communications physics

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42005-025-02361-1

Quantum Chebyshev probabilistic models for fragmentation functions

Jorge J. Martínez de Lejarza [1,6] ✉, Hsin-Yu Wu [2,3,6], Oleksandr Kyriienko [4], Germán Rodrigo [1] & Michele Grossi [5]

Quantum generative modeling is emerging as a powerful tool for advancing data analysis in high-energy physics, where complex multivariate distributions are common. However, efficiently learning and sampling these distributions remains challenging. We propose a quantum protocol for a bivariate probabilistic model based on shifted Chebyshev polynomials, trained as a circuit-based representation of two correlated variables, with sampling performed via quantum Chebyshev transforms. As a key application, we study fragmentation functions (FFs) of charged pions and kaons from single-inclusive hadron production in electron-positron annihilation. We learn the joint distribution of momentum fraction z and energy scale Q, and infer their correlations from the entanglement structure. Building on the generalization capabilities of the quantum model and extended register architecture, we perform fine-grid multivariate sampling for FF dataset augmentation. Our results highlight the growing potential of quantum generative modeling to advance data analysis and scientific discovery in high-energy physics.

The ability to analyze scientific data relies on designing advanced machine learning tools, inferring correlations, and performing generative modeling by sampling synthetic distributions[1]. Multivariate differential distributions represent a major interest in high-energy physics (HEP), specifically in the context of quantum chromodynamics (QCD) and studies of scattering processes[2]. Quantum computing shows promise for accelerating calculations of physical properties and is believed to be vital for large-scale QCD simulations in the future[3–5]. Moreover, quantum machine learning (QML) methods become increasingly more developed[6–8]. Here, HEP data analysis is a suitable example of learning processes that are inherently quantum. For instance, to date, QML has been applied to jet clustering[9–12], elementary particle process integration[13–15], anomaly detection[16–18], data classification[19,20], etc. QML was also used to analyze the causal structure of multi-loop Feynman diagrams and integrate them to predict decay rates at higher orders in perturbation theory[21–26]. Finally, an increasing part of QML studies is targeting generative modeling and preparing quantum circuits that can mimic relevant probability distributions[27–33].

Fragmentation functions (FFs) represent an essential component in describing quantum particle processes, encoding the transition from partons—quarks and gluons—to hadrons after a hard-scattering event[34–36]. Unlike partonic cross-sections, FFs cannot be predicted perturbatively and must be obtained from experimental data via global analysis across diverse processes[37,38]. Several approaches were proposed to extract FFs from experimental data using statistical tools and heuristics, for instance, a polynomial-based ansatz in the hadron's momentum fraction[39–42], and Euler Beta function distributions with additional parameters[34–36]. Despite the progress, the determination of FFs can still be affected by procedural bias, including limited functional form flexibility and uncertainty estimation challenges. Machine learning (ML) techniques, particularly based on neural networks, have emerged to potentially overcome these limitations[43–46].

While supervised ML approaches are effective in determining FFs and Parton Density Functions (PDF) in hadronic collisions[47], some challenges remain. A primary issue lies in the energy-scale evolution. Models are typically trained for a specific range of momentum fraction z and a fixed energy scale Q. However, evolving a given FF to different energy scales requires solving the DGLAP evolution equations[48–51], a task that is computationally expensive. Additionally, once the FFs are determined, their true value are only known for specific points, necessitating interpolation to estimate the function at other regions, increasing the complexity of the method. Finally, given the inherent nature of FFs as probability distributions, which require inversion to get samples, and the need for multivariate descriptions based on both momentum fraction (z) and the energy scale (Q), alternative methods are being explored.

[1]Instituto de Física Corpuscular, Universitat de València—Consejo Superior de Investigaciones Científicas, Parc Científic, Valencia, Spain. [2]Department of Physics and Astronomy, University of Exeter, Exeter, UK. [3]Pasqal, Massy, France. [4]School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, UK. [5]European Organization for Nuclear Research (CERN), Geneva, Switzerland. [6]These authors contributed equally: Jorge J. Martínez de Lejarza, Hsin-Yu Wu. ✉e-mail: jormard@ific.uv.es

Quantum computing naturally fits generative modeling due to the probabilistic nature of projective measurements and natural sampling abilities from strongly correlated distributions[52,53]. Following the Born rule, the probability of measuring a bitstring x is $p(x) = |\langle x|\psi\rangle|^2$, where a wavefunction $|\psi\rangle$ is prepared by some parametrized quantum circuit. We refer to models that represent such circuit-based distributions as quantum probabilistic models of an implicit type[30,31]. These models can be sampled, offering generative modeling capabilities as commonly needed in ML[54]. Quantum generative modeling was demonstrated with quantum generative adversarial networks (QGANs)[55–58], quantum Boltzmann machines (QBMs)[59–61], and quantum circuit born machines (QCBMs)[62–67]. To date, quantum generative modeling has shown intriguing results for HEP analysis[3,32,68–71]. Also, quantum-inspired approaches based on tensor networks are explored as probabilistic models to analyze multivariate distributions[72–74].

Both Fourier and Chebyshev expansions are common methods for approximating smooth functions using global basis functions. Fourier expansions, which use the complex exponential form of sinusoidal functions, are ideal for periodic functions on uniform grids. In contrast, Chebyshev expansions, which utilize Chebyshev polynomials, are particularly suited for non-periodic functions defined on finite intervals. In our recent works[33,75], we have developed quantum Chebyshev and Hartley feature map circuits. These circuits are designed to encode input data as real-valued quantum states associated with real expansion coefficients. We have also built their corresponding transform circuits responsible for basis mappings, which are essential for generative modeling. For this specific study, we are focusing primarily on Chebyshev expansions. We have chosen this approach because it offers the best polynomial approximation to a continuous function under the maximum norm, making it particularly well-suited for approximating the FFs relevant to our work.

In this work, we propose a multivariate quantum probabilistic model that represents HEP data involving several correlated variables, and apply it to describe FFs for the hadronization process. Specifically, we develop quantum Chebyshev probabilistic models (QCPMs) with shifted Chebyshev feature maps and a correlation circuit for multivariate probability distributions[33,75]. Our approach allows encoding models *explicitly* in the basis of orthogonal Chebyshev polynomials, and transforms trained QML models of variables z and Q into the bivariate distribution $p(z, Q)$ ready for sampling. Such an approach provides an ability to infer correlations between variables, giving an insight into HEP data, as well as augmenting HEP datasets with generative modeling on extended grids that grow exponentially with the system size. While classical approaches, such as[44], focus on fitting frameworks that provide statistically robust estimations of FF uncertainties, our approach differs in scope and goals. We do not aim to quantify uncertainties; instead, we go beyond traditional polynomial parametrizations of FFs. Using the Chebyshev feature map, we build a more flexible and expressive representation that can capture complex structures missed by standard fits. Furthermore, our method allows efficient interpolation, with resolution growing exponentially with the number of qubits. QPCM training uses small circuits to keep optimization tractable. After training, we exploit the exponential growth of Hilbert space with qubit count to sample at higher resolutions. This separation of training and sampling demonstrates QPCM's interpolation ability without exceeding classical simulation limits, highlighting its potential for efficient, high-resolution sampling with low training cost. The small number of qubits required for both training and sampling makes hardware execution feasible and also allows efficient simulation using tensor networks. However, scaling the sampling to a larger number of qubits for extremely high resolution would require a bond dimension that grows exponentially, which may limit the performance of the tensor network implementation.

Results and discussion

FFs and hadronization

Our goal is to study HEP processes which: (1) lead to correlated data; (2) require working with multivariate distributions that are hard to study classically; (3) benefit from native abilities of quantum devices for sampling non-trivial probability distributions[53]. Here, we concentrate on parton hadronization. These processes are described by FFs, being two-dimensional probability distributions that possess the properties outlined above. In practice, FFs define differential cross-sections for processes that lead to specific hadrons (e.g., pions or kaons) as a result (Fig. 1a). For the single-inclusive production of a hadron h in electron-positron annihilation, the differential cross-section is defined from

$$\frac{d\sigma^{\mathrm{h}}}{dz}(z, Q^2) = \frac{4\pi\alpha^2(Q)}{Q^2} F^{\mathrm{h}}(z, Q^2), \tag{1}$$

where $F^{\mathrm{h}}(z, Q^2)$ is the fragmentation *structure* function, and $\alpha(Q)$ represents the quantum electrodynamics (QED) running coupling.

The structure function $F^{\mathrm{h}}(z, Q^2) \propto \sum_{i,j} C_j(z, \alpha_s(Q)) \otimes D_i^{\mathrm{h}}(z, Q)$ is defined from a convolution of coefficient functions $C_j(z, \alpha_s(Q))$ and FFs $D_i^{\mathrm{h}}(z, Q)$, where indices i and j denote relevant configurations (e.g., singlet or non-singlet), and $\alpha_s(Q)$ is the QCD running coupling. The convolution operation $\otimes$ is defined as $f(z) \otimes g(z) = \int_z^1 (dy/y) f(y) g(z/y)$. The FFs $D_i^{\mathrm{h}}(z, Q)$ follow the DGLAP evolution equations[50], where one needs to perform the evolution of the FFs with the energy scale Q to solve Eq. (1). As an example, the evolution of the non-singlet component according to the DGLAP equation is given by

$$\frac{\partial}{\partial \log(Q^2)} D_{\mathrm{NS}}^{\mathrm{h}}(z, Q^2) = P_+(z, \alpha_s) \otimes D_{\mathrm{NS}}^{\mathrm{h}}(z, Q^2), \tag{2}$$

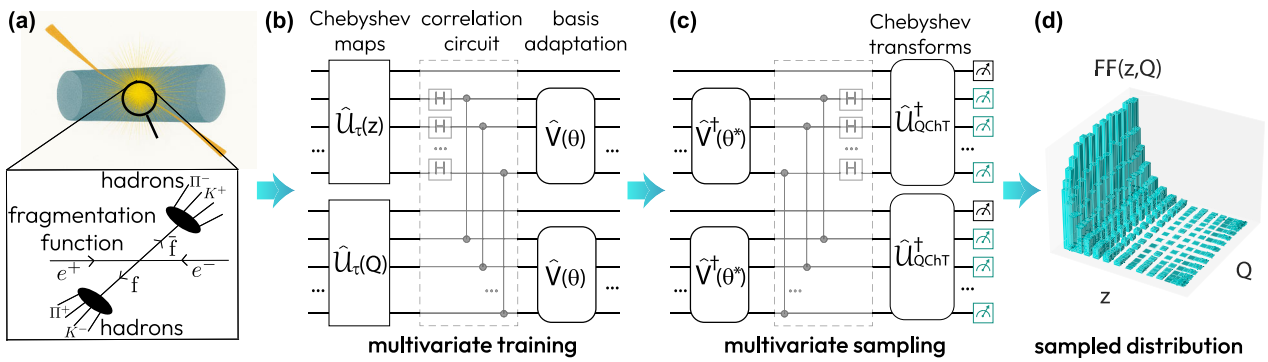

Fig. 1 | Workflow for describing FFs with quantum probabilistic models based on shifted Chebyshev polynomials. a The input data for the FF $D_i^{\mathrm{h}}(z, Q)$ is produced for a grid of z and Q. **b** The quantum probabilistic model is composed of two Chebyshev feature maps for encoding z and Q, a correlation circuit that entangles both registers, and basis adaptation circuits to be trained on $D_i^{\mathrm{h}}(z, Q)$. **c** For sampling, we perform the inverse of basis adaptation, the correlation circuit, followed by parallel inverse quantum Chebyshev transforms for mapping the model into the bit basis. **d** Sampling results assembled in a 2-dimensional plot that represents $D_i^{\mathrm{h}}(z, Q)$.

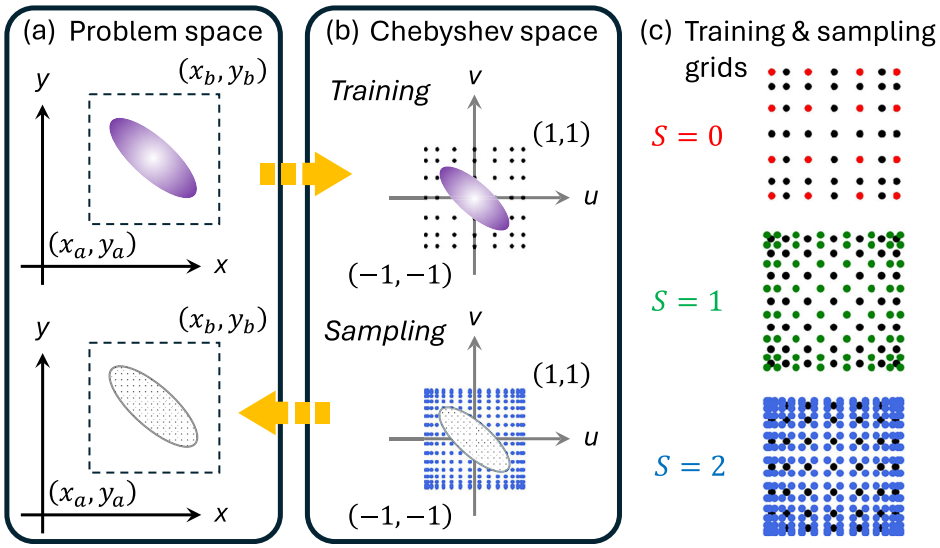

Fig. 2 | Applying linear maps and setting up QCPM. Application of shifted Chebyshev polynomials to map between **a** the problem and **b** Chebyshev spaces. **c** Visualization of training (black) and sampling (color) grids for different extended registers.

where $P_+(z, \alpha_s)$ is a splitting function having a perturbative expansion in the strong coupling $\alpha_s(Q)$ (see Supplementary Note 1). The integro-differential equation in Eq. (2) is, in general, difficult to solve. In our generative modeling, we take a data-driven approach to bypass the DGLAP propagation.

QCPMs

We proceed to introduce a framework that is suitable for describing FFs and associated probabilistic processes. The resulting models need to provide access to samples of multivariate distributions yet have a form such that they can be analyzed[76,77]. To match the desired requirements, we propose to build QCPMs. These models are based on the quantum Chebyshev toolbox[75], which allows building models in the basis of orthogonal polynomials. Given the typical shape of distributions, the Chebyshev basis[78] fits naturally for describing FFs. We extend this to a generalized setting, specifically developing quantum circuits for two variables, correlating them, and demonstrating sampling capabilities on the extended grids.

The core of QCPM is represented by a probability distribution parametrized as $p(x) = |\langle\tau(x)|\psi\rangle|^2$, where $|\psi\rangle \equiv \hat{V}|0\rangle$ is a quantum state prepared by a unitary circuit $\hat{V}$, and $|\tau(x)\rangle \equiv \hat{\mathcal{U}}_\tau(x)|0\rangle$ is a Chebyshev state generated by the corresponding feature map $\hat{\mathcal{U}}_\tau$. The Chebyshev state is given by ref. 75

$$|\tau(x)\rangle = \frac{1}{2^{N/2}} T_0(x)|0\rangle + \frac{1}{2^{(N-1)/2}} \sum_{k=1}^{2^N-1} T_k(x)|k\rangle, \tag{3}$$

where $T_k(x) \equiv \cos(k \arccos(x))$ is a degree k Chebyshev polynomials of the first kind. These are typically defined in the domain of $[-1, 1]$. At the Chebyshev grid points, $\{|\tau(x_j)\rangle\}_{j=0}^{2^N-1}$ there is a set of orthonormal states, which are mapped by a unitary transform to the computational basis states, $\{|x_j\rangle\}_{j=0}^{2^N-1}$. Importantly, by applying Chebyshev feature maps in parallel, we generate explicit models of several variables, and can introduce correlations by a circuit that connects the registers. This leads to QCPM of the form $p(x, y) = \sum_{k,l=0}^{2^N-1} c_{k,l} T_k(x) T_l(y)$, where $c_{k,l}$ are defined by quantum circuits $\hat{V}$. Our goal is to train such models on FF-based datasets and enable sampling.

The workflow for QCPM-based HEP analysis is shown in Fig. 1. In Fig. 1a, we visualize a hadronization process, described by the FF $D_j^h(z, Q)$ and the associated data. These data are transformed into the Chebyshev space and fed into the training (Fig. 1b). Once training is complete, the inferred model is used for sampling a multivariate back in the problem space (Fig. 1c). This leads to an augmented FF distribution visualized in Fig. 1d.

As an important methodological detail for FF modeling, we address the challenge of generalizing our models from the original problem domain to the Chebyshev native domain $\Omega_C = [-1, 1] \times [-1, 1]$ and back. This is illustrated in Fig. 2. Given a 2D probability density function $p(x, y)$ defined in a box domain delimited by coordinate tuples, $\Omega_P = [x_a, x_b] \times [y_a, y_b]$, we express it in the scaled domain as $p(u, v) = \sum_{k,l=0}^{2^N-1} c_{k,l} T_k(u) T_l(v)$, where $u(x) = [2x - (x_a + x_b)]/(x_b - x_a)$ and $v(y) = [2y - (y_a + y_b)]/(y_b - y_a)$. The transformation is shown in Fig. 2a, b (top). Using the Chebyshev feature maps for the scaled variables $\hat{\mathcal{U}}_\tau(u)$ and $\hat{\mathcal{U}}_\tau(v)$, we build the parameterized QCPM $p_Q(u, v)$ to approximately represent the transformed probability density function $p(u, v)$. The quantum model is trained on a 2D grid of Chebyshev nodes $\{u_i^{\mathrm{Ch}}\}_{i=0}^{2^N-1} \times \{v_j^{\mathrm{Ch}}\}_{j=0}^{2^N-1}$ plus additional half-index points $\{u_{i+1/2}^{\mathrm{Ch}}\}_{i=0}^{2^N-2} \times \{v_{j+1/2}^{\mathrm{Ch}}\}_{j=0}^{2^N-2}$. The training grid is schematically illustrated in Fig. 2b (top) as a black-dotted grid within Ω_C. Once the model is trained, sampling is carried out through projective measurements in the same domain (Fig. 2b, bottom), using inverse quantum Chebyshev transform circuits $\hat{\mathcal{U}}_{\mathrm{QChT}}^\dagger$. Details of implementing QCPM are provided in the "Methods" section.

The sampled $p_Q(u, v)$ is mapped back to the problem domain to obtain the sampled histogram $p_Q(x, y)$, rescaled as $x(u) = [(x_b - x_a)u + (x_a + x_b)]/2$ and $y(v) = [(y_b - y_a)v + (y_a + y_b)]/2$ and reshaped in a suitable form. This is shown in Fig. 2(a, bottom). Crucially, the quantum model is trained on a sparse training grid given by the available training dataset, but can provide samples (predictions) for unseen data points on a controllable, denser *sampling grid*, which is not necessarily a superset of the *training grid*. This is based on extending the register to $2(N + S)$, where S is the number of added qubits on which we act with the Chebyshev transform. In this case, the learned distribution from the training data can closely mimic the physical distribution for the input data, generalizing (interpolating) beyond the training grid.

In this work, we applied shifted Chebyshev polynomials to study the FFs as a testbed. Shifted Chebyshev polynomials can approximate any continuous function beyond the Chebyshev domain and thus are not limited to FFs. Because the coefficients in the Chebyshev expansion are purely real, it is straightforward that a purely real quantum state should be used. The Chebyshev encoding provides a convenient way to encode each input datum, v, as a distinct real quantum state, $\tau(v)$, in terms of orthonormal bases. Most importantly, the proposed QCPMs allow us to systemically and modularly extend to a larger system size and allow us to explore the underlying data correlations inferred by QCPMs. The correlation circuit shown in Fig. 1b is a problem-specific ansatz designed to efficiently train the FFs. The ansatz circuit can be tailored to fit various applications in different fields.

In addition, the modular capability of QCPM allows us to easily scale up the model dimension at the expense of increasing the circuit complexity, mainly contributed by quantum Chebyshev transform circuits. Please refer to our previous work[75] for details of the quantum circuits.

Application of QCPM to FFs

We have applied QCPM to FFs by using data accessible via the LHAPDF6 interface[79,80]. The FF datasets are derived from hadron production in electron-positron single-inclusive annihilation (SIA), one of the cleanest processes for studying hadron production, as it does not require simultaneous knowledge of PDFs. We use the next-to-next-to-leading order (NNLO) datasets NNFF10_PIsum_nnlo for pions ($\pi^{\pm}=\pi^{+}+\pi^{-}$) and NNFF10_KAsum_nnlo for kaons ($K^{\pm}=K^{+}+K^{-}$)[45]. Regarding the specific FFs analyzed, we consider five independent combinations of FFs for each hadron h, $\{D_g^h, D_{b^+}^h, D_{c^+}^h, D_{d^+ + s^+}^h, D_{u^+}^h\}$, where combined FFs are defined as $D_{q+}^h \equiv D_q^h + D_{\bar{q}}^h$ and $D_{d^+ + s^+}^h \equiv D_{d^+}^h + D_{s^+}^h$ for different quarks. The problem domain is set as $\Omega_P = [10^{-2}, 1] \times [1, 10{,}000]$ GeV for (z, Q).

We present the results of one particular FF as an example of the proposed quantum protocol, $D_g^{K^{\pm}}(z, Q)$, which corresponds to the sum of FFs of the gluon g fragmenting into the kaon K^+ and its antiparticle K^-. Samplings of other FFs for $h = K^{\pm}, \pi^{\pm}$ are available in Figs. S1 and S2.

The results in Fig. 3 show that QCPM correctly captures the behavior of the 2-dimensional function $D_g^{K^{\pm}}(z, Q)$ in the region of interest. Figure 3a depicts the target distribution, shown on the training grid. Figure 3b demonstrates the ability of the model to generate samples using the same quantum registers employed during training. Figure 3c, d display the sampling performance when the model utilizes one ($S = 1$) and two ($S = 2$) additional qubits per variable, respectively, where the model makes predictions in untrained regions, and overall shows excellent correspondence to the ground truth.

At this point, it is of particular interest to analyze how correlations between variables affect the training process of QCPMs. The introduction of correlations between variables z and Q in the training circuit is motivated by their analytical relationship through the DGLAP evolution equations. In this work, we use a heuristic-based approach, where we explicitly introduce a correlation circuit $\hat{C}$, as illustrated in Fig. 1b and detailed in Fig. 5, to entangle the registers that load z and Q, which we refer to as $\mathcal{Z}$ and $\mathcal{Q}$. This circuit combines Hadamard gates (H) applied to the first variable with controlled-Z gates (CZ). Our aim is to evaluate the impact of these correlations on QCPM performance and infer correlations.

We compare the accuracy of the models with (w/ CC) and without (w/o CC) correlations between the variables, all conditions being equal (following Fig. 3 and with the same number of qubits, trainable parameters, and optimizer epochs). We use the coefficient of determination R^2 as a metric to evaluate the goodness of the fit for the explicit QCPMs after the training. The results in Fig. 4a demonstrate that QCPM with the correlation circuit $\hat{C}$ consistently outperforms those without, across all the FFs studied. These findings point to the valuable insight from the developed quantum probabilistic models—we can infer (indirectly) the degree of correlations between physical variables by studying the performance of the quantum model and the need for entanglement between $\mathcal{Z}$ and $\mathcal{Q}$. Although typically correlations are inferred from samples and require significant dataset sizes, we suggest quantifying correlations by unraveling the entanglement properties between registers of QCPMs. This can be achieved by various means (entanglement entropy, purity test, and mutual information), and we concentrate on purity measurement[81] as the one that can be readily obtained from the SWAP test[82].

Specifically, to quantify the cross-variable correlations, we define the nonpurity as $C = |1 - \gamma_{\mathcal{Z}}|$, where $\gamma_{\mathcal{Z}}(\rho) \equiv \mathrm{Tr}_{\mathcal{Q}}(\rho_{\mathcal{Z}}^2)$ is the purity of the subsystem $\mathcal{Z}$, with $\mathcal{Q}$ being traced out. We track the value of C during the sampling stage for QCPM that includes the fixed correlation circuit [c.f. Fig. 1c], and show the change of C for a hundred epochs in Fig. 4b [two solid curves for selected processes]. Note that the correlations observed in all the FFs are very similar, as can be seen in Figs. S1 and S2. Therefore, here we focus on two FFs for illustrative purposes. We observe that from a large value of $C_0 \approx 0.7$ during the learning stage, the model settles nonpurity to a small but finite value $C_{100} \approx 2 \times 10^{-2}$ [red and blue lines in Fig. 4b as a guide]. Note that the same values of C are achieved with different circuit initialization chosen to start with uncorrelated registers [dashed curves in Fig. 4b]. While being modest, this value reflects the presence of non-negligible z–Q correlations that are required to reach high accuracy, as shown in Fig. 4a. As such, the analysis of QCPM trained on multidimensional datasets leads to quantifiable measures of cross-correlations to high accuracy.

Conclusions

We introduced QCPMs for learning on multivariate HEP distributions, inferring relevant properties, and performing generative modeling by sampling of quantum circuits. We applied QCPMs to describe FFs being probability distributions that describe hadronization processes during high-energy collisions. The developed models in the Chebyshev basis over a generalized grid are particularly suitable for describing FFs that are known to exhibit a polynomial dependence on the momentum fraction. Additionally, QCPM models can be pretrained to describe the required marginal distributions, and tuned further to capture correlations. Our results on inference show that introducing entanglement between (z, Q) quantum registers significantly enhances training performance, reinforcing the growing body of evidence that quantum correlations can improve model accuracy. Our results on generative modeling demonstrate excellent generalization of models between training points (interpolation), where Chebyshev models enable fine grid sampling, where quantum registers extended by S qubits lead $2^{2(S-1)}$-fold increase in sampling grid size. This work paves the way for further exploration of quantum computing techniques to gain deeper insights into fundamental distributions in particle physics.

Methods

Implementation details of QCPMs

In this section, we provide implementation details of QCPMs to ensure the reproducibility of our results. The training and sampling quantum circuits employed are presented in Fig. 5. The training circuit (Fig. 5a) is composed of two Chebyshev feature map circuits $\hat{\mathcal{U}}_\tau(u)$ and $\hat{\mathcal{U}}_\tau(v)$ for encoding two independent variables, u and $v \in \Omega_C$, in parallel registers, followed by a correlation circuit $\hat{\mathcal{C}}$ and two separate variational Anzätze $\hat{\mathcal{V}}(\theta)$ and $\hat{\mathcal{V}}(\vartheta)$. The correlation circuit generates Bell-like entangled states to make two otherwise independent latent variables correlated for efficient training of FFs. The quantum model $p_Q(u, v) = p_{\theta,\vartheta}(u, v) = \alpha|\langle 0_a \phi 0_a \phi|\big(\hat{\mathrm{I}} \otimes \hat{\mathcal{V}}(\theta) \otimes \hat{\mathrm{I}} \otimes \hat{\mathcal{V}}(\vartheta)\big)\hat{\mathcal{C}}\big(\hat{\mathcal{U}}_\tau(u) \otimes \hat{\mathcal{U}}_\tau(v)\big)|0_a \phi 0_a \phi\rangle|^2 + \beta$ is trained to search for optimal angles (θ^* and ϑ^*) and classical weighting parameters (α_{opt} and β_{opt}), so that a mean squared error loss function is minimized. Because of $p_{\theta^*,\vartheta^*}(u, v) \simeq |\langle u_s u_a u_j v_s v_a v_j|(\hat{\mathcal{U}}_{\mathrm{QChT}}^\dagger \otimes \hat{\mathcal{U}}_{\mathrm{QChT}}^\dagger)\hat{\mathcal{C}}_{\mathrm{ext}}^\dagger(\hat{\mathrm{I}}^{\otimes(S+a)} \otimes \hat{\mathcal{V}}^\dagger(\theta^*) \otimes \hat{\mathrm{I}}^{\otimes(S+a)} \otimes \hat{\mathcal{V}}^\dagger(\vartheta^*))|0_s 0_a \phi 0_s 0_a \phi\rangle|^2$, the sampling circuit (Fig. 5b) consists of the inverse operations of the trained variational Anzätze, followed by an extended inverse correlation circuit $\hat{\mathcal{C}}_{\mathrm{ext}}^\dagger$ and two identical sets of extended inverse quantum Chebyshev transform $\hat{\mathcal{U}}_{\mathrm{QChT}}^\dagger$ circuits. This procedure allows the basis transformation from the Chebyshev to computational basis spaces. In this context, we assume that the problem domain of interest is expressed in terms of the momentum fraction and energy scale $(z, Q) \in \Omega_P$.

It is worth mentioning that Eq. (3) represents the (unnormalized) Chebyshev state. Its normalized version can be prepared by a quantum Chebyshev feature map circuit $\hat{\mathcal{U}}_\tau(u)$ with respect to any continuous input variable u. The Quantum Chebyshev transform $\hat{\mathcal{U}}_{\mathrm{QChT}}$ is a basis transformation between the bitstring and the Chebyshev state at the Chebyshev nodes. Details of both quantum circuits are given in our previous work[75].

The setup we consider for training the QCPM is as follows. We fix the epoch count of the ADAM[83] optimizer to 10^4, the number of Ansatz layers to 3, and the number of qubits per variable to 4. Hence, both circuits have the same number of free parameters (16 for each variable) and require the same training

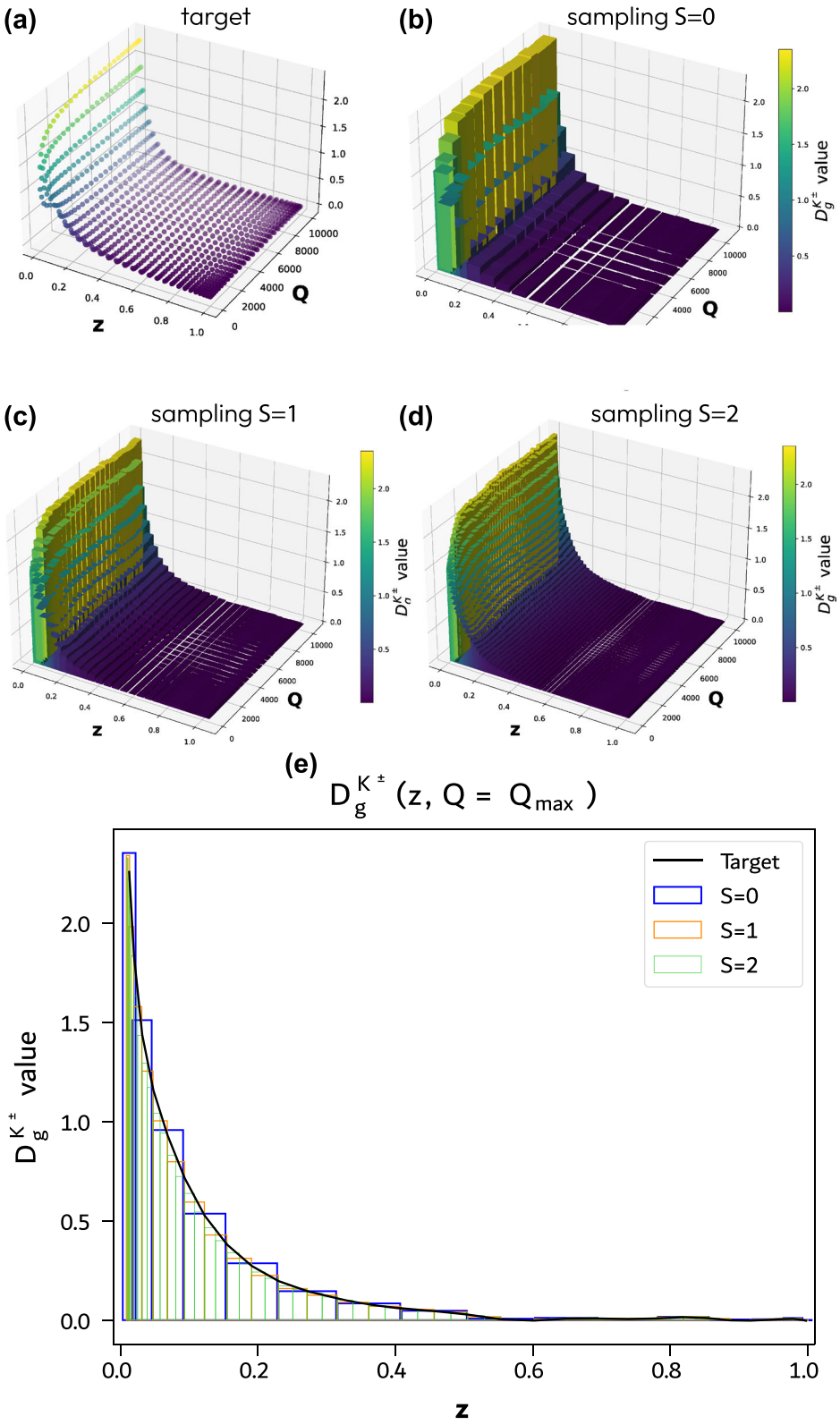

Fig. 3 | Sampling of $D_g^{K^\pm}(z, Q)$, the FF of a gluon fragmenting into kaons. a–d The sampling is performed with $S = 0, 1, 2$ additional qubits for each variable to interpolate in untrained regions. **a** Target distribution $D_g^{K^\pm}(z, Q)$. **b** Samples from trained QCPM of $D_g^{K^\pm}(z, Q)$ with the same number of qubits ($S = 0$). **c**, **d** Samples from $D_g^{K^\pm}(z, Q)$ with extended register ($S = 1, 2$). **e** Contour overlay of the FF projected onto variable Q with $S = 0$, 1, and 2 and the target function. The histogram bins represent an average of the $D_g^{K^\pm}$ value for a range of values of z.

time. We perform a sweep over the learning rate of the ADAM optimizer across the range [0.1, 1.0] and select the model with the best accuracy in learning the functions $D_i^h(z, Q)$. All quantum simulations are performed using Pennylane[84], and the training process is accelerated with JAX[85]. The Ansatz used for the variational quantum circuit is depicted in Fig. 6.

We use the Hardware-Efficient Ansatz (HEA) primarily for its simplicity and compatibility with NISQ devices, which facilitates future implementation on quantum hardware. This choice is not motivated by any particular mathematical structure or claim of optimality for the QPCM. More structured circuit families may offer better performance or reduced resource requirements, especially when considering compilation to fault-tolerant gate sets. However, such optimization is outside the scope of this work.

To provide a baseline reference, we include resource estimates for compiling our HEA circuit into a fault-tolerant Clifford + T gate set using

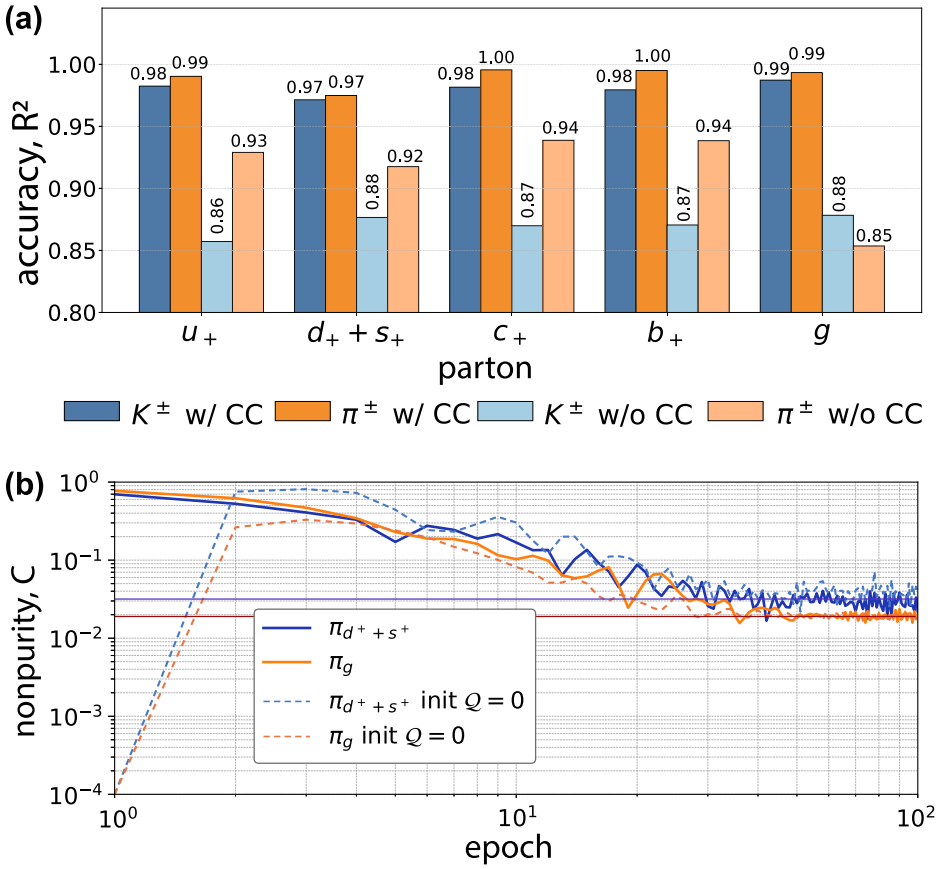

Fig. 4 | Training performance and nonpurity evaluation of the models for different FF. **a** Accuracy (R^2) comparison of QCPMs for learning FFs $D_i^h(z, Q)$ w/ CC and w/o CC correlations between z and Q. **b** Nonpuriry coefficient $C = |1 - \gamma_{\mathcal{Z}}|$ of system $\mathcal{Z}$ for 100 training epochs with the fixed correlation circuit (solid curves, log-log curves). Dashed curves represent C when the system starts in a product state without entanglement between registers (no correlations). Solid curves highlight the values of nonpurity after training.

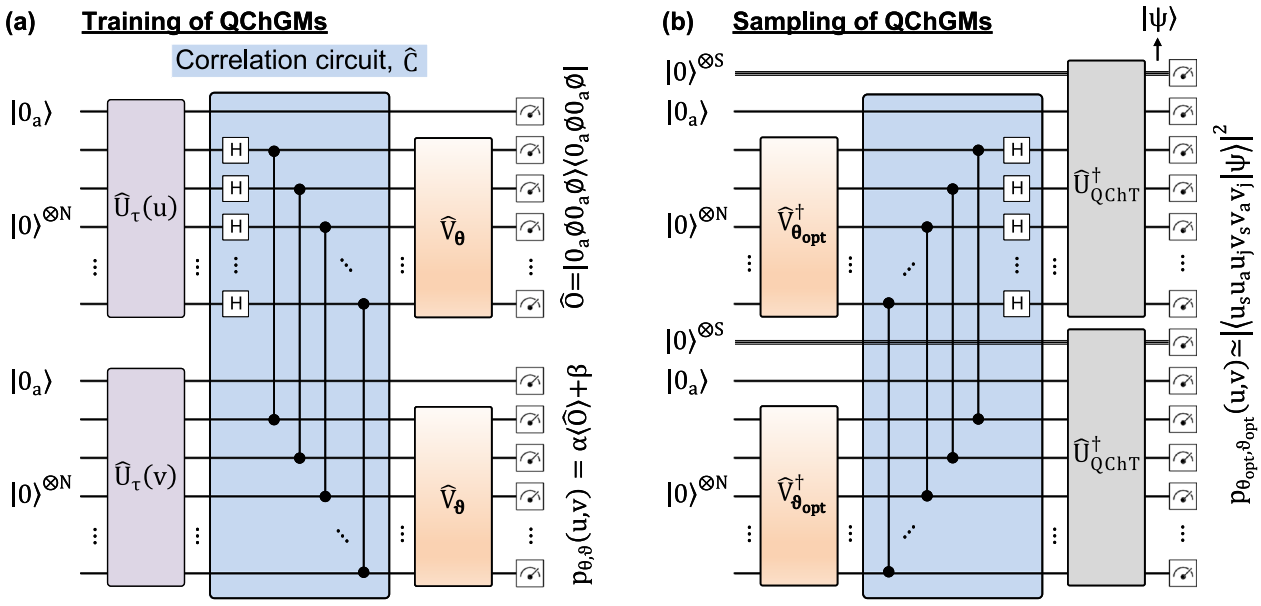

Fig. 5 | Quantum circuits for implementation of QCPMs. a Quantum circuit used to train the multivariate distribution in the QCPM latent space, where a correlation circuit $\hat{C}$ is sandwiched between two identical sets of feature map circuits and variational Ansätze. Measured observable is defined as $\hat{\mathcal{O}} = |0_a\phi 0_a\phi\rangle\langle 0_a\phi 0_a\phi|$, where $|\phi\rangle \equiv |0\rangle^{\otimes N}$. Here, α and β are trainable scaling and bias parameters. **b** Quantum circuit used to sample the multivariate distribution from the trained model, where the inverse versions of the same parameterized circuits are applied with θ^* and ϑ^* being retrieved after the optimization procedure, followed by the inverse versions of the same correlation and two identical sets of quantum Chebyshev transform circuits associated with extended registers of S qubits ($|0\rangle^{\otimes S}$) in parallel, for fine sampling in the computational basis $|u_s u_a u_j v_s v_a v_j\rangle$. The quantum state prior to measurement is denoted as $|\psi\rangle$.

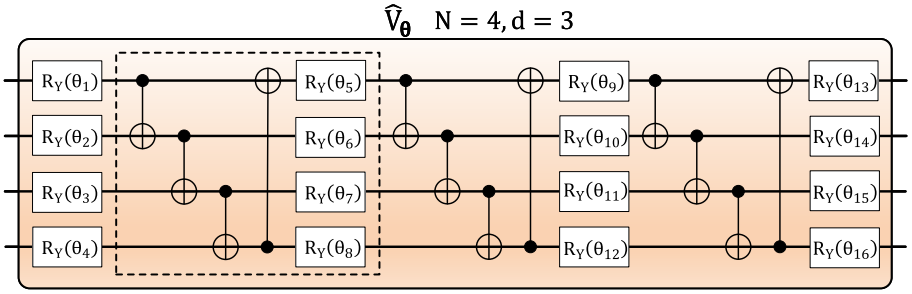

Fig. 6 | Hardware-efficient real-amplitude (HERA) ansatz circuit $\hat{V}_\theta$. The HERA for $N = 4$ qubits with a circuit quantum depth of $d = 3$ is shown. The first block ($d = 1$) is framed by a dashed box. The HERA is composed of $N(d + 1)$ tunable single-qubit R_Y gates and Nd entangling (CNOT) gates. The training parameters are $\{\theta_i\}_{i=1}^{N(d+1)}$.

Table 1 | Resource estimation for transpiled HEA using Ross–Selinger method

ϵ	Total depth	T-count	Total gate count
10^{-2}	415	360	1506
10^{-3}	566	512	2092
10^{-4}	716	654	2681
10^{-5}	841	698	2704

the Ross–Selinger decomposition method implemented in `Pennylane`[86] in Table 1. We define the approximation accuracy by the maximum allowable operator norm error ϵ for the full circuit. The resource estimates include total circuit depth, T-count, and total number of gates. T-count is especially relevant in fault-tolerant quantum computing due to its impact on qubit overhead and magic state distillation costs[87–90]. For context, recent fault-tolerant proposals report that decomposing a single gate may require on the order of 100 T gates[91], indicating that the T counts we obtain are reasonable for the scale of circuits used in this work.

Unlike the structure of GAN represented by the simultaneous training of two neural networks, QCPM separates the training and sampling stages. Because the probability distribution is encoded in the orthogonal Chebyshev basis states provided by the quantum Chebyshev feature map circuit, the model effectively represents a wide range of functions with sufficient training data and learns patterns with a decent number of variational parameters. The number of variational parameters used for training in this work is 32 ($= 2N(d + 1)$) with $N = 4$ and $d = 3$. The trained model generates new data that fall within the trained distribution. In addition to generation and sampling, the QCPM efficiently encodes probability distributions into quantum states, making it a valuable tool for a wider range of quantum algorithms and applications.

Data availability

The raw data that support the findings of this study are available at https://github.com/CERN-IT-INNOVATION/QChPM.

Code availability

The code and the dataset used in this work are open source and available at https://github.com/CERN-IT-INNOVATION/QChPM.

Received: 7 April 2025; Accepted: 8 October 2025;
Published online: 19 November 2025

References

1. Wang, H. et al. Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).
2. Guest, D., Cranmer, K. & Whiteson, D. Deep learning and its application to LHC physics. *Annu. Rev. Nucl. Part. Sci.* **68**, 161–181 (2018).
3. Delgado, A. et al. Quantum computing for data analysis in high-energy physics. Preprint at https://arxiv.org/abs/2203.08805 (2022).
4. Di Meglio, A. et al. Quantum computing for high-energy physics: state of the art and challenges. *PRX Quantum* **5**, 037001 (2024).
5. Rodrigo, G. Quantum algorithms in particle physics. *Acta Phys. Pol. Supp.* **17**, 2–A14 (2024).
6. Biamonte, J. et al. Quantum machine learning. *Nature* **549**, 195–202 (2017).
7. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
8. Du, Y. et al. Quantum machine learning: a hands-on tutorial for machine learning practitioners and researchers. Preprint at https://arxiv.org/abs/2502.01146 (2025).
9. Wei, A. Y., Naik, P., Harrow, A. W. & Thaler, J. Quantum algorithms for jet clustering. *Phys. Rev. D.* **101**, 094015 (2020).
10. Delgado, A. & Thaler, J. Quantum annealing for jet clustering with thrust. *Phys. Rev. D.* **106**, 094016 (2022).
11. Martínez de Lejarza, J. J., Cieri, L. & Rodrigo, G. Quantum clustering and jet reconstruction at the LHC. *Phys. Rev. D.* **106**, 036021 (2022).
12. de Lejarza, J. J. M., Cieri, L. & Rodrigo, G. Quantum jet clustering with LHC simulated data. *PoS* **ICHEP2022**, 241 (2022).
13. de Lejarza, J. J. M., Grossi, M., Cieri, L. & Rodrigo, G. Quantum Fourier iterative amplitude estimation. In *2023 International Conference on Quantum Computing and Engineering* (IEEE, 2023).
14. Agliardi, G., Grossi, M., Pellen, M. & Prati, E. Quantum integration of elementary particle processes. *Phys. Lett. B* **832**, 137228 (2022).
15. Williams, I. & Pellen, M. A general approach to quantum integration of cross sections in high-energy physics. *Quantum Sci. Technol.* **10**, 045017 (2025).
16. Belis, V. et al. Quantum anomaly detection in the latent space of proton collision events at the LHC. *Commun. Phys.* **7**, 334 (2024).
17. Schuhmacher, J. et al. Unravelling physics beyond the standard model with classical and quantum anomaly detection. *Mach. Learn. Sci. Technol.* **4**, 045031 (2023).
18. Kyriienko, O. & Magnusson, E. B. Unsupervised quantum machine learning for fraud detection. Preprint at https://arxiv.org/abs/2208.01203 (2022).
19. Belis, V. et al. Guided quantum compression for high dimensional data classification. *Mach. Learn. Sci. Tech.* **5**, 035010 (2024).
20. Belis, V. et al. Higgs analysis with quantum classifiers. *EPJ Web Conf.* **251**, 03070 (2021).
21. Ramírez-Uribe, S., Rentería-Olivo, A. E., Rodrigo, G., Sborlini, G. F. R. & Vale Silva, L. Quantum algorithm for Feynman loop integrals. *JHEP* **05**, 100 (2022).
22. Clemente, G. et al. Variational quantum eigensolver for causal loop Feynman diagrams and directed acyclic graphs. *Phys. Rev. D.* **108**, 096035 (2023).
23. Ramírez-Uribe, S., Rentería-Olivo, A. E. & Rodrigo, G. Quantum querying based on multicontrolled Toffoli gates for causal Feynman loop configurations and directed acyclic graphs. Preprint at https://arxiv.org/abs/2404.03544 (2024).
24. de Lejarza, J. J. M., Cieri, L., Grossi, M., Vallecorsa, S. & Rodrigo, G. Loop Feynman integration on a quantum computer. *Phys. Rev. D.* **110**, 074031 (2024).

25. de Lejarza, J. J. M., Rentería-Estrada, D. F., Grossi, M. & Rodrigo, G. Quantum integration of decay rates at second order in perturbation theory. *Quantum Sci. Technol.* **10**, 025026 (2025).
26. Pyretzidis, K., de Lejarza, J. J. M. & Rodrigo, G. Unlocking multi-dimensional integration with quantum adaptive importance sampling. Preprint at https://arxiv.org/abs/2506.19965 (2025).
27. Romero, J. & Aspuru-Guzik, A. Variational quantum generators: generative adversarial quantum machine learning for continuous distributions. *Adv. Quantum Technol.* **4**, 2000003 (2021).
28. Paine, A. E., Elfving, V. E. & Kyriienko, O. Quantum quantile mechanics: solving stochastic differential equations for generating time-series. *Adv. Quantum Technol.* **6**, 2300065 (2023).
29. Zhu, E. Y. et al. Generative quantum learning of joint probability distribution functions. *Phys. Rev. Res.* **4**, 043092 (2022).
30. Kyriienko, O., Paine, A. E. & Elfving, V. E. Protocols for trainable and differentiable quantum generative modeling. *Phys. Rev. Res.* **6**, 033291 (2024).
31. Kasture, S., Kyriienko, O. & Elfving, V. E. Protocols for classically training quantum generative models on probability distributions. *Phys. Rev. A* **108**, 042406 (2023).
32. Rudolph, M. S. et al. Trainability barriers and opportunities in quantum generative modeling. *npj Quantum Inf.* **10**, 116 (2024).
33. Wu, H.-Y., Elfving, V. E. & Kyriienko, O. Multidimensional quantum generative modeling by quantum Hartley transform. *Adv. Quantum Technol.* https://doi.org/10.1002/qute.202400337 (2024).
34. de Florian, D., Sassot, R. & Stratmann, M. Global analysis of fragmentation functions for protons and charged hadrons. *Phys. Rev. D.* **76**, 074033 (2007).
35. Aidala, C. A., Ellinghaus, F., Sassot, R., Seele, J. P. & Stratmann, M. Global analysis of fragmentation functions for Eta mesons. *Phys. Rev. D.* **83**, 034002 (2011).
36. de Florian, D., Sassot, R. & Stratmann, M. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev. D.* **75**, 114010 (2007).
37. Rojo, J. et al. The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II. *J. Phys. G* **42**, 103103 (2015).
38. Albino, S. et al. Parton fragmentation in the vacuum and in the medium. Preprint at https://arxiv.org/abs/0804.2021 (2008).
39. Glück, M., Reya, E. & Vogt, A. Dynamical parton distributions revisited. *Eur. Phys. J. C.* **5**, 461–470 (1998).
40. Kretzer, S. Fragmentation functions from flavor inclusive and flavor tagged e+ e- annihilations. *Phys. Rev. D.* **62**, 054001 (2000).
41. Kniehl, B. A., Kramer, G. & Potter, B. Fragmentation functions for pions, kaons, and protons at next-to-leading order. *Nucl. Phys. B* **582**, 514–536 (2000).
42. Hirai, M., Kumano, S., Nagai, T. H. & Sudoh, K. Determination of fragmentation functions and their uncertainties. *Phys. Rev. D.* **75**, 094009 (2007).
43. Nocera, E. R. Towards a neural network determination of charged pion fragmentation functions. In *22nd International Symposium on Spin Physics* (2017).
44. Bertone, V., Carrazza, S., Nocera, E. R., Hartland, N. P. & Rojo, J. Towards a neural network determination of pion fragmentation functions. In *Proc Parton Radiation and Fragmentation from LHC to FCC-ee* 19–25 (CERN, Geneva, 22–23, 2016); http://inspirehep.net/record/1512989/files/1512294_19-25.pdf.
45. Bertone, V., Carrazza, S., Hartland, N. P., Nocera, E. R. & Rojo, J. A determination of the fragmentation functions of pions, kaons, and protons with faithful uncertainties. *Eur. Phys. J. C.* **77**, 516 (2017).
46. Soleymaninia, M., Hashamipour, H., Khanpour, H. & Spiesberger, H. Fragmentation functions for Ξ^-/Ξ^-+ using neural networks. *Nucl. Phys. A* **1029**, 122564 (2023).
47. Ball, R. D. et al. An open-source machine learning framework for global analyses of parton distributions. *Eur. Phys. J. C.* **81**, 958 (2021).
48. Gribov, V. N. & Lipatov, L. N. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.* **15**, 438–450 (1972).
49. Lipatov, L. N. The parton model and perturbation theory. *Yad. Fiz.* **20**, 181–198 (1974).
50. Altarelli, G. & Parisi, G. Asymptotic freedom in parton language. *Nucl. Phys. B* **126**, 298–318 (1977).
51. Dokshitzer, Y. L. Calculation of the structure functions for deep inelastic scattering and e+ e− annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP* **46**, 641–653 (1977).
52. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2011); https://www.amazon.com/Quantum-Computation-Information-10th-Anniversary/dp/1107002176?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=1107002176.
53. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
54. Bond-Taylor, S., Leach, A., Long, Y. & Willcocks, C. G. Deep generative modelling: a comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 7327–7347 (2022).
55. Lloyd, S. & Weedbrook, C. Quantum generative adversarial learning. *Phys. Rev. Lett.* **121**, 040502 (2018).
56. Dallaire-Demers, P.-L. & Killoran, N. Quantum generative adversarial networks. *Phys. Rev. A* **98**, 012324 (2018).
57. Zoufal, C., Lucchi, A. & Woerner, S. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Inf.* **5**, 103 (2019).
58. Huang, H.-L. et al. Experimental quantum generative adversarial networks for image generation. *Phys. Rev. Appl.* **16**, 024051 (2021).
59. Amin, M. H., Andriyash, E., Rolfe, J., Kulchytskyy, B. & Melko, R. Quantum Boltzmann machine. *Phys. Rev. X* **8**, 021050 (2018).
60. Zoufal, C., Lucchi, A. & Woerner, S. Variational quantum boltzmann machines. *Quantum Machine Intell.* https://doi.org/10.1007/s42484-020-00033-7 (2021).
61. Coopmans, L. & Benedetti, M. On the sample complexity of quantum Boltzmann machine learning. *Commun. Phys*. **7**, 274 (2024).
62. Liu, J.-G. & Wang, L. Differentiable learning of quantum circuit born machines. *Phys. Rev. A* **98**, 062324 (2018).
63. Benedetti, M. et al. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Inf.* **5**, 45 (2019).
64. Benedetti, M., Lloyd, E., Sack, S. & Fiorentini, M. Erratum: parameterized quantum circuits as machine learning models (2019 quant. sci. tech. 4 043001). *Quantum Sci. Technol.* **5**, 019601 (2019).
65. Coyle, B., Mills, D., Danos, V. & Kashefi, E. The born supremacy: quantum advantage and training of an Ising born machine. *npj Quantum Inf.* **6**, 60 (2020).
66. Kiss, O., Grossi, M., Kajomovitz, E. & Vallecorsa, S. Conditional born machine for Monte Carlo event generation. *Phys. Rev. A* **106**, 022612 (2022).
67. Hibat-Allah, M., Mauri, M., Carrasquilla, J. & Perdomo-Ortiz, A. A framework for demonstrating practical quantum advantage: comparing quantum against classical generative models. *Commun. Phys.* **7**, 68 (2024).
68. Delgado, A., Rios, F. & Hamilton, K. E. Identifying overparameterization in Quantum Circuit Born machines. Preprint at https://arxiv.org/abs/2307.03292 (2023).
69. Delgado, A., Venegas-Vargas, D., Huynh, A. & Carroll, K. Towards designing scalable quantum-enhanced generative networks for neutrino physics experiments with liquid argon time projection chambers. Preprint at https://arxiv.org/abs/2410.12650 (2024).
70. Bermot, E. et al. Quantum generative adversarial networks for anomaly detection in high energy physics. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01, 331–341 (IEEE, 2023).

71. Tüysüz, C. et al. Learning to generate high-dimensional distributions with low-dimensional quantum Boltzmann machines. Preprint at https://arxiv.org/abs/2410.16363 (2024).
72. Pérez-Salinas, A., Cruz-Martinez, J., Alhajri, A. A. & Carrazza, S. Determining the proton content with a quantum computer. *Phys. Rev. D.* **103**, 034027 (2021).
73. Cruz-Martinez, J. M., Robbiati, M. & Carrazza, S. Multi-variable integration with a variational quantum circuit. *Quantum Sci. Technol.* **9**, 035053 (2024).
74. Kang, Z.-B., Moran, N., Nguyen, P. & Qian, W. Partonic distribution functions and amplitudes using tensor network methods. Preprint at https://arxiv.org/abs/2501.09738 (2025).
75. Williams, C. A., Paine, A. E., Wu, H.-Y., Elfving, V. E. & Kyriienko, O. Quantum Chebyshev transform: mapping, embedding, learning and sampling distributions. Preprint at https://arxiv.org/abs/2306.17026 (2023).
76. Schuld, M., Sweke, R. & Meyer, J. J. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Phys. Rev. A* **103**, 032430 (2021).
77. Kyriienko, O. & Elfving, V. E. Generalized quantum circuit differentiation rules. *Phys. Rev. A* **104**, 052417 (2021).
78. Kyriienko, O., Paine, A. E. & Elfving, V. E. Solving nonlinear differential equations with differentiable quantum circuits. *Phys. Rev. A* **103**, 052416 (2021).
79. Buckley, A. et al. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J. C.* **75**, 132 (2015).
80. Buckley, A. et al. LHAPDF: main page (LHAPDF, accessed 16 Jan 2025, 2015); https://www.lhapdf.org.
81. Scali, S., Umeano, C. & Kyriienko, O. The topology of data hides in quantum thermal states. *APL Quantum* **1**, 036106 (2024).
82. Paine, A. E., Elfving, V. E. & Kyriienko, O. Quantum kernel methods for solving regression problems and differential equations. *Phys. Rev. A* **107**, 032428 (2023).
83. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980 (2017).
84. Bergholm, V. et al. PennyLane: automatic differentiation of hybrid quantum-classical computations. Preprint at https://arxiv.org/abs/1811.04968 (2018).
85. Bradbury, J. et al. JAX: composable transformations of Python +NumPy programs. Github http://github.com/jax-ml/jax (2018).
86. Xanadu. qml.ops.rs_decomposition—PennyLane—docs.pennylane.ai (accessed 04 August 2025); https://docs.pennylane.ai/en/stable/code/api/pennylane.ops.rs_decomposition.html.
87. Gosset, D., Kliuchnikov, V., Mosca, M. & Russo, V. An algorithm for the t-count. Preprint at https://arxiv.org/abs/1308.4134 (2013).
88. Postler, L. et al. Demonstration of fault-tolerant universal quantum gate operations. *Nature* **605**, 675–680 (2022).
89. Gheorghiu, V., Mosca, M. & Mukhopadhyay, P. T-count and T-depth of any multi-qubit unitary. *npj Quantum Inf.* **8**, 141 (2022).
90. Bravyi, S. & Kitaev, A. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*. https://doi.org/10.1103/PhysRevA.71.022316 (2005).
91. Loaiza, I. et al. Simulating near-infrared spectroscopy on a quantum computer for enhanced chemical detection. Preprint at https://arxiv.org/abs/2504.10602 (2025).

Acknowledgements

The authors thank Antonio A. Gentile for useful discussions and suggestions on the manuscript. This work is supported by the Spanish Government (Agencia Estatal de Investigación MCIU/AEI/10.13039/501100011033), grant no. PID2023-146220NB-I00, no. PID2020-114473GB-I00 and no. CEX2023-001292-S, and Generalitat Valenciana grants no. ASFAE/2022/009 (Planes Complementarios de I + D + i, NextGenerationEU). This work is also supported by the Ministry of Economic Affairs and Digital Transformation of the Spanish Government and NextGenerationEU through the Quantum Spain project and by the CSIC Interdisciplinary Thematic Platform on Quantum Technologies (PTI-QTEP+). J.J.M.d.L. is supported by the Generalitat Valenciana grant no. ACIF/2021/219. H.-Y.W. and O.K. acknowledge the funding from the UK EPSRC award under the Agreements EP/Z53318X/1 (QCi3 Hub) and EP/Y005090/1. M.G. is supported by CERN through the CERN Quantum Technology Initiative.

Author contributions

J.J.d.M.L., G.R., and M.G. conceived the initial idea of analyzing FFs with quantum Chebyshev models. J.J.d.M.L. generated the data and executed the simulations. O.K. proposed the probabilistic modeling regime for inferring correlations and the multivariate architecture. H.-Y.W. conceived the idea of applying shifted Chebyshev polynomials to the quantum Chebyshev models and provided the initial demo codes. All authors discussed the results and contributed to the final manuscript.

Competing interests

Michele Grossi is a Guest Editor for Communications Physics, but was not involved in the editorial review of, or the decision to publish this article. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-025-02361-1.

Correspondence and requests for materials should be addressed to Jorge J. Martínez de Lejarza.

Peer review information *Communications Physics* thanks Jay Chan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.