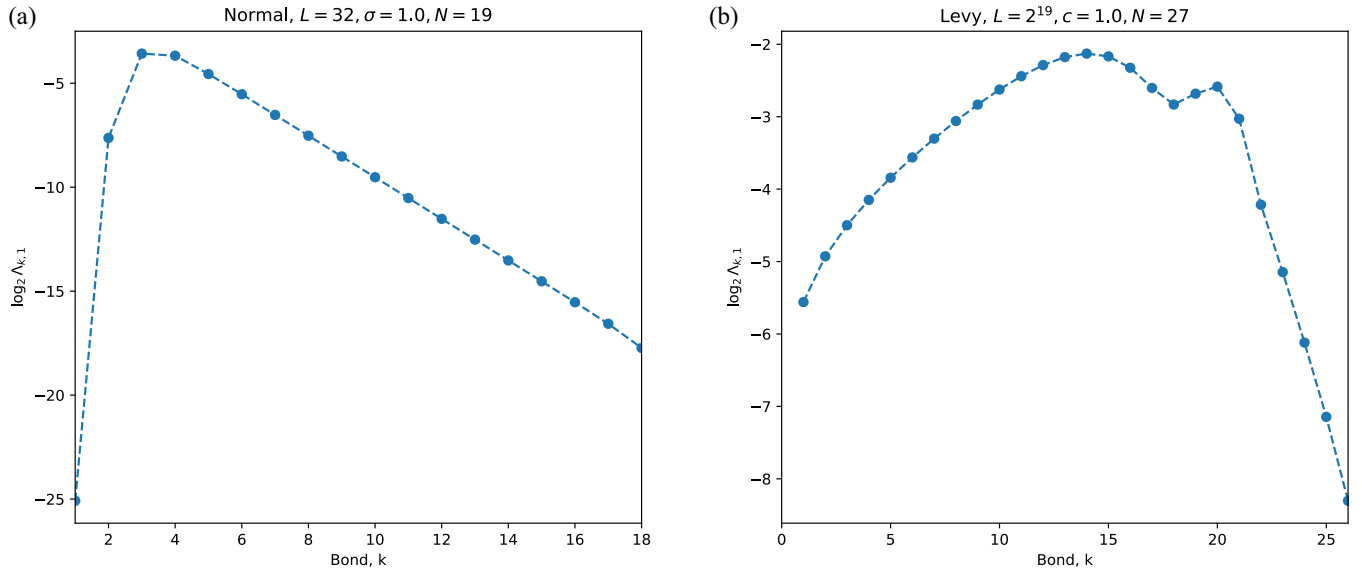


FIG. 3. Entanglement decay for exponential and polynomial localization: (a) The entanglement spectra $\Lambda_{k,1}$ as a function of the bond position k (lines are for guidance only) for the exponentially localized one-sided normal ($\mu = 0$, $\sigma = 1$) distribution truncated to the interval $[0, 32]$, for $N = 19$ qubits. Note the superexponential decay of entanglement for small k , and the universal, exponential, smoothness-controlled regime at large k . (b) The entanglement spectra $\Lambda_{k,1}$ as a function of the bond position k for the Lévy distribution (with $c = 1$ and interval $[0, 2^{19}]$), on $N = 27$ qubits. For small k , the entanglement decays superexponentially, but, due to only polynomial localization, slower than for normal distribution, delaying the onset of the universal regime at large k .

achieved between $k = 3, 4$, consistent with the condition $k \approx \log_2(L/2\sigma)$ derived in Sec. III A. For the Lévy case the small- k decay of entanglement is much slower (nearly exponential, as expected for polynomial localization), and correspondingly the onset of the universal regime of entanglement decay at small scales is pushed to much larger values of k : We required a support truncated to a very large interval $[0, 2^{19}]$ in Fig. 3(b) to clearly demonstrate both the low- and high- k regimes.

Both Figs. 3(a) and 3(b) demonstrate the characteristic triangular shape, which is also seen for other localized distributions (e.g., log-normal, Gamma, Cauchy). We emphasize that while the large- k behavior is universal, the small- k is not, and depends, among others, on the detailed localization properties of the function. The slow decay of entanglement for heavy-tailed distributions (e.g., all α -stable distributions except normal) has practical implications for the design of encoding circuits, as it indicates more qubits with significant entanglement will be needed, before the onset of the universal exponential decay regime controlled by smoothness.

IV. CONSTRUCTING THE ENCODING CIRCUIT USING THE MPS REPRESENTATION

We now assume that an MPS encoding of a smooth function is provided, and describe how to construct a shallow [using $O(N)$ gates] encoding quantum circuit taking advantage of the above results. For small qubit numbers (less than 30), a vector of discretized function values can be converted to MPS using the standard SVD method. For large scales, we employ the TCI algorithm [45], which is sampling based and *does not* require explicit storage of the full inputs in the classical memory. This TCI version is compute- and

memory-efficient, allowing us to generate utility-scale encoding circuits, which we executed on QPU for up to 156 qubits (see Figs. 10 and 11).

Our base algorithm extends the approach of Refs. [21–23,25,26], which we now briefly review. Recall first that an MPS of bond dimension $\chi = 2$ can be converted exactly, in a purely analytic procedure, into a quantum circuit utilizing only two-qubit gates [9,10]. Though for larger bond dimensions an exact procedure requires arbitrarily large multiqubit gates and is thus infeasible, an approximate procedure is constructed as follows. (1) The input MPS state is *truncated* to bond dimension $\chi = 2$. The truncation can be done by, e.g., setting to zero all but the two largest elements of the diagonal Λ matrices in the canonical form representation of MPS [see Eq. (5)]. (2) The MPS with bond dimension $\chi = 2$ can be converted exactly into a quantum circuit. (3) The Hermitian conjugate of this quantum circuit is contracted with the original MPS input state (at full bond dimension) on an MPS quantum circuit simulator (we used the Qiskit MPS simulator [48]) *disentangling* it. (4) The procedure is iterated with the updated MPS state. Note that this approach is approximate and greedy. As MPS of bond dimension 2 is represented exactly by a quantum circuit, contracting it with its Hermitian conjugate yields a product state, removing all entanglement. For MPS of higher bond dimension, contracting Hermitian conjugate of the circuit obtained from a truncated representation removes the leading part of the state, and yields a state of lower entanglement, requiring a lower bond dimension for the next iteration. In practice, this method generally converges to the exact circuit for state preparation for a large number of iterations.

We improve upon the above procedure in two distinct ways. First, motivated by our theoretical results, we modify it to take advantage of the structure of the entan-

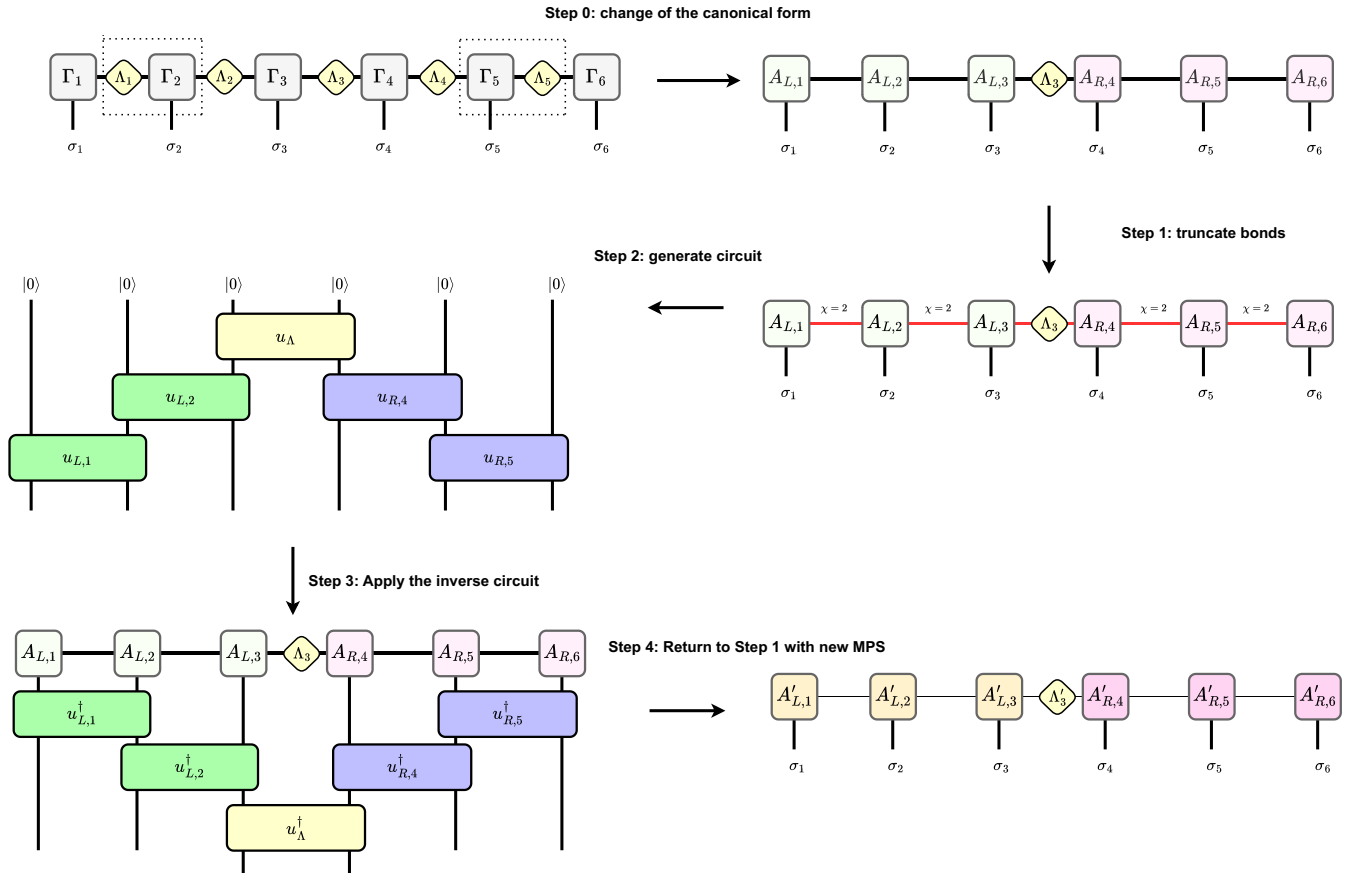


FIG. 4. An overview of our MPS-based-circuit construction algorithm. First, the standard canonical form MPS is transformed into the central canonical form (Step 0). All of the bonds are then truncated to bond dimension $\chi = 2$, shown in red (Step 1). The truncated MPS is decomposed exactly into a unitary circuit using V layers (see the main text), which may also be incomplete. The central gate may be variationally optimized (Step 2). On an MPS simulator, the inverse of the circuit is appended to the original nontruncated MPS to serve as a disentangler (Step 3). A new MPS is obtained by contracting the tensor network, which is now less entangled. The procedure is repeated for the new state.

gument in the input MPS. Second, to further improve the practical performance we note that specific operators in the procedure admit more efficient gate decompositions than those used in Refs. [21–23,25,26], resulting in overall shallower circuits.

For the former, note that the exact procedure of Refs. [9,10], used also in all subsequent works, exclusively produces complete linear staircase layers as disentangling circuits. Each layer, multiple of which are generated in disentangling iterations, uses $N - 1$ gates on N qubits, and has a depth also of $N - 1$. While this already produces a circuit with $O(N)$ gates, we show that an improved “double-staircase” or “V”-layer can also be generated (see Fig. 4), which has two advantages: First, the layer depth can be as low as $N/2 - 1$ for a centrally placed origin, which is practically a very significant advantage on real noisy QPUs. Second, since, as we demonstrated, the entanglement is *not* spatially uniform across the bonds, the greedy iterative disentangling procedure may in practice produce a better approximation when equipped with a freedom to move the origin of disentangling layers, even in the absence of noise. This is indeed the case, as we demonstrate numerically in Fig. 5, where up to three layers are used to approximate the Lévy distribution, and the error

of approximation is plotted as a function of the layer origin. Note that the gain due to shifting the origin can be as large as adding another disentangling layer in the standard staircase scheme, thus drastically reducing the number of gates and the circuit depth.

V layers are generated as follows: First, a central bond is chosen and the MPS is converted to the central canonical form with respect to this bond, so that the tensors to the left and to the right of the central bond are isometries [see Eq. (10) and Step 0 in Fig. 4]. The central bond can be chosen either to minimize the depth of the resulting V layer, or it can be placed in the region with largest bond entanglement. The MPS is then truncated to bond dimension 2. For this truncation, we can either briefly restore the full canonical form and set all but the two largest diagonal elements in all Λ matrices to zero, or we can truncate directly in the center canonical form by moving the position of central bond and sequentially truncating the Λ matrix of the bond (the latter method is described in detail in Ref. [7]). In either case, the resulting left and right tensors A_L and A_R are isometries, as defined in Eqs. (11) and (12), of dimension $2 \times \chi = 4$ by $\chi = 2$. The isometries can be completed into unitary matrices of dimension 4×4 , and interpreted as explicit matrix representations of two-qubit

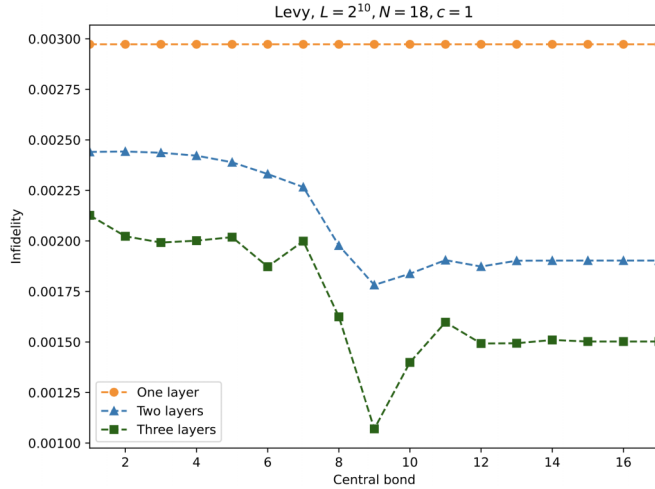


FIG. 5. The dependence of the infidelity on the number and position of the origin of the V layers in the encoding quantum circuits on an ideal simulator (lines between points for guidance only). Here, we show data for the Lévy distribution with $c = 1$ on support truncated to $L = 2^{10}$ using $N = 18$ qubits. The accuracy increases with additional disentangling layers (but see Fig. 12 for the effects of noise), and significantly depends on the origin of the V layers. The effect of the latter, which also reduces to overall circuit depth, has effect comparable to adding a new disentangling layer.

gates (as in Ref. [9]). Completion here means that the resulting two-qubit gates reduce to the original isometries if one of the qubits on which gate acts is in the $|0\rangle$ state. This is shown in Step 2 in Fig. 4, where the unitary matrix $u_{L,1}$ is obtained from the completion of the contraction of tensors $A_{L,1}$ and $A_{L,2}$, $u_{L,2}$ is found as a completion of $A_{L,3}$, $u_{R,4}$ is the completion of $A_{R,4}$, and $u_{R,5}$ is the completion of contraction of $A_{R,5}$ and $A_{R,6}$.

It is also necessary to prepare the initial state $\sum_i \Lambda_i |ii\rangle$ on the central bond (the gate u_Λ in Fig. 4). This requires a distinct approach. In the first layer, u_Λ acts on the $|00\rangle$ state and can be constructed using a single cnot [49]. In subsequent layers, we make this (and only this) gate variational. As for general two-qubit isometries, only up to two cnot gates are needed [49], we take such an ansatz with one-qubit parametrized rotations. Their parameters are fixed by minimizing the second Renyi entropy on the central bond of the contracted disentangled MPS [50].

The above procedure generates “full” disentangling V layers with $N - 1$ gates each. This is, however, not necessary as we have shown that the entanglement decays exponentially fast with bond numbers. To take advantage of this fact, we introduce a hyperparameter $\varepsilon_{\text{trunc}} > 0$, on whose value we condition the addition of gates on individual bonds in the V layer. Namely, for each bond k we compute $\sum_{i \geq 1} \Lambda_{k,i}^2$ in the contracted MPS. If this is smaller than the chosen small $\varepsilon_{\text{trunc}}$, then the state is essentially already disentangled, and the bond can be truncated to $\chi = 1$, i.e., no new gate is needed. In this way, in subsequent iterations we generate “incomplete” and diminishing V layers, as depicted in Fig. 1(d), resulting in overall fewer gates.

We have also modified the gate synthesis procedure for the isometric tensors. In the standard procedure, the two-qubit isometry is extended into a full two-qubit unitary matrix, from

which the gate representation is synthesized. Such a general unitary synthesis requires, however, three cnot gates and is not optimal. In fact, to synthesize a general isometry only two cnots are needed and one-qubit ry and rz gates [49], which we use, further reducing the overall and two-qubit gate count. For real-valued functions, one may even further simplify the circuits by only using the ry rotations [51].

V. NUMERICAL TESTS OF THE ENCODING CIRCUITS ON SIMULATORS AND QPUS

Here, we numerically study the quantum circuits constructed with our MPS-based procedure both on ideal and noisy simulators and on real IBM QPUs.

To measure the accuracy of the prepared quantum circuit, we use the fidelity defined as $F = |\langle \psi | f \rangle|^2$, where $|f\rangle$ is the exact state and $|\psi\rangle$ is the state prepared by the circuit. For a single-layer disentangling circuit, we can analytically predict the fidelity of the approximation using the truncated singular values. Truncating a single MPS bond k to $\chi = 2$ results in fidelity $1 - \sum_{i \geq 2} \Lambda_{k,i}^2$, where $\Lambda_{k,i}$ are the entanglement spectra/singular values on the k th bond. With multiple bonds truncated, to lowest order the fidelity is

$$F \approx 1 - \sum_{k > 0, i \geq 2} \Lambda_{k,i}^2, \quad (20)$$

which can be effectively computed up to an exact leading term via Theorem 2 (see Appendix B for details). The multilayer decomposition does not easily admit such analysis, and we only show numerical results.

In Fig. 6(a), we show the infidelity of a one-layer approximating quantum circuit as a function of the truncated support size L , assuming a fixed discretization size (thus the qubit number is $N = 10 + \lceil \log_2 L \rceil$). Here, we consider the example of a normal distribution with $\sigma = 2$. The infidelity is computed on an MPS simulator with a very large bond dimension (i.e., essentially exact simulation). Overall, the infidelity is excellent already for this extremely shallow circuit, no higher than 10^{-3} . Note also that for smaller L it perfectly follows the analytical estimate from Eq. (20). For larger supports, this simple approximation is not accurate: As discussed in Sec. III A, the smoothness-controlled regime is valid from bond $k = \log_2(L/2\sigma)$ upward, and below the entanglement is localization controlled [and, for this exponentially localized function, much smaller; see Fig. 3(a)]. For large L , thus, Eq. (20) overestimates the infidelity, which can be corrected by only summing contributions from bonds $k \geq \log_2(L/2\sigma)$, i.e., the ones for which the asymptotic scaling is valid. This corrected estimate agrees perfectly with the numerical results for all L .

In Fig. 6(b), we show the numerical infidelity of the one-layer quantum circuit for the Lévy distribution as a function of the truncated support L . We used a fixed discretization step ($N = 10 + \lceil \log_2 L \rceil$ qubits). Note again the very low infidelity achieved already with a single layer of gates, saturating at 4×10^{-3} . While for this case the analytical estimates based on Eq. (20) are less reliable, Lévy distribution being only weakly localized, the qualitative behavior of infidelity saturation at large L is expected: The nonuniversal contributions from a (finite) set of larger scales giving place to

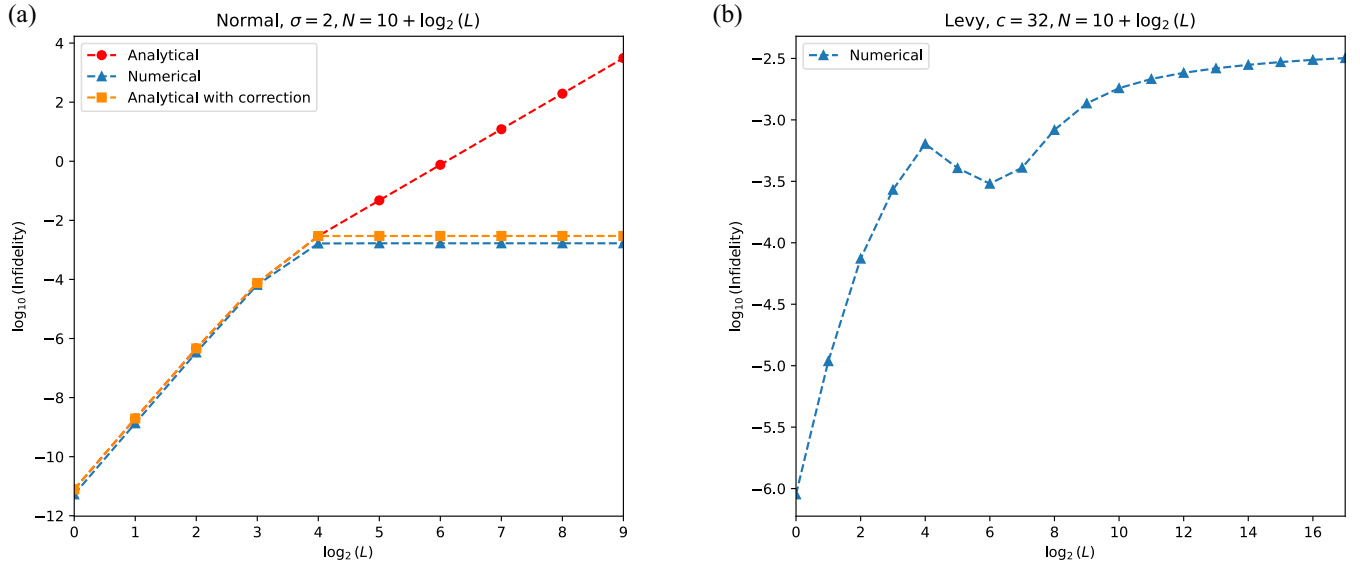


FIG. 6. (a) The log-log plot of infidelity of a one-layer quantum circuit as a function of the size of the truncated support interval L . We fix the discretization step size, with the qubit number given by $N = 10 + \lceil \log_2 L \rceil$. Here, we consider the truncated normal distribution with $\sigma = 2$, show the numerically computed infidelity, and compare it to the analytical estimates. The label “Analytical” refers to Eq. (20) and “Analytical with correction” to prediction taking into account localization properties (see the main text). Note the very low infidelity obtained already for this shallow circuit. (b) The numerical infidelity of the one-layer quantum circuit for the truncated Lévy distribution with parameter $c = 2^5$ as a function of the support interval L . The discretization step size is fixed, with the qubit number given by $N = 10 + \lceil \log_2 L \rceil$.

smoothness-controlled regime at larger scales where the entanglement systematically vanishes.

In Fig. 5, we examine the influence of using multiple disentangling V layers in the quantum circuit construction, and the sensitivity to the V layer origin position, on the example of the Lévy distribution. We observe that the fidelity can be systematically improved by adding additional disentangling layers (note, though, that in the presence of noise this is not true, as the increased circuit depth may result in a lower fidelity

in practice). Furthermore, for more than a single layer we also see a strong dependence on the layer origin, as expected from the entanglement analysis in Sec. III A: The effect is comparable in size to adding another disentangling layer, and hence the freedom to choose it is very important in practice. Note also that the minimal circuit depth is achieved with the origin on the middle qubits, further emphasizing the usefulness of this strategy.

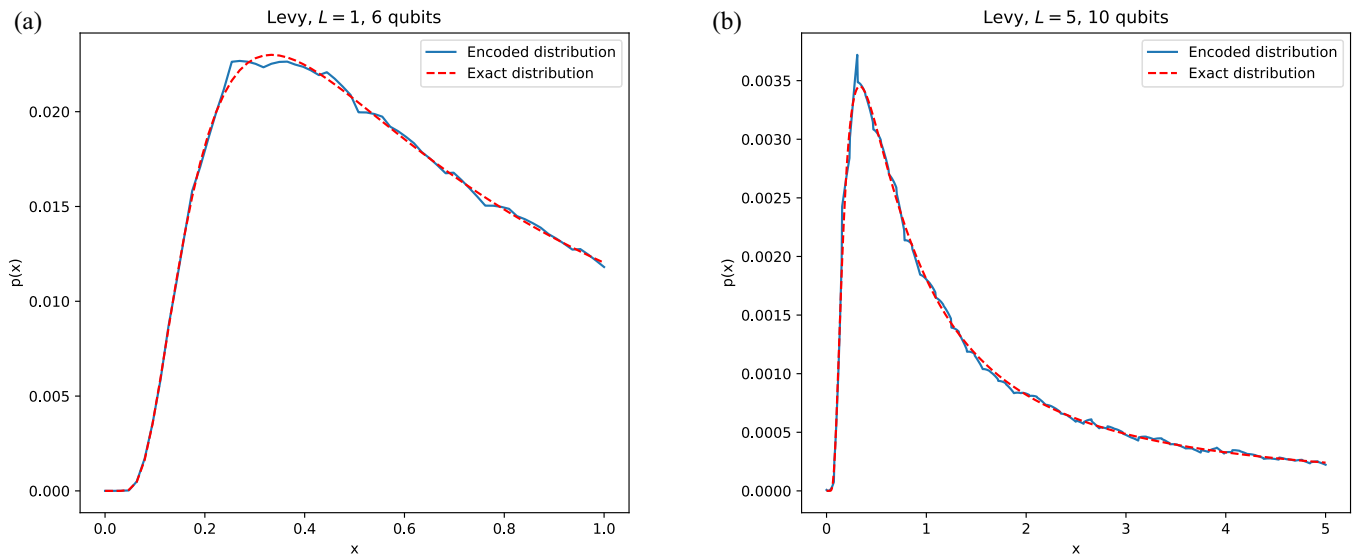


FIG. 7. The encoded Lévy distribution ($c = 1$) for different truncated support and discretization (number of qubits) sizes on an ideal noiseless simulator. Amplitudes of the quantum state created using a two-layer encoding circuit are compared to the exact distribution. In panel (a), for $N = 6$ qubits and $L = 1$; in panel (b), for $N = 10$ qubits and $L = 5$.

TABLE I. Statistical tests of the encoded distribution on an ideal simulator and the encoding circuit's (two-layer) properties with a $\varepsilon_{\text{trunc}} = 10^{-3}$ cutoff.

	Qubits	KL divergence (mean $\pm$ std)	KS test p value (mean $\pm$ std)	Depth (mean)	CNOTS (mean)
Normal	6	0.000 31 $\pm$ 0.000 48	0.238 39 $\pm$ 0.234 89	17.2	9.8
	10	0.000 45 $\pm$ 0.000 69	0.492 84 $\pm$ 0.289 43	26.2	17.8
	15	0.000 46 $\pm$ 0.000 73	0.504 86 $\pm$ 0.282 03	41.2	28.0
Log-normal	6	0.000 97 $\pm$ 0.000 75	0.089 86 $\pm$ 0.169 66	15.6	9.0
	10	0.001 48 $\pm$ 0.001 35	0.319 05 $\pm$ 0.323 49	28.4	18.2
	15	0.001 15 $\pm$ 0.001 12	0.473 37 $\pm$ 0.285 52	40.2	27.0
Lévy	6	0.001 75 $\pm$ 0.000 86	0.266 24 $\pm$ 0.223 99	15.6	9.0
	10	0.001 72 $\pm$ 0.000 53	0.500 47 $\pm$ 0.279 75	25.9	17.0
	15	0.001 49 $\pm$ 0.000 48	0.503 45 $\pm$ 0.284 11	41.0	27.0

We show the encoded distributions themselves in Fig. 7, where the amplitudes of the state created with a two-layer disentangling circuit on an ideal simulator are compared to the exact distributions for different support intervals L and discretizations (specified by the number of qubits N). We observe

a very good agreement, and emphasize that this is achieved using only linearly deep circuits, which is crucial when considering execution on realistic noisy devices. A more rigorous test of performance is obtained by computing statistical metrics for the samples drawn from the encoded distributions.

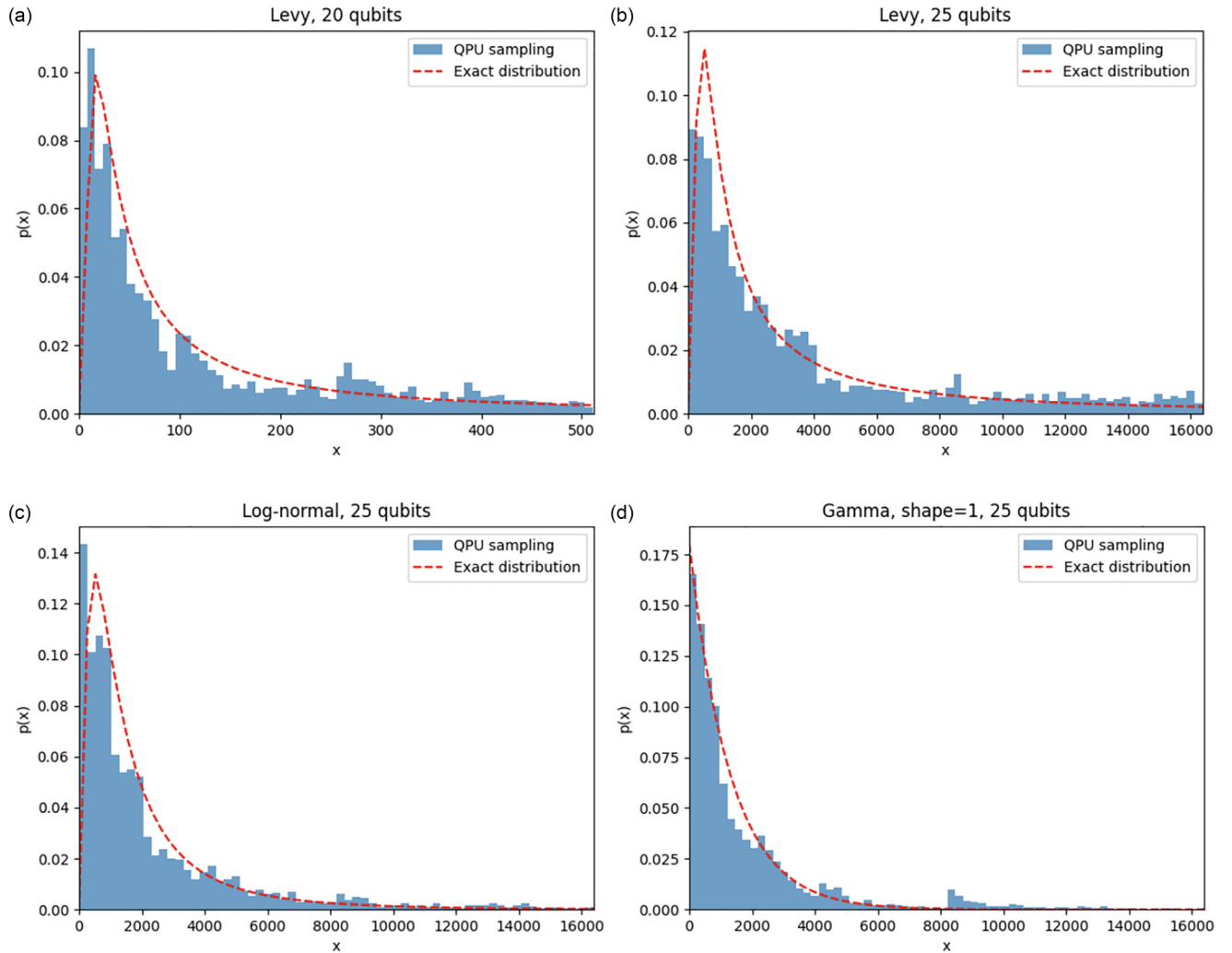


FIG. 8. Encoding circuits execution on the `ibm_torino` QPU. (a) Lévy distribution with $c = 50$ truncated to $L = 2^9$ with $N = 20$ qubits. (b) Lévy distribution with $c = 1310$ on $L = 2^{14}$ with $N = 25$ qubits. (c) Log-normal distribution with $N = 25$ qubits. (d) Gamma distribution with shape parameter $k = 1$ on $N = 25$ qubits. In each case, 5000 shots were used. All executions pass the KS test with 200 samples. The better of the one- and two-layer circuits used is shown.

TABLE II. Statistical tests of the encoded distribution on an ideal simulator and the encoding circuit's (two-layer) properties with a $\varepsilon_{\text{trunc}} = 10^{-4}$ cutoff.

	Qubits	KL divergence (mean $\pm$ std)	KS test p value (mean $\pm$ std)	Depth (mean)	CNOTS (mean)
Normal	6	0.000 23 $\pm$ 0.000 39	0.244 17 $\pm$ 0.243 45	18.2	10.4
	10	0.000 38 $\pm$ 0.000 63	0.490 01 $\pm$ 0.285 71	27.5	18.7
	15	0.000 37 $\pm$ 0.000 61	0.505 92 $\pm$ 0.282 07	41.6	29.2
Log-normal	6	0.000 76 $\pm$ 0.000 64	0.096 07 $\pm$ 0.177 82	19.5	12.1
	10	0.001 27 $\pm$ 0.001 20	0.326 50 $\pm$ 0.327 61	31.1	21.8
	15	0.001 06 $\pm$ 0.001 08	0.476 56 $\pm$ 0.290 33	43.8	32.7
Lévy	6	0.001 01 $\pm$ 0.000 97	0.263 29 $\pm$ 0.223 42	20.6	11.8
	10	0.001 12 $\pm$ 0.000 73	0.496 81 $\pm$ 0.278 33	30.5	20.8
	15	0.000 99 $\pm$ 0.000 64	0.492 13 $\pm$ 0.287 45	42.8	31.8

We use the Kullback-Leibler (KL) divergence between the encoded and ideal distributions and the Kolmogorov-Smirnov (KS) test. The results are reported in Tables I and II. For these tests, we use a two-layer encoding circuit constructed by our complete algorithm as described in Sec. IV, including truncation of gates in V layers set by the parameter $\varepsilon_{\text{trunc}}$. We report the depth and amount of cnot gates of the circuit, which are computed in the RX, RY, RZ, and CNOT basis (the depth includes all one-qubit gates). To collect the data, we varied the interval L and scale factors of the distributions, and we report the mean value and the standard deviation. The constructed distributions mostly pass the statistical tests with a wide margin, but we note that for very small amount of qubits KS test may occasionally fail (p value < 0.05) due to extreme discretization of the distribution, as the test is performed against the ideal truncated cumulative density function, which is continuous. In Table I, we set the hyperparameter $\varepsilon_{\text{trunc}}$ to 10^{-3} , while in Table II to $\varepsilon_{\text{trunc}}$ to 10^{-4} . Changing this value represents an accuracy-depth trade-off, which will be important on noisy QPUs.

To emphasize the above mentioned depth-accuracy trade-off induced by realistic noise, we perform noisy simulations on the IBM FakeToronto backend. We compare the results of the execution of a one-layer and five-layer encoding circuits

in Fig. 12. While in the absence of noise the latter is more accurate, in this more realistic setting the distribution features are largely destroyed by the noise, while the nominally less accurate one-layer encoding performs much better. We leave detailed examination of this trade-off to future work.

Finally, we put our encoding circuits to test on real QPUs. Motivated by the observed theoretical accuracy-depth trade-off, we executed shallow one- and two-layer circuits generated by our procedure ($\varepsilon_{\text{trunc}} = 10^{-4}$) on the *ibm_torino* superconducting device. In Fig. 8, we show the distributions sampled from the better of the two encoding circuits for the Lévy, log-normal, and Gamma distributions for up to $N = 25$ qubits. For each plot, we used 5000 shots. All these experiments pass the KS test at the standard significance level 0.05 (with 200 samples).

Furthermore, using the TCI approach, we also generate large-scale circuits, for up to 156 qubits, which we execute on IBM devices. First, we validate the TCI function loading itself. Toward this goal, we generated a 40-qubit encoding circuit for the Lévy distribution and interval size $L = 10^9$, using the TCI algorithm of Ref. [45] and *torchTT* package [52] to generate the MPS representation directly from the PDF. Specifically, we build a map $[0, 1] \mapsto \sqrt{L \cdot p_{\text{Levy}}(xL)/R}$, where $\sqrt{L}$ accounts for the renormalization after changing the

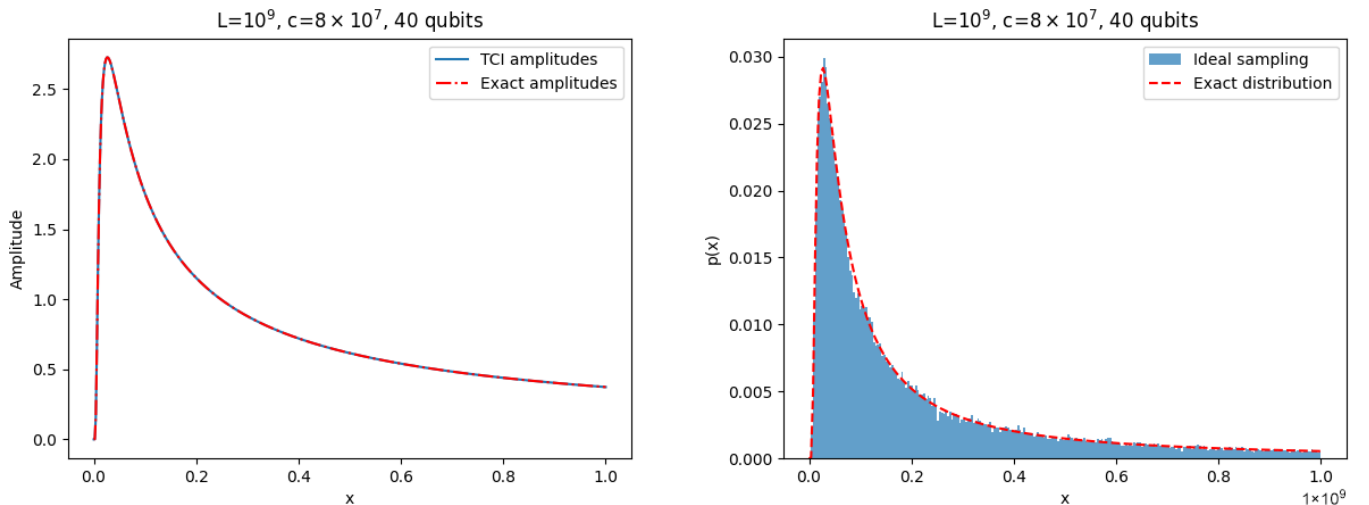


FIG. 9. (Left) The amplitudes of the quantum state encoding the Lévy distribution on 40 qubits obtained using TCI, compared to the exact MPS amplitudes (right). The results of sampling from the 40-qubit quantum circuit constructed using TCI approach on an ideal simulator.

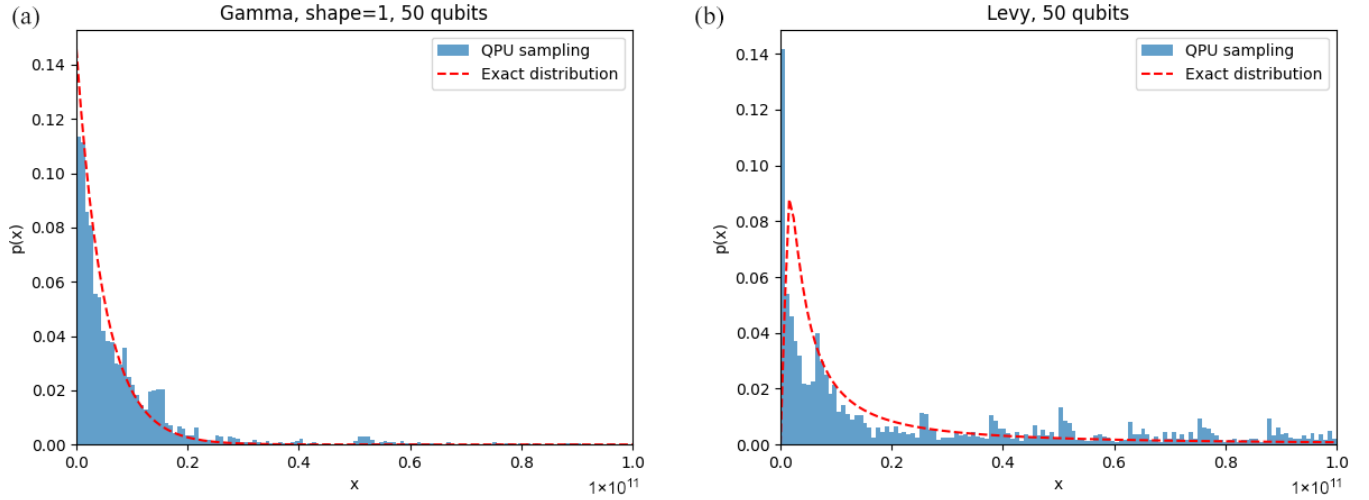


FIG. 10. TCI constructed two-layer encoding circuits for the (a) Gamma and (b) Lévy distributions on 50 qubits, executed on the *ibm_marrakesh* QPU. We used in both cases, $L = 10^{11}$, the scale factor $c = 5 \times 10^9$, and 5000 shots on the QPU. While these very large circuits no longer pass the quantitative statistical tests, the qualitative features of the distributions are still reproduced. Compare, however, the results on higher-quality QPU in Fig. 11.

variables, and a constant factor R accounts for a truncation of the distribution to a finite interval and can be factored out of the cumulative distribution function (CDF). In Fig. 9, we show the encoded state and sampling results from a 40-qubit encoding quantum circuit executed on an ideal simulator. The large scaling factor c is chosen so that the distribution tail does not vanish in the truncated range. Numerically, we confirm that the relative error between the true PDF and the TCI uploaded value in the MPS is 6×10^{-8} on average and 4×10^{-7} at most. We note that 8×10^4 calls to the PDF map sufficed to build the state to this accuracy.

Next, in Fig. 10 we show the results of executing the TCI-generated two-layer quantum circuits for the Gamma and Lévy distributions at 50 qubits, executed on the *ibm_marrakesh* QPU with 5000 shots. While these large-scale distributions no longer pass the statistical tests due

to device noise, the qualitative features of the distributions are reproduced. To showcase how the improvement in device quality directly impacts the sampling results, in Fig. 11 we show the results on the *ibm_kingston* QPU. While this is also a Heron r2-type device, its two-qubit gate fidelities were substantially better ($1.88\text{e-}3$ vs $2.67\text{e-}3$ for *ibm_marrakesh*). This allowed to sample from Gamma and Lévy distributions on up to 156 qubits, which we performed using 2000 measurement shots.

VI. CONCLUSIONS

We studied entanglement properties of MPS representations of smooth functions, and their applications to constructing shallow encoding quantum circuits. In particular, we derived the universal properties of their entanglement spectra.

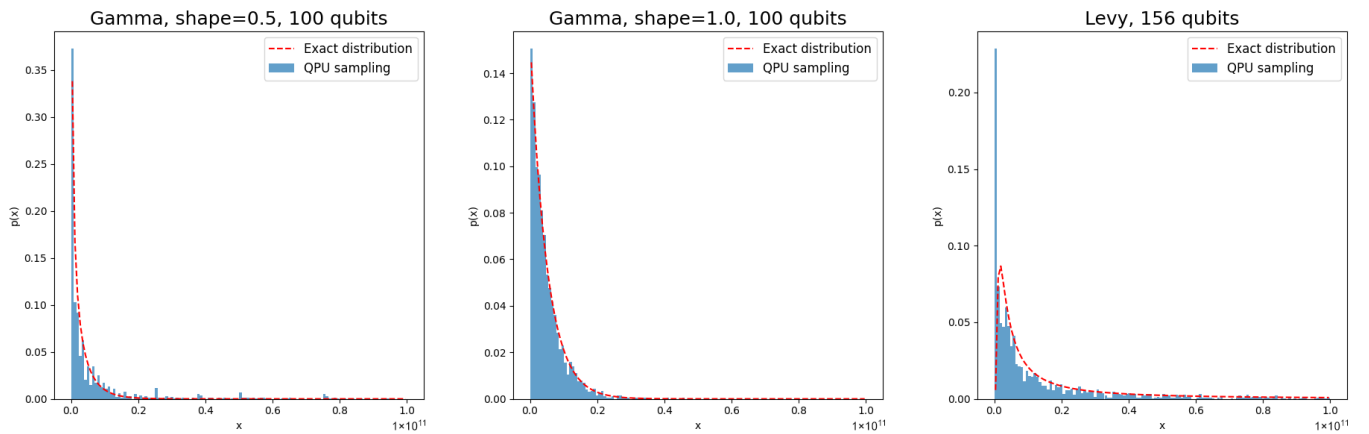


FIG. 11. Sampling from two-layer encoding circuits on the *ibm_kingston* QPU. We used TCI in circuit generation, and measured 2000 shots on the QPU; we also used dynamical decoupling. In all plots, we used $L = 10^{11}$ and the scale factor $c = 5 \times 10^9$. (Left) Gamma distribution with shape parameter $a = 1$ sampled on 100 qubits. (Middle) Gamma distribution with shape parameter $a = 0.5$ sampled on 100 qubits. (Right) Lévy distribution sampled on 156 qubits. Note that while both *ibm_marrakesh* and *ibm_kingston* are QPUs of type Heron r2, the latter has a significantly lower two-qubit gate error ($2.67\text{e-}3$ vs $1.88\text{e-}3$ median error), which is reflected in the much higher quality of results compared to Fig. 10.

We used these to estimate the errors of the quantum circuit representations of these functions, and to build an improved MPS-based algorithm generating shallow quantum circuits [$O(N)$ gates] encoding them. The resulting circuits were benchmarked on real and simulated IBM QPUs. The practical performance of our algorithm was validated experimentally on the important example of preparing states encoding PDF, in particular, heavy-tailed ones. Both the analytical and algorithmic results apply, however, to general smooth functions. One further domain where this may be of importance is the preparation of initial states for quantum simulation of PDE, used, e.g., in quantum computational fluid dynamics [15].

For the important case of Lévy distributions, the quantum circuits we constructed reproduced well their qualitative behavior for up to 156 qubits. Moreover, for up to 25 qubits the accuracy was sufficient to pass quantitative statistical tests even on the current noisy devices. The results on distribution loading may have importance to various domains of financial risk analysis, risk management, and decision-making that include series of financial data. By developing methods to efficiently work with Lévy distributions on a quantum computer, we pave the way for more precise modeling of market behaviors, particularly in capturing heavy tails, skewness, and volatility clustering.

It is of key importance that our procedure does not require the explicit storage of the input, i.e., the potentially exponentially many discretized function values, in memory. Instead, it is able to construct the circuit directly, by locally sampling from the PDF through the use of TCI [36]. Thus, our method is both classical compute- and memory-efficient, and scalable to utility sizes.

Several open research directions remain. First, the density matrix-based analysis can be extended to multivariate functions. It is particularly interesting to derive the optimal tensor network representation of such functions based on their entanglement (for multivariate functions, there are many possible tensor network geometries, which may represent the function; from the point of quantum computing, these geometries correspond to different qubit orderings, or qubits arranged in topologies more complicated than linear). Analytic results on entanglement of multivariate functions may lead both to better quantum circuit encodings of such functions and also may improve existing tensor network algorithms. Finally, from a practical standpoint it would also be very important to carry out a more detailed analysis of the influence of hardware noise on the function preparation.

ACKNOWLEDGMENTS

We thank Vladyslav Los and Oleksandra Hutor for their technical support in executing the quantum circuits on the IBM quantum hardware.

This paper was prepared for information purposes and is not a product of HSBC Bank Plc. or its affiliates. Neither HSBC Bank Plc. nor any of its affiliates make any explicit or implied representation or warranty and none of them accept any liability in connection with this paper, including, but not limited to, the completeness, accuracy, reliability of information contained herein and the potential legal, compliance, tax or accounting effects thereof. Copyright HSBC Group 2024.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: PROOFS OF THEOREMS

1. Proof of Theorem 1

To obtain the purities, we first construct the reduced density matrix ρ_k in terms of the tensor representation:

$$\rho_k = \frac{1}{2^N} \sum_{\mathbf{v}_k} f_{\mathbf{u}_k, \mathbf{v}_k} f_{\mathbf{u}'_k, \mathbf{v}_k}^*, \quad (\text{A1})$$

where we introduced collective indices $\mathbf{u}_k = \sigma_1 \sigma_2 \dots \sigma_k$ and $\mathbf{v}_k = \sigma_{k+1} \sigma_{k+2} \dots \sigma_N$ for the bipartition of the system (across spatial scales). Here, f^* denotes a complex conjugate of f . Then, the purity is given by

$$p_k = \frac{1}{4^N} \sum_{\mathbf{u}_k, \mathbf{u}'_k, \mathbf{v}_k, \mathbf{v}'_k} f_{\mathbf{u}_k, \mathbf{v}_k} f_{\mathbf{u}'_k, \mathbf{v}'_k}^* f_{\mathbf{u}_k, \mathbf{v}_k}^* f_{\mathbf{u}'_k, \mathbf{v}'_k}. \quad (\text{A2})$$

We further write the value $f_{\mathbf{u}_k, \mathbf{v}_k}$ as $f(\mathbf{u}_k + \mathbf{v}_k/2^k)$, where by a slight abuse of the notation, in the argument of the function f , we have identified the index $\mathbf{u}_k$ with the real number $0.\mathbf{u}_k \equiv 0.\sigma_1 \dots \sigma_k$ and $\mathbf{v}_k$ with $0.\mathbf{v}_k \equiv 0.\sigma_{k+1} \dots \sigma_N$. In the limit $N \rightarrow \infty$, the sums over $\mathbf{v}_k, \mathbf{v}'_k$ may be rewritten as integrals over continuous v, v' :

$$p_k = \frac{1}{4^k} \sum_{\mathbf{u}_k, \mathbf{u}'_k} \int_0^1 \int_0^1 [f(\mathbf{u}_k + v/2^k) f^*(\mathbf{u}_k + v'/2^k) \times f(\mathbf{u}'_k + v'/2^k) f^*(\mathbf{u}'_k + v/2^k)] dv dv'. \quad (\text{A3})$$

The sums over $\mathbf{u}_k, \mathbf{u}'_k$ cannot be equally converted to integrals by taking the limit $k \rightarrow \infty$, as we want the purities at finite k . To this end, we use the Euler-MacLaurin resummation formulas, which approximate the finite $1/2^k$ resolution sum with continuous integrals augmented with additional boundary terms. In our specific case, we use

$$\sum_{u_n} \frac{f(u_n)}{2^k} = \int_0^1 f(u) du + \frac{f(0) - f(1)}{2^{k+1}} + B_2 \frac{f'(1) - f'(0)}{2^{2k+1}} + O(1/2^{3k}), \quad (\text{A4})$$

where B_2 is the second Bernoulli number. Then, the purity is given by

$$p_k = \int_0^1 \int_0^1 \left(\int_0^1 f\left(u + \frac{v}{2^k}\right) f^*\left(u + \frac{v'}{2^k}\right) du + b_1(v, v') \right) \times \left(\int_0^1 f^*\left(u' + \frac{v}{2^k}\right) f\left(u' + \frac{v'}{2^k}\right) du' + b_2(v, v') \right) dv dv', \quad (\text{A5})$$

where b_1, b_2 are the boundary terms. There is a small subtlety here, as the function here is effectively extended beyond $[0, 1]$ to the interval $[0, 1 + 1/2^k]$. We assume that such an

extension exists and is smooth. Indeed, any smooth function on $[0, 1]$ can be easily extended. Moreover, on practice it is likely that the function is already defined beyond this interval, which is used for encoding.

To evaluate the v, v' integrals, for large k we can expand in the small parameter $v/2^k$:

$$f(u + v/2^k) = f(u) + f'(u)\frac{v}{2^k} + f''(u)\frac{v^2}{2^{2k+1}} + O(1/2^{3k}). \quad (\text{A6})$$

This effectively leads to an expansion in $1/2^k$ for the purities p_k , which we will now show. First, the leading term in p_k is equal to one, where we use the fact that boundary terms come at order at least $1/2^k$ and the function is normalized:

$$\left(\int_0^1 f(u)f^*(u)du\right)\left(\int_0^1 f^*(u')f(u')du'\right) = 1. \quad (\text{A7})$$

The contributions to the p_k expansion at order $1/2^k$ come from the corresponding terms in the expansion [Eq. (A6)] and the boundary terms. Introducing the notation

$$h_{n,m} = \int_0^1 f^{(n)}(u)f^{*(m)}(u)du. \quad (\text{A8})$$

We now can see that the $1/2^k$ terms in the p_k expansion vanish:

$$h_{0,1} + h_{1,0} + |f(0)|^2 - |f(1)|^2 = 0, \quad (\text{A9})$$

since $h_{0,1} + h_{1,0}$ is the integral of a total derivative from the integration by parts formula, which cancels the boundary terms.

Finally, we evaluate the contributions to p_k at $1/4^k$, which come from gathering the terms in the expansion of the du and du' integrals in Eq. (A5) to second order. From a product of the zeroth-order term in u and second-order terms in u' , we obtain

$$\begin{aligned} & (h_{2,0} + h_{0,2})/6 + h_{1,1}/4 - (|f(1)|^2)' - |f^2(0)|'/6 \\ &= (h_{2,0} + 2h_{1,1} + h_{0,2})/6 - h_{1,1}/12 - (|f(1)|^2)' \\ & \quad - |f^2(0)|'/6 \\ &= -h_{1,1}/12. \end{aligned} \quad (\text{A10})$$

Above, in the second line, we used the fact that $h_{2,0} + 2h_{1,1} + h_{0,2}$ is a total derivative, which cancels the boundary terms. We obtain an exactly equal contribution from the term with the roles of u and u' exchanged. Finally, the contribution of a product of first-order terms in both u and u' gives

$$\begin{aligned} & 2h_{1,0} \times h_{0,1}/3 + h_{1,0}^2/4 + h_{0,1}^2/4 - (|f(0)|^2 - |f(1)|^2)^2/4 \\ &= (h_{1,0} + h_{0,1})^2/4 + h_{1,0} \times h_{0,1}/6 \\ & \quad - (|f(0)|^2 - |f(1)|^2)^2/4 \\ &= h_{1,0}h_{0,1}/6, \end{aligned} \quad (\text{A11})$$

where in the second line we collected a square of the total derivative, which cancels the boundary terms.

Gathering all of the contributions, we obtain the final result for the purity p_k :

$$p_k = 1 - \frac{g_1(f)}{6 \times 4^k} + O(1/8^k), \quad (\text{A12})$$

with

$$g_1(f) \equiv h_{1,1} - h_{1,0}h_{0,1} = \langle f^{(1)}|\hat{P}_0|f^{(1)}\rangle, \quad (\text{A13})$$

where $\hat{P}_0$ is the operator, which projects out the function $f(u)$, and scalar product $\langle f^{(1)}|f^{(1)}\rangle$ should be considered to be defined consistently with the L_2 norm on the interval of function definition. ■

Note that the functional $g_1(f)$ obeys several properties, which serve as a consistency check for the p_k computation. First, it is positive definite for all functions $f(x)$, which corresponds to the purities p_k being smaller or equal to unity, as required. Second, the value is real. Finally, we may ask when it vanishes, i.e., when $g_1(f) = 0$. For this to hold, the derivative of $f(x)$ should be in the kernel of $\hat{P}_0$, which implies that the derivative is proportional to the function itself.

The only solution of such differential equation is the exponential function: $f(x) = c \exp(bx)$. This is consistent with the fact that exponential functions can be exactly expressed as MPS with bond dimension equal to one [37], and hence their purities must vanish because of trivial factorizability on every bond.

Note also that reduced density matrices may be used to construct other entanglement measures. Some of them, like higher-order generalizations of purity $\text{Tr}\rho^4, \text{Tr}\rho^6$ are amenable to the calculations along the lines developed above.

2. Proof of Theorem 2

Let us consider the reduced density matrix of the first k qubits ρ_k and compute its particular element:

$$\rho_k(\mathbf{u}_k, \mathbf{u}'_k) = \frac{1}{2^k} \int_0^1 f(\mathbf{u}_k + v/2^k)f^*(\mathbf{u}'_k + v/2^k)dv, \quad (\text{A14})$$

where $\mathbf{u}_k, \mathbf{u}'_k$ are the combined indices of the reduced density matrix. Expanding the functions f, f^* in $v/2^k$, we can write the reduced density matrix as

$$\rho_k = \frac{1}{2^k} \sum_{m,n \geq 0} \frac{|f^{(m)}(u)\rangle\langle f^{(n)}(u')|}{(1+m+n)n!m!2^{k(m+n)}}, \quad (\text{A15})$$

here $f^{(n)}(u)$ are derivatives of order n , and we used the bracket notation. This is a derivative expansion, but it is crucial that this is also an expansion in $1/2^k$, which implies that only derivatives of small order will contribute. Note that, as in the previous theorem, to obtain asymptotic expansions we consider a situation where $N \rightarrow \infty$.

Note that the states $|f^{(n)}(u)\rangle$ are not normalized and not orthogonal, but can be made so using the Gram-Schmidt method. To this end, we define projectors P_m , projecting out all derivatives up to order m at consecutive steps of the Gram-Schmidt procedure. Particularly, P_0 projects out the function itself. The new orthogonal basis is now given by $\{|f(u)\rangle, P_0|f^{(1)}(u)\rangle, P_1|f^{(2)}(u)\rangle, \dots\}$. The basis change in the reduced density matrix can now be performed by iteratively replacing

$$|f^{(1)}(u)\rangle \rightarrow P_0|f^{(1)}(u)\rangle + (1 - P_0)|f^{(1)}(u)\rangle \quad (\text{A16})$$

in Eq. (A15), collecting the terms, and repeating recursively for higher derivatives. Note that this transformation is *tri-*

angular in a sense that after applying it, only the density matrix coefficients with lower indices are affected [e.g., $(1 - P_0)|f^{(1)}(u)\rangle \propto |f(u)\rangle$ will generate correction to the $|f^{(0)}(u)\rangle$]. Since the initial coefficients of $|f^{(m)}(u)\rangle\langle f^{(n)}(u')|$ are of order $1/2^{k(m+n)}$, the orthogonalization only changes them at subleading orders in $1/2^k$, and can be ignored.

The new basis vectors must be normalized, which gives additional $\sqrt{\langle f^n(u)|P_{n-1}|f^n(u)\rangle}$ factors in the density matrix coefficients. After a simple but tedious calculation, we obtain the following expression for $\rho_{k,m,n}$ in the orthonormal basis, to leading order in $1/2^k$:

$$\rho_{k,m,n} = \frac{\sqrt{\langle f^n(u)|P_{n-1}|f^n(u)\rangle\langle f^m(u)|P_{m-1}|f^m(u)\rangle}}{(m+n+1)m!n!2^{k(m+n)}}, \quad (\text{A17})$$

where $m, n \in \mathbb{N}_0$ and $\rho_{k,m,n}$ is a recomputed term under the sum in Eq. (A15). The scalar products $\langle f^{(m)}(u)|P_{m-1}|f^{(m)}(u)\rangle$ appearing in the density matrix are defined for the function discretized at a finite resolution $1/2^k$. However, as before, via the Euler-MacLaurin formula they can be approximated by continuous integrals with boundary terms (which, as we work to leading order in $1/2^k$ throughout, we ignore). In particular, we have $\langle f^{(1)}(u)|P_0|f^{(1)}(u)\rangle = g_1(f)$.

We will now perturbatively diagonalize the density matrix terms $\rho_{k,m,n}$. As a warm-up, consider the restriction of the problem to the index range $m, n = 0, 1$, where we can diagonalize the matrices. To leading order in $1/2^k$, we obtain the eigenvalues $\rho_{k,0} \approx 1$, $\rho_{k,1} \approx g_1(f)/(12 \times 4^k)$, which precisely agree with the predictions obtained from the expansions of purities (the density matrix eigenvalues are given by $\rho_{k,n} = \Lambda_{k,n}^2$).

Based on this example, we conjecture that eigenvectors are such that restrictions of the matrix to its lower indices preserve the eigenvalues to leading order. Consequently, for the first eigenvector we make the ansatz $e_1 = (1, c_1/2^k, c_2/4^k, c_3/8^k, \dots)$, where c_s are of order 1, and do not depend on k to leading order in $1/2^k$. Substituting this into the eigenvector equation, we indeed find that all coefficients c_s can be determined perturbatively in terms of $\langle f^n(u)|P_{n-1}|f^n(u)\rangle$, while the first eigenvalue is 1, to leading order. Similarly, for the second eigenvector we take $e_2 = (d_0/2^k, 1, d_2/2^k, d_3/4^k, \dots)$, where d_s are again order one coefficients. Substituting into the eigenvector equation allows to find the coefficients d_s , and the corresponding eigenvalue, which, to leading order, is indeed $g_1(f)/(12 \times 4^k)$. The generalization to the m th eigenvector is as follows:

$$e_m = (q_0/2^{mk}, q_1/2^{(m-1)k}, \dots, q_{m-1}/2^k, 1, q_{m+1}/2^k, \dots) \\ = (\vec{q}, 1, \vec{Q}), \quad (\text{A18})$$

where we have introduced vectors $\vec{q}, \vec{Q}$ with elements $\tilde{q}_n = q_n/2^{(m-n)k}$, $\tilde{Q}_n = q_{m+1+n}/2^{k(n+1)}$. We now need to solve the eigenvalue equation

$$\hat{\rho}_k e_m = \rho_{k,m} e_m, \quad (\text{A19})$$

where $\rho_{k,m}$ is the m th eigenvalue of the k th qubit reduced density matrix. To this end, one can write the matrix $\hat{\rho}_k$ in

the block form

$$\hat{\rho}_k = \begin{pmatrix} M & l & D \\ l^\dagger & p & w^\dagger \\ D^\dagger & w & G \end{pmatrix}, \quad (\text{A20})$$

where M is $m \times m$ size matrix, l is a column vector of size m , p is a number, w is a column vector, and D and G are matrices. In this notation, the eigenvalue equation may be written in the following form:

$$M\tilde{q} + l + D\tilde{Q} = \rho_{k,m}\tilde{q}, \quad (\text{A21})$$

$$l^\dagger \cdot \tilde{q} + p + w^\dagger \cdot \tilde{Q} = \rho_{k,m}, \quad (\text{A22})$$

$$D^\dagger \tilde{q} + w + G\tilde{Q} = \rho_{k,m}\tilde{Q}. \quad (\text{A23})$$

We begin solving the system from the second equation. Both $l \cdot \tilde{q}$ and p are of the leading order $1/2^{2mk}$, while $w^\dagger \cdot \tilde{Q} = O(1/2^{(2m+2)k})$. Hence, to the leading order, we can safely neglect the term $w^\dagger \cdot \tilde{Q}$. Note also that for consistency $\rho_{k,m} = O(1/2^{2km})$. Consider now the n th equation (with indexation starting from zero) from the matrix equation $M\tilde{q} + l + D\tilde{Q} = \rho_{k,m}\tilde{q}$. To the leading order, we have $l_n = O(1/2^{k(m+n)})$, $M_n\tilde{q} = O(1/2^{k(m+n)})$, $D_n\tilde{Q} = O(1/2^{k(m+n+2)})$, and $\rho_{k,m}\tilde{q}_n = O(1/2^{3km-nk})$ (note that $n < m$). Only first two terms remain, and to the leading order we obtain

$$M\tilde{q} = -l, \quad (\text{A24})$$

which allows to find the vector $\tilde{q}$. Note that all the factors $1/2^k$ cancel out in these equations, and we obtain a simple linear system on coefficients q_n with $n < m$. The eigenvalue $\rho_{k,m}$ can be found by substituting this solution for $\tilde{q}$ into $\rho_{k,m} = l^\dagger \cdot \tilde{q} + p$. Finally, we should find the vector $\tilde{Q}$. Let us estimate the orders of elements of the n th equation from $D^\dagger \tilde{q} + w + G\tilde{Q} = \rho_{k,m}\tilde{Q}$. We have $D_n^\dagger \tilde{q} = O(1/2^{k(2m+1+n)})$, $w_n = O(1/2^{k(2m+1+n)})$, and $\rho_{k,m}\tilde{Q}_n = O(1/2^{k(2m+1+n)})$, while $G_n\tilde{Q} = O(1/2^{k(2m+n+3)})$. Hence, we can neglect the term $G\tilde{Q}$ and just obtain $\rho_{k,m}\tilde{Q} = D^\dagger \tilde{q} + w$, where all $1/2^k$ factors also cancel out to the leading order.

We can now return to the equation for eigenvalues $\rho_{k,m}$ and further simplify them:

$$\rho_{k,m} = p - l^\dagger M^{-1} l = \frac{\det N}{\det M}, \quad (\text{A25})$$

where the $(m+1) \times (m+1)$ matrix N is defined as follows:

$$N = \begin{pmatrix} M & l \\ l^\dagger & p \end{pmatrix}. \quad (\text{A26})$$

The matrix N is the analog of the matrix M for the eigenvalue $\rho_{k,m+1}$. To simplify the determinants, we notice that the matrices N, M can be factorized as follows: $N = F_{m+1}H_{m+1}F_{m+1}$, $M = F_mH_mF_m$, where F_m is a $m \times m$ diagonal matrix with diagonal elements given by

$$F_{m,n} = \frac{\sqrt{\langle f^n(u)|P_{n-1}|f^n(u)\rangle}}{n!2^{kn}} \quad (\text{A27})$$

for $n = 0, 1, \dots, m-1$, while H_m is an $m \times m$ matrix with matrix elements $H_{m,nl} = \frac{1}{n+l+1}$. With these factorizations, we

obtain the following formula for $\rho_{k,m}$:

$$\begin{aligned}\rho_{k,m} &= \left(\frac{\det F_{m+1}}{\det F_m} \right)^2 \times \frac{\det H_{m+1}}{\det H_m} \\ &= \frac{\langle f^{(m)}(u) | P_{m-1} | f^{(m)}(u) \rangle}{m! 2^{4km}} \times \frac{\det H_{m+1}}{\det H_m}.\end{aligned}\quad (\text{A28})$$

Here, the first factor contains information on the function and the $1/4^{km}$ scaling, while the factor $\frac{\det H_{m+1}}{\det H_m}$ is just a positive numerical coefficient, independent of k and of the details of the function. ■

3. Proof of Corollaries 1 and 2

We prove both results simultaneously.

Let the subleading part of the entanglement spectrum be of order s . Hereafter, we can make additional assumption that the third in size entanglement coefficient is at least of order $O(s^2)$. This assumption follows from Theorem 2. Then, from the state normalization $\sum_i \Lambda_{k,i}^2 = 1$, it follows that $\Lambda_{k,0} = 1 - s^2/2 + O(s^4)$. Using Eq. (9), the purity has now the form

$$p_k = (1 - s^2/2 + O(s^4))^4 + O(s^4) = 1 - 2s^2 + O(s^4). \quad (\text{A29})$$

Comparing with the expansion of purity in $1/2^k$ in Eqs. (A12) and (A13), we obtain the result of the first corollary.

To prove the result for the entanglement entropy, it is enough to compute

$$\begin{aligned}S_k &= -(1 - s^2) \log(1 - s^2) - s^2 \log s^2 \\ &\approx s^2(1 - \log s^2) \sim Ck/4^k\end{aligned}\quad (\text{A30})$$

for some constant C , which one can compute explicitly from other terms.

Note that in both results we used an assumption that s is small enough and asymptotic expansions can be applied. The latter is supported by Theorem 1 for large enough k . ■

4. Breaking the smoothness assumption

Here, we briefly comment on how our results change when the smoothness assumption is not satisfied. Consider the case where the function f has only l absolutely continuous derivatives. Then, by Taylor's theorem it can be expanded in series $f(u + v/2^k) = f(u) + f'(u)v/2^k + \dots + f^{(l)}(u)v^l/l!2^{kl} + R_l(u, v)$, where R_l is the remainder term in the integral representation $R_l(u, v) = \int_0^v f^{(l+1)}(u + z/2^k)(v - z)^l dz/l!2^{k(l+1)}$. This remainder term is on the order of $O(1/2^{k(l+1)})$, and hence it cannot spoil the hierarchy of matrix elements in the density matrix. As a result, all the results that depended only on hierarchy of magnitudes of the matrix elements, like the formula for the leading entanglement spectra, will not change. In particular, the first $(l + 1)$ density matrix eigenvalues will follow Eq. (A28). In numerical experiments, we observed an even stronger result, where the first $l + 2$ eigenvalues behaved followed the scaling derived for the smooth functions. This may be connected to the fact that piecewise polynomials do have exact MPS representations of low rank. We leave a more detailed study of entanglement of nonsmooth functions to future research.

APPENDIX B: EXAMPLES, EXPANSION CONSTANTS, AND PROBABILITY DISTRIBUTIONS

1. Analytical examples for the theoretical smoothness results

The expansions in $1/2^k$ from Theorem 1 are valid when $g_1(f)/(6 \times 4^k) < 1$ [i.e., when the second expansion term in Eq. (13) is smaller than 1]. To shed light on the intuitive understanding of this condition, consider the example of a Gaussian $f(x) = \exp[-(x - 0.5)^2/4\sigma^2]$. Then, $g_1(f) = \langle (x - 0.5)^2/4\sigma^4 \rangle = 1/4\sigma^2$ (we ignored the interval finiteness, which is valid for low enough σ) and the validity condition reduces to $k > \log_2 1/2\sigma - \log_2 6$. As 2σ is the characteristic length scale of the Gaussian, the condition implies that the asymptotic results are valid for qubits k responsible for scales $1/2^k$ smaller than the characteristic length scale of the distribution. Thus, while for small k , corresponding to large-scale function variations, entanglement entropies and other measures can be large, for large k (small-scale variations) they decay exponentially in k .

Next, consider the function $f(x) = \sin(\pi x)$, for which the integrals in the density matrix definition can be computed exactly in full [16]. The density matrix is given by

$$\begin{aligned}\rho_k &= |\sin \pi x\rangle \left(\frac{1}{2} + \frac{2^k}{4\pi} \sin(2\pi/2^k) \right) \langle \sin \pi x| \\ &+ |\cos \pi x\rangle \left(\frac{1}{2} - \frac{2^k}{4\pi} \sin(2\pi/2^k) \right) \langle \cos \pi x| \\ &+ \frac{2^k}{2\pi} \sin^2(\pi/2^k) (|\cos \pi x\rangle \langle \sin \pi x| \\ &+ |\sin \pi x\rangle \langle \cos \pi x|).\end{aligned}\quad (\text{B1})$$

Then,

$$\begin{aligned}\rho_k &= |\sin \pi x\rangle (1 + 2^k \sin(2\pi/2^k)/(2\pi))/2 \langle \sin \pi x| \\ &+ |\cos \pi x\rangle (1 - 2^k \sin(2\pi/2^k)/(2\pi))/2 \langle \cos \pi x| \\ &+ 2^k \sin^2(\pi/2^k)/(2\pi) (|\cos \pi x\rangle \langle \sin \pi x| \\ &+ |\sin \pi x\rangle \langle \cos \pi x|).\end{aligned}\quad (\text{B2})$$

Since $|\sin \pi x\rangle$ and $|\cos \pi x\rangle$ are orthogonal for all k and have the same norm, we do not need the Gram-Schmidt procedure. Diagonalizing the density matrix, we obtain the following eigenvalues:

$$\rho_{k,\pm} = \frac{1}{2} \pm \frac{2^k}{2\pi} \sin(\pi/2^k). \quad (\text{B3})$$

Expanding the smaller one we obtain to leading order $\pi^2/(12 \times 4^k)$, which precisely agrees with Theorem 2, with $g_1(f) = \pi^2$. Note that entanglement spectra decrease with k , so for large enough k , we can truncate the bond dimensions to 1.

2. The third eigenvalue and its sum

For the error analysis, we compute the expression for the third entanglement coefficient $\Lambda_{k,2}$. To the leading order for large enough k , we have

$$\Lambda_{k,2} \approx \sqrt{\frac{\langle f^{(2)}(u) | P_1 | f^{(2)}(u) \rangle}{720 \times 16^k}} = \sqrt{\frac{g_2(f)}{720 \times 16^k}}, \quad (\text{B4})$$

where we defined $g_2(f) = \langle f^{(2)}(u) | P_1 | f^{(2)}(u) \rangle$.

The sum of the remaining eigenvalues appears in computing fidelities between an ideal state and a circuit's state after a single layer of the analytical decomposition. For the entanglement spectra, from Theorem 2 we know that $\Lambda_{k,i} = O(1/2^{ki})$. Using these scaling relations, we can derive the leading estimate of the fidelity after one layer of a generalized analytic decomposition:

$$\sum_{\substack{k \geq m \\ i \geq 2}} \Lambda_{k,i}^2 \approx g_2(f)/720 \times 15 \times 16^{m-1} + \dots, \quad (\text{B5})$$

where m is the first bond with truncation. It can be the second bond (the first bond always has the bond dimension 2 and is not truncated) or some larger bond if we use analytic decomposition with three-site gates for the first qubits. Note also that for one-layer analytic decomposition, the resulting fidelity does not depend on the position of the canonical form center, as it does not change the error of truncation.

3. Exact functional coefficients for probability distributions

Here, we give explicit expressions for the $g_1(f)$ functions for the symmetrical normal ($\mu = L/2$), log-normal, and Lévy distributions, obtained by performing the integrals in Eq. (A13) for these examples:

$$\begin{aligned} g_1(\sqrt{p(x)_{\text{normal}}}) &= \frac{L^2}{4\sigma^2} \left[1 - \sqrt{\frac{L^2}{2\sigma^2}} \frac{\exp(-\frac{L^2}{8\sigma^2})}{(\Gamma(1/2) - \Gamma(1/2, L^2/8\sigma^2))} \right], \end{aligned} \quad (\text{B6})$$

$$\begin{aligned} g_1(\sqrt{p(x)_{\text{log-normal}}}) &= \frac{L^2 \exp(-2\mu + 2\sigma^2)}{4\sigma^2 [\sqrt{2\pi} - \frac{1}{\sqrt{2}} \Gamma(\frac{1}{2}, \frac{((\log L - \mu)/\sigma)^2}{2})]} \\ &\times \left[\frac{1 + \sigma^2}{\sqrt{2}} \left(2\sqrt{\pi} - \Gamma\left(\frac{1}{2}, \frac{((\log L - \mu)/\sigma + 2\sigma)^2}{2}\right) \right) \right. \\ &+ \left. \left((2 - \sqrt{2})\sigma - \frac{\log L - \mu}{\sqrt{2}\sigma} \right) \exp\left[\frac{(\frac{\log L - \mu}{\sigma} + 2\sigma)^2}{-2}\right] \right] \\ &- \frac{L^2 \exp\{-2\mu + \sigma^2 - ((\log L - \mu)/\sigma + \sigma)^2\}}{2\sqrt{2}\sigma^2 [2\sqrt{\pi} - \Gamma(\frac{1}{2}, \frac{((\log L - \mu)/\sigma)^2}{2})]^2}, \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} g_1(\sqrt{p(x)_{\text{Levy}}}) &= \frac{21L^2}{8c^2} + \frac{4L^2 \exp[-c/2L]}{c^2 \Gamma(1/2, c/2L)} \left[\frac{21}{16} \left(\frac{c}{2L}\right)^{1/2} + \frac{7}{8} \left(\frac{c}{2L}\right)^{3/2} \right. \\ &+ \left. \frac{1}{8} \left(\frac{c}{2L}\right)^{5/2} + \frac{1}{4} \left(\frac{c}{2L}\right)^{7/2} \right] \\ &- \frac{L^2}{c^2 \Gamma(1/2, c/2L)^2} \left(\frac{c}{2L}\right)^3 \exp[-c/L]. \end{aligned} \quad (\text{B8})$$

For the Lévy distribution, we computed $g_1(f)$ for the discretization on the finite interval L . The leading term $21L^2/8c^2$ defines the characteristic scale $k = \log L/c$, where the entanglement spectra become small. Other terms define the

subleading corrections due to the finiteness of the interval. Here, $\Gamma(s, x)$ is the incomplete Γ function, which comes from the normalization of the Lévy distribution on the finite interval.

Analogous explicit formulas can also be derived for the $g_2(f)$ coefficient. Here, we give the one for the same normal distribution (the others result in impractically large expressions):

$$\begin{aligned} g_2(\sqrt{p(x)_{\text{normal}}}) &= \frac{L^4}{8\sigma^4} - \frac{L^4}{4\sigma^4} \left(\sqrt{\frac{L^2}{32\sigma^2}} + \sqrt{\frac{L^2}{8\sigma^2}} \right) \frac{\exp[-\frac{L^2}{8\sigma^2}]}{\Gamma(\frac{1}{2}) - \Gamma(\frac{1}{2}, \frac{L^2}{8\sigma^2})} \\ &- \frac{L^6}{32\sigma^6} \frac{\exp[-\frac{L^2}{4\sigma^2}]}{[\Gamma(\frac{1}{2}) - \Gamma(\frac{1}{2}, \frac{L^2}{8\sigma^2})]^2}. \end{aligned} \quad (\text{B9})$$

If we drop the finiteness of the interval, that is, consider the large L limit, the expression simplifies to

$$g_2(\sqrt{p(x)_{\text{normal}}}) = \frac{1}{8\sigma^4}. \quad (\text{B10})$$

APPENDIX C: ENTANGLEMENT AND FUNCTION LOCALIZATION

In Sec. III A, we showed that the entanglement on the smallest scales exhibits universal exponential decay. We can similarly ask how the entanglement behaves on the qubits corresponding to the largest scales.

It is easy to see that this behavior *cannot be universal*, by the following construction: For any given quantum state on $n < N$ qubits, one can convert its amplitudes into the corresponding tensor of 2^n discrete “function values,” which can further be polynomially interpolated to yield a smooth function. This function can have an arbitrarily complicated behavior down to the scale $L/2^n$. On the other hand, a function can equally well be arbitrarily simple on large scales, such as for instance in the case of a periodic function with a period T . It is easy to see that the qubits responsible for the larger scales L' , i.e., $T < L' < L$, will be in an equal weight superposition (a product of $|+\rangle$ states), and completely disentangled from the qubits corresponding to the smaller scales. These two opposite extremes show that there is no universality for large scales.

While we do not provide here rigorous results, it is possible to characterize the large-scale entanglement slightly better for specific classes of functions. An important case is the *localized* functions. Informally, a function is localized exponentially or polynomially if the integral of the function over increasing subsets containing the domain of localization approaches unity (we assumed normalized function) with the remainder decaying polynomially or exponentially in the size of the subsets.

We can follow an analogous path to talk about localization on a finite domain, but the presence of a binary grid, which we work with, introduces additional subtleties. For the purpose of clarity, and at the expense of full generality, we restrict the discussion to a localization on a subinterval I_m of size $L/2^m$ (and scale m), which is “compatible” with the binary grid. We define a compatible interval $I_{\sigma_1 \dots \sigma_m} \equiv I_m$ to be the set of

all points sharing a fixed first m bits $\sigma_1 \dots \sigma_m$ of the binary expansion. A sequence of increasing intervals $I_m \subset I_{m-1} \subset \dots \subset I_1$ exists, defined by freeing the successive digits in the binary expansion. For the function localized exponentially (resp. polynomially) to I_m , we now have

$$\int_{I_l} |f(u)|^2 du \approx 1 - C(|I_l|), \quad (\text{C1})$$

for $l = m, \dots, 1$, where the correction $C(|I_l|)$ decreases exponentially (polynomially) in the size of I_l .

Note now that $\int_{I_l} |f(u)|^2 du$ is in fact [see Eq. (A14)] the diagonal element of the reduced density matrix $\rho_{l,u_l u_l}$, with $u_l \equiv \sigma_1 \dots \sigma_l$. Thus, $\rho_{l,u_l u_l}$ is exponentially or polynomially close to one, for localized functions. Conversely, the off-diagonal elements of the reduced density matrix by Eq. (A14) will be exponentially/polynomically suppressed, which immediately implies that the entanglement across bonds $m, \dots, 1$ is also similarly suppressed (see Fig. 3). A delta distribution is an extreme example of this: All but on elements of the density matrix vanish, and the distribution can be represented as a product state. Thus, localization properties influence the behavior of the entanglement on the bonds representing the largest spatial scales.

APPENDIX D: ACCURACY-DEPTH TRADE-OFF IN THE PRESENCE OF NOISE

As discussed in the main text, noise induces a practical trade-off, where theoretically better but deeper circuits result

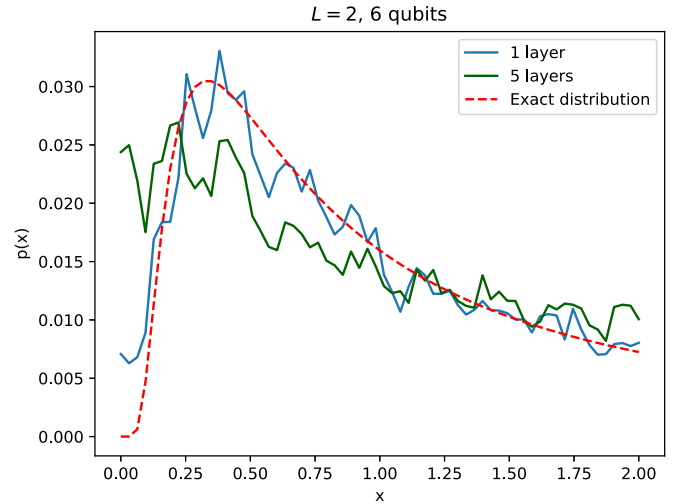


FIG. 12. The Lévy distribution encoded with one- and five-layer circuits on a noisy simulated quantum device (IBM FakeToronto). The output of the theoretically more accurate but deeper circuit is much more strongly affected by the noise. Such trade-off effects are very important in practice.

in worse approximation of the distribution. To probe this, we performed noisy simulations on the IBM FakeToronto backend. We compare the results of the execution of a one-layer and five-layer encoding circuits in Fig. 12. While in the absence of noise the latter is more accurate, with noise the distribution features are largely destroyed, while the nominally less accurate one-layer encoding performs much better.

- [1] V. V. Shende, S. S. Bullock, and I. L. Markov, Synthesis of quantum-logic circuits, *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **25**, 1000 (2006).
- [2] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, *npj Quantum Inf.* **5**, 103 (2019).
- [3] J. J. García-Ripoll, Quantum-inspired algorithms for multivariate analysis: From interpolation to partial differential equations, *Quantum* **5**, 431 (2021).
- [4] K. Endo, T. Nakamura, K. Fujii, and N. Yamamoto, Quantum self-learning Monte Carlo and quantum-inspired Fourier transform sampler, *Phys. Rev. Res.* **2**, 043442 (2020).
- [5] M. Moosa, T. W. Watts, Y. Chen, A. Sarma, and P. L. McMahon, Linear-depth quantum circuits for loading Fourier approximations of arbitrary functions, *Quantum Sci. Technol.* **9**, 015002 (2024).
- [6] K. Dasgupta and B. Paine, Loading probability distributions in a quantum circuit, [arXiv:2208.13372](https://arxiv.org/abs/2208.13372).
- [7] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011).
- [8] J. Ignacio Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, *Rev. Mod. Phys.* **93**, 045003 (2021).
- [9] S.-J. Ran, Encoding of matrix product states into quantum circuits of one- and two-qubit gates, *Phys. Rev. A* **101**, 032310 (2020).
- [10] M. S. Rudolph, J. Chen, J. Miller, A. Acharya, and A. Perdomo-Ortiz, Decomposition of matrix product states into shallow quantum circuits, *Quantum Sci. Technol.* **9**, 015012 (2024).
- [11] P.-F. Zhou, R. Hong, and S.-J. Ran, Automatically differentiable quantum circuit for many-qubit state preparation, *Phys. Rev. A* **104**, 042601 (2021).
- [12] M. Ben-Dov, D. Shnaiderov, A. Makmal, and E. G. Dalla Torre, Approximate encoding of quantum states using shallow circuits, *npj Quantum Inf.* **10**, 65 (2024).
- [13] D. Malz, G. Styliaris, Z.-Y. Wei, and J. Ignacio Cirac, Preparation of matrix product states with log-depth quantum circuits, *Phys. Rev. Lett.* **132**, 040404 (2024).
- [14] K. C. Smith, A. Khan, B. K. Clark, S. M. Girvin, and T.-C. Wei, Constant-depth preparation of matrix product states with adaptive quantum circuits, *PRX Quantum* **5**, 030344 (2024).
- [15] B. Ljubomir, Quantum algorithm for the Navier–Stokes equations by using the streamfunction-vorticity formulation and the lattice Boltzmann method, *Int. J. Quantum Inf.* **20**, 2150039 (2022).

- [16] I. Oseledets, Constructive representation of functions in low-rank tensor formats, *Construct. Approxim.* **37**, 1 (2013).
- [17] A. Holmes and A. Y. Matsuura, Efficient quantum circuits for accurate state preparation of smooth, differentiable functions, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, USA, 2020), pp. 169–179.
- [18] A. Holmes and A. Y. Matsuura, Entanglement properties of quantum superpositions of smooth, differentiable functions, [arXiv:2009.09096](https://arxiv.org/abs/2009.09096).
- [19] G. Marin-Sanchez, J. Gonzalez-Conde, and M. Sanz, Quantum algorithms for approximate function loading, *Phys. Rev. Res.* **5**, 033114 (2023).
- [20] M. Lindsey, Multiscale interpolative construction of quantized tensor trains, [arXiv:2311.12554](https://arxiv.org/abs/2311.12554).
- [21] P. Gundlapalli and J. Lee, Deterministic and entanglement-efficient preparation of amplitude-encoded quantum registers, *Phys. Rev. Appl.* **18**, 024013 (2022).
- [22] J. Iaconis, S. Johri, and E. Y. Zhu, Quantum state preparation of normal distributions using matrix product states, *npj Quantum Inf.* **10**, 15 (2024).
- [23] A. A. Melnikov, A. A. Termanova, S. V. Dolgov, F. Neukart, and M. R. Perelshtein, Quantum state preparation using tensor networks, *Quantum Sci. Technol.* **8**, 035027 (2023).
- [24] I. Y. Akhalwaya, A. Connolly, R. Guichard, S. Herbert, C. Kargi, A. Krajenbrink, M. Lubasch, C. M. Keever, J. Sorci, M. Spranger, and I. Williams, A modular engine for quantum Monte Carlo integration, [arXiv:2308.06081](https://arxiv.org/abs/2308.06081).
- [25] Y. Sano and I. Hamamura, Quantum state preparation for probability distributions with mirror symmetry using matrix product states, *Phys. Rev. Appl.* **24**, 034062 (2025).
- [26] J. Gonzalez-Conde, T. W. Watts, P. Rodriguez-Grasa, and M. Sanz, Efficient quantum amplitude encoding of polynomial functions, *Quantum* **8**, 1297 (2024).
- [27] S. Woerner and D. J. Egger, Quantum risk analysis, *npj Quantum Inf.* **5**, 15 (2019).
- [28] N. Stamatopoulos, D. J. Egger, Y. Sun, C. Zoufal, R. Iten, N. Shen, and S. Woerner, Option pricing using quantum computers, *Quantum* **4**, 291 (2020).
- [29] J.-P. Bouchaud and M. Potters, *Théorie Des Risques Financiers* (Aléa-Saclay, Eyrolles, Paris, 1997).
- [30] R. Cont and P. Tankov, Non-parametric calibration of jump-diffusion option pricing models, *J. Comput. Finance* **7**, 1 (2004).
- [31] M. Z. Bazant, Random walks and diffusion, MIT OpenCourseWare (2003), <https://ocw.mit.edu/>.
- [32] B. Mandelbrot, *Fractals and Scaling in Finance* (Springer, New York, 1997).
- [33] P. J. S. De Brouwer, Maslowian portfolio theory: An alternative formulation of the behavioural portfolio theory, *J. Asset Manag.* **9**, 359 (2009).
- [34] P. J. S. De Brouwer, Target-oriented investment advice, *J. Asset Manag.* **13**, 102 (2011).
- [35] R. N. Mantegna and H. Eugene Stanley, Scaling behavior in the dynamics of an economic index, *Nature (London)* **376**, 46 (1995).
- [36] Y. Núñez Fernández, M. K. Ritter, M. Jeannin, J.-W. Li, T. Kloss, T. Louvet, S. Terasaki, O. Parcollet, J. von Delft, H. Shinaoka, and X. Waintal, Learning tensor networks with tensor cross interpolation: New algorithms and libraries, *SciPost Phys.* **18**, 104 (2025).
- [37] B. N. Khoromskij, $O(d \log N)$ -quantics approximation of N -d tensors in high-dimensional numerical modeling, *Construct. Approxim.* **34**, 257 (2011).
- [38] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
- [39] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, *Quantum Inf. Comput.* **7**, 401 (2007).
- [40] T. Xiang, Matrix product states, in *Density Matrix and Tensor Network Renormalization* (Cambridge University Press, Cambridge, 2023), pp. 154–165.
- [41] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, *Phys. Rev. A* **101**, 010301(R) (2020).
- [42] J. J. Rodríguez-Aldavero, P. García-Molina, L. Tagliacozzo, and J. J. García-Ripoll, Chebyshev approximation and composition of functions in matrix product states for quantum-inspired numerical analysis, [arXiv:2407.09609](https://arxiv.org/abs/2407.09609).
- [43] I. Oseledets and E. Tyrtshnikov, TT-cross approximation for multidimensional arrays, *Linear Algebra Appl.* **432**, 70 (2010).
- [44] M. K. Ritter, Y. Núñez Fernández, M. Wallerberger, J. von Delft, H. Shinaoka, and X. Waintal, Quantics tensor cross interpolation for high-resolution parsimonious representations of multivariate functions, *Phys. Rev. Lett.* **132**, 056501 (2024).
- [45] D. Savostyanov and I. Oseledets, Fast adaptive interpolation of multi-dimensional arrays in tensor train format, in *The 2011 International Workshop on Multidimensional (nD) Systems* (IEEE, New York City, 2011), pp. 1–8.
- [46] M. Bebendorf, Approximation of boundary element matrices, *Numer. Math.* **86**, 565 (2000).
- [47] P. Carr, H. Geman, D. Madan, and M. Yor, The fine structure of asset returns: An empirical investigation, *J. Business* **75**, 305 (2002).
- [48] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta, Quantum computing with Qiskit, [arXiv:2405.08810](https://arxiv.org/abs/2405.08810).
- [49] R. Iten, R. Colbeck, I. Kukuljan, J. Home, and M. Christandl, Quantum circuits for isometries, *Phys. Rev. A* **93**, 032318 (2016).
- [50] J. Hauschild, E. Leviatan, J. H. Bardarson, E. Altman, M. P. Zaletel, and F. Pollmann, Finding purifications with minimal entanglement, *Phys. Rev. B* **98**, 235163 (2018).
- [51] H.-R. Wei and Y.-M. Di, Decomposition of orthogonal matrix and synthesis of two-qubit and three-qubit orthogonal gates, *Quantum Inf. Comput.* **12**, 0262 (2012).
- [52] I. G. Ion, torchTT, GitHub (2024), <https://github.com/ion-g-ion/torchTT>.