

# Multidimensional Quantum Generative Modeling by Quantum Hartley Transform

Hsin-Yu Wu,\* Vincent E. Elfving, and Oleksandr Kyriienko

An approach for building quantum models based on the exponentially growing orthonormal basis of Hartley kernel functions is developed. First, a differentiable Hartley feature map parameterized by real-valued argument that enables quantum models suitable for solving stochastic differential equations and regression problems is designed. Unlike the naturally complex Fourier encoding, the proposed Hartley feature map circuit leads to quantum states with real-valued amplitudes, introducing an inductive bias and natural regularization. Next, a quantum Hartley transform circuit is proposed as a map between computational and Hartley basis. The developed paradigm is applied to generative modeling from solutions of stochastic differential equations, and utilize the quantum Hartley transform for fine sampling from parameterized distributions through an extended register. Finally, the capability of multivariate quantum generative modeling is demonstrated for both correlated and uncorrelated distributions. As a result, the developed quantum Hartley-based generative models (QHGMs) offer a distinct quantum approach to generative AI at increasing scale.

quantum Fourier transform (QFT).<sup>[11–14]</sup> The QFT is the implementation of the classical discrete Fourier transform (DFT) on a quantum circuit (QC) and it can be derived in an equivalent way from a fast Fourier transform based on the Danielson–Lanczos lemma.<sup>[15–17]</sup> Using the phase kick-back trick, one can show that both resulting QFT circuits are identical. As the vast majority of applications is concerned with the processing of real-valued datasets, the Hermitian-symmetric QFT manifests itself that there is redundancy in the spectral expansion of a real-valued signal. In addition, one complex multiplication requires the application of four real multiplications and three additions/subtractions, while one real multiplication just needs two multiplications and one addition/subtraction. Consequently, the use of the Fourier-related algorithms capable of naturally performing real-valued unitary transformations between the real and reciprocal spaces is highly

desirable. Namely, it adds an inductive bias, thus reducing resources required for building quantum models and generating quantum probability distributions.

Quantum machine learning (QML) is a burgeoning interdisciplinary field that integrates quantum computing with machine learning. Usually, it refers to the use of variational quantum algorithms to undertake classical learning tasks and solve practically relevant problems.<sup>[18–26]</sup> Motivated by the great success of their classical counterparts in machine learning,<sup>[27–31]</sup> various protocols for quantum generative modeling (QGM) have recently been developed to exploit parameterized QCs for modeling of the probability distribution of a given dataset, including quantum Boltzmann machines (QBMs), quantum generative adversarial networks (QGANs) and quantum generative diffusion models (QGDMS). Once successfully trained, these models can generate synthetic samples closely resembling the data distribution. QBMs belong to energy-based models that use a parameterized QC as a Hamiltonian operator to prepare quantum Gibbs state whose sampling behavior reflects the probability distribution of given training data.<sup>[32,33]</sup> Due to the nature the minmax loss function, QGANs rely on the simultaneous training of two separate parameterized QCs in alternating optimization steps to reach a Nash equilibrium where the quantum generator can produce samples from the trained distribution.<sup>[34–36]</sup> Inspired by non-equilibrium thermodynamics, QGDMS, comprising two separable phases - the forward diffusion with a specific noise schedule

## 1. Introduction

Quantum computing offers speedup in solving certain problems, with a promise of outperforming classical solvers thanks to natively quantum effects of entanglement and superposition.<sup>[1]</sup> Quantum algorithms that take advantage of these properties cover various application areas, ranging from cryptography<sup>[2]</sup> and quantum simulation,<sup>[3,4]</sup> to linear equation solvers,<sup>[5–7]</sup> and optimization.<sup>[8–10]</sup> Most of the protocols mentioned above rely on subroutines that involve the quantum phase estimation and the

H.-Y. Wu, O. Kyriienko  
Department of Physics and Astronomy  
University of Exeter  
Stocker Road, Exeter EX4 4QL, UK  
E-mail: [h.wu@exeter.ac.uk](mailto:h.wu@exeter.ac.uk)  
H.-Y. Wu, V. E. Elfving, O. Kyriienko  
PASQAL  
7 Rue Léonard de Vinci, Massy 91300, France

The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/qute.202400337>

© 2024 The Author(s). Advanced Quantum Technologies published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/qute.202400337

and the backward denoising with a time-parameterized QC, have been demonstrated to offer faster convergence, better scalability and higher fidelity than QGANs.<sup>[37–39]</sup>

Without a classical analogue, quantum circuit Born machines (QCBMs) are quantum-inspired generative models and this model family leverages the inherent probabilistic nature of quantum measurement<sup>[40]</sup> and implicitly treats data represented by binary strings as an observable (thus building an implicit model<sup>[41]</sup>). As soon as QCBMs are successfully trained, one can directly produce samples from the parameterized probability distribution through projective measurements using the same circuit.<sup>[42–47]</sup> By contrast, the explicit models allow for separation of the training and sampling stages of QGM<sup>[48,49]</sup> and rely on a differentiable feature map encoding of continuous or discrete distributions of data at the training stage.<sup>[50]</sup> A typical workflow for the quantum explicit models starts with encoding classical input data of the form  $\mathbf{x} = \{x_0, \dots, x_{2^N-1}\}$  to quantum states via a  $N$ -qubit quantum feature map circuit that acts on a zero product state  $|\phi\rangle \equiv |0\rangle^{\otimes N}$ . Specifically, a feature map is a unitary operator that takes an input argument  $\mathbf{x}$  and maps it to a distinct quantum state living in a  $2^N$  Hilbert space,  $\mathbf{x} \mapsto |\varphi(\mathbf{x})\rangle = \hat{U}_\varphi(\mathbf{x})|\phi\rangle$ . One feature map that associates an input feature with a quantum state in the phase space is the Fourier (phase) feature map.<sup>[48]</sup> It is formed by an initial layer of Hadamard gates on each qubit followed by a layer of scaled phase shift gates,  $P_l(\mathbf{x}) = \text{diag}\{1, \exp(i2\pi\mathbf{x}/2^l)\}$  applied to the qubit index  $l \in [1, \dots, N]$ . The set of the output states  $\{|\varphi(x_j)\rangle\}_{j=0}^{2^N-1}$ , evaluated at  $\{x_j\}_{j=0}^{2^N-1}$  representing consecutive integers in the range of  $[0, 2^N - 1]$  corresponding to length- $N$  binary strings, forms a complete orthonormal Fourier basis. These Fourier basis states can be mapped to the set of computational basis through an inverse quantum Fourier transform (QFT) circuit,  $\{|\varphi(x_j)\rangle\}_{j=0}^{2^N-1} = \hat{U}_{\text{QFT}}^\dagger \{|\phi(x_j)\rangle\}_{j=0}^{2^N-1}$ . Unlike other quantum encoding techniques, such as amplitude and basis embeddings,<sup>[51,52]</sup>  $\hat{U}_\varphi(\mathbf{x})$  can be differentiated with respect to a continuous variable  $\mathbf{x} \in \mathbb{R}_{2^N-1}$ , allowing solving differential equations modeled in the explicit form. Different from Fourier feature encoding, the Chebyshev feature map has recently gathered attention as it generates quantum states with amplitudes proportional to Chebyshev polynomials of the first kind, forming an orthonormal Chebyshev basis on non-equidistant nodes.<sup>[53,54]</sup> Current classical generative architecture concentrates on training large-scale models of rapidly increasing complexity. Quantum computing may offer a powerful alternative platform for multidimensional generative modeling with a potential ability to prepare correlated distributions.<sup>[48,55]</sup> However, it remains an open problem how to build quantum models that can efficiently express, learn, and sample from a high dimensional probability distribution, enabling quantum generative artificial intelligence (AI) in a large scale.

In this paper, we propose an orthonormal Hartley feature map that enables explicit quantum model building in the Hartley basis.<sup>[56]</sup> Unlike DFT, the discrete Hartley transform (DHT) maps a real input to a real output and has the convenient property of being its own inverse<sup>[57]</sup> as well, particularly suitable for physics-informed quantum machine learning.<sup>[54]</sup> The DHT has been shown to offer computational advantages over DFT in applications of power spectrum and convolution computations.<sup>[58,59]</sup>

The orthonormal Hartley feature map as a quantum circuit parameterized by a continuous variable  $\mathbf{x}$  prepares quantum states with amplitudes proportional to the so-called Hartley kernel function, which facilitates building and training quantum models in the real-valued Hartley (latent) space with an exponentially large basis set and allows for model differentiation. We apply the developed quantum protocols for learning probability distributions motivated by financially-relevant processes. We demonstrate the efficient sampling of the Hartley-based quantum model by mapping the Hartley basis to the computational basis via a quantum Hartley transform (QHT) circuit. We then employ the developed tools to solve differential equations and compare the results obtained from the different quantum models constructed by Fourier and Hartley bases. Finally, we proceed to extend and implement the proposed strategies to multivariate encoding and sampling for multidimensional quantum generative modeling.

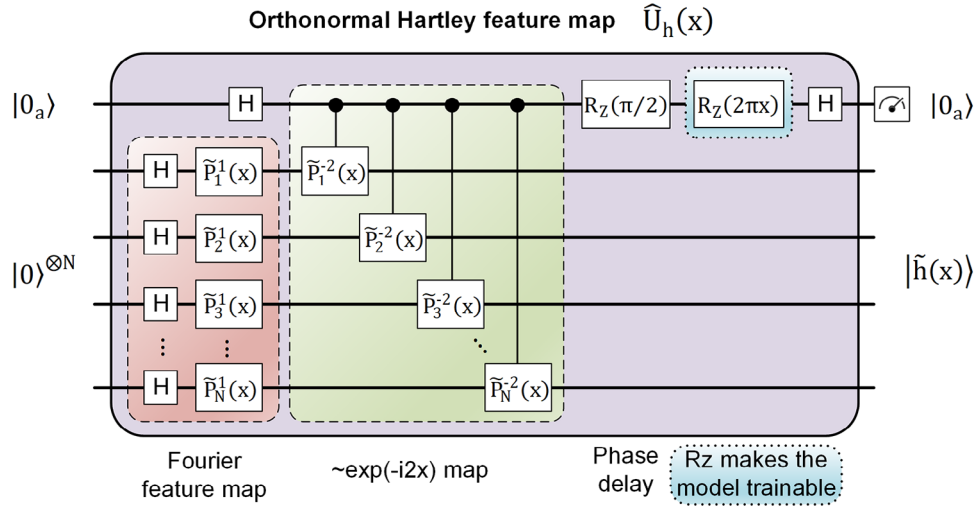
## 2. Results and Discussion

### 2.1. Orthonormal Hartley Feature map Circuit

We design Hartley feature maps to facilitate the learning process for quantum probability distribution, and make models trainable, while allowing easy sampling from the trained model via a unitary transformation to the computational basis. The computational basis states refer to orthonormal states  $\{|\mathbf{x}_j\rangle\}_{j=0}^{2^N-1}$  such that overlaps equate to the Kronecker delta function,  $\langle \mathbf{x}_{j'} | \mathbf{x}_j \rangle = \delta_{j'j}$ . In the case where the initial input states are  $|\phi\rangle$ ,  $|\mathbf{x}_j\rangle$  can be easily generated by applying a Pauli  $\hat{X}$  gate to the  $l^{\text{th}}$  qubit if the  $l^{\text{th}}$  classical bit is 1, often referred to as the basis encoding.<sup>[60]</sup> Specifically, we want to generate a  $N$ -qubit quantum state  $|h(\mathbf{x})\rangle$  with amplitude proportional to an equally weighted sum of  $\mathbf{x}$ -dependent cosine and sine wave functions, namely the Hartley kernel  $\text{cas}(2\pi k\mathbf{x}/2^N) \equiv \cos(2\pi k\mathbf{x}/2^N) + \sin(2\pi k\mathbf{x}/2^N)$ . This state can be written as  $|h(\mathbf{x})\rangle = 2^{-N/2} \sum_{k=0}^{2^N-1} \text{cas}(2\pi k\mathbf{x}/2^N) |k\rangle$ , where  $\{|k\rangle\}$  are  $2^N$  computational basis states. For  $x_j \in [0, \dots, 2^N - 1]$ , the  $\text{cas}(\cdot)$  function satisfies the following orthogonality conditions,

$$\sum_{j=0}^{2^N-1} \text{cas}(2\pi kx_j/2^N) \text{cas}(2\pi \ell x_j/2^N) = \begin{cases} 2^N & k = \ell \\ 0 & k \neq \ell \end{cases} \quad (1)$$

As a consequence, the states  $|h(\mathbf{x})\rangle$  are orthonormal on the integer grid points. Namely, the set of Hartley states  $\{|h(\mathbf{x}_j)\rangle\}_{j=0}^{2^N-1}$  satisfies  $\langle h(\mathbf{x}_{j'}) | h(\mathbf{x}_j) \rangle = \delta_{j'j}$  with  $\delta_{j'j}$  being the Kronecker delta function. We note that the states  $|h(\mathbf{x})\rangle$  also fulfill this orthonormal condition for half-integer points,  $\{x_{(j+1/2)}\} \forall j \in [0, 2^N - 1]$ . Since the  $\text{cas}(\cdot)$  function can be alternatively expressed as a delayed cosine function,  $\text{cas}(\mathbf{x}) = \sqrt{2} \cos(\mathbf{x} - \pi/4)$ , Hartley states can be prepared using a combination of exponents  $\cos(\mathbf{x}) = [\exp(i\mathbf{x}) + \exp(-i\mathbf{x})]/2$  for some scaled variable  $\mathbf{x}$ , where each amplitude is embedded via the phase feature map.<sup>[48]</sup> Therefore, the implementation of Hartley feature map circuit  $\hat{U}_h(\mathbf{x})$  that prepares normalized Hartley state for any continuous variable  $\mathbf{x}$  can be achieved by the linear combination of unitary (LCU) approach,<sup>[61]</sup> as shown in **Figure 1**. The circuit begins with a Hadamard gate acting on the ancilla register, being the most significant bit, with



**Figure 1.** Quantum Hartley feature map. Quantum Hartley feature map  $\hat{U}_h(x)$  circuit creates a Hartley state via  $x$ -parameterized isometry — a  $N$ -qubit Fourier feature map followed by controlled phase shift gates to embed complex exponents and a constant  $R_z$  gate to provide a global phase delay. A  $x$ -dependent  $R_z$  gate is acting on the ancilla and sandwiched between two Hadamard gates to ensure favorable mid-point behavior while retaining the real amplitude of  $|\tilde{h}(x)\rangle$  after the ancillary measurement yields  $|0\rangle$  outcome. Here, scaled single-qubit phase shift gate is defined as  $\tilde{P}_l^m(x) = \text{diag}\{1, \exp(im2\pi x/2^l)\}$ , where  $l \in [1, \dots, N]$  is the qubit index and  $m$  takes values of 1 and  $-2$ , for any continuous variable  $x \in \mathbb{R}_{2^N-1}$ .

a Fourier feature map applied to  $N$  qubits to distribute the scaled relative phases  $\exp(ix)$  to all of the states in Fourier space, and then a series of controlled phase gates are applied to contribute relative phases based on  $\exp(-ix) = \exp(ix) \exp(-i2x)$ . A constant  $R_z(\pi/2)$  gate is utilized to produce a  $\pi/4$  phase delay in the cosine argument. A  $x$ -dependent  $R_z(2\pi x)$  gate is appended to the ancillary qubit to minimize the contributions of the quantum state associated with non-integer points on the state overlap (see Appendix A.1 for detailed discussion), followed by a Hadamard gate. Under the condition that the ancilla register collapses to  $|0\rangle$  outcome, the quantum state prepared by  $\hat{U}_h(x)$  reads  $|\tilde{h}(x)\rangle = |h(x)\rangle/\mathcal{N}(x)$ , where the unnormalized state  $|h(x)\rangle$  is expressed as

$$|h(x)\rangle = \frac{1}{2^{N/2}} \sum_{k=0}^{2^N-1} \text{cas}[(2\pi k/2^N - \pi)x] |k\rangle \quad (2)$$

and  $\mathcal{N}(x) = \sqrt{\langle h(x)|h(x)\rangle} = \sqrt{1 - \sin(2\pi x)/2^N}$  is  $x$ -dependent when not evaluated at the half-integer and integer grid points.

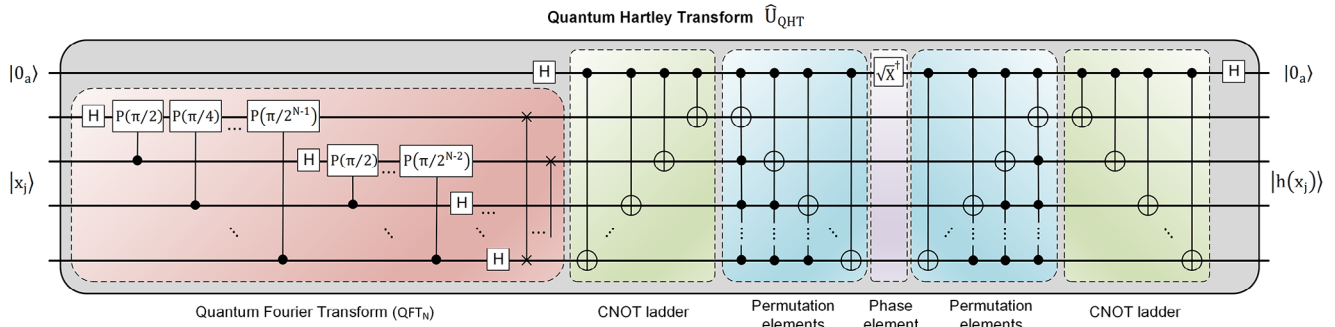
## 2.2. Quantum Hartley Transform Circuit

For sampling purposes, we need to develop a corresponding transformation circuit that enables the mapping between Hartley states  $\{|h(x_j)\rangle\}_{j=0}^{2^N-1}$  and the computational states  $\{|x_j\rangle\}_{j=0}^{2^N-1}$  (and reverse). This unitary transformation reads  $|h(x_j)\rangle = (-1)^j \hat{U}_{\text{QHT}} |x_j\rangle$ , where  $\hat{U}_{\text{QHT}}$  represents quantum Hartley transform (QHT) and the sign flips for odd integers results from the introduction of  $x$ -dependent  $R_z$  gate in the Hartley feature map circuit. For general sampling, the effect of this phase modulation on projective measurements can be ignored. As a result, the matrix representation of QHT can be expressed as  $\hat{U}_{\text{QHT}} = \sum_{j=0}^{2^N-1} |h(x_j)\rangle \langle x_j|$ . We note that QHT is the quan-

tum analogue of DHT. Namely, the vector amplitude of  $|h(x_j)\rangle$  corresponds to the  $(j+1)^{\text{th}}$  column of the DHT matrix defined as  $\text{DHT}_N := 2^{-N/2} \left\{ \text{cas}(2\pi kj/2^N) \right\} \forall k, j \in [0, 2^N - 1]$ . It is worth mentioning that  $\hat{U}_{\text{QHT}}$  is an involutory matrix  $\hat{U}_{\text{QHT}}^2 = \hat{U}_{\text{QHT}}^{2\dagger} = \hat{I}$  in contrast to  $\hat{U}_{\text{QFT}}^2 = \hat{U}_{\text{QFT}}^{2\dagger} \neq \hat{I}$ . We realize that  $\hat{U}_{\text{QHT}} = \left[ (1-i)\hat{U}_{\text{QFT}} + (1+i)\hat{U}_{\text{QFT}}^\dagger \right] / 2$  is strongly related to the  $\hat{U}_{\text{QFT}}$  and thus suggest the use of an extended QFT circuit to build QHT circuit,<sup>[62]</sup> as shown in Figure 2. The circuit begins with a Hadamard gate applied to the ancilla, being the most significant bit, and a QFT circuit applied to  $N$  qubits, which maps the input binary state  $|x_j\rangle$  from the computational basis to Fourier basis. The combination of a CNOT ladder and a permutation circuit is used to reorder the amplitudes of the conditioned states and is equivalent to a controlled  $\hat{U}_{\text{QFT}}^2$ . A Hermitian adjoint of square root of Pauli  $\hat{X}$  gate is introduced to adjust the relative phases and the intermediate state at this stage is in the form of  $(|0_a\rangle \hat{U}_{\text{QHT}} |x_j\rangle + |1_a\rangle \hat{U}_{\text{QFT}}^2 \hat{U}_{\text{QHT}} |x_j\rangle) / \sqrt{2}$ . The circuit is concluded with the adjoint (conjugate transpose) version of CNOT ladder and permutation circuits, followed by a Hadamard gate to ensure that  $|0_a h(x_j)\rangle$  is left alone for any input states  $|0_a x_j\rangle$ , and the amplitude of  $|h(x_j)\rangle$  is purely real. We note that the ancilla starts and ends in  $|0\rangle$  state ('clean run'). In the following sections, the symbol  $\hat{U}_{\text{QHT}}$  will be treated as a QHT circuit rather than matrix itself, providing the unitary transformation as

$$\hat{U}_{\text{QHT}} = \sum_{j=0}^{2^N-1} |0_a h(x_j)\rangle \langle 0_a x_j| \quad (3)$$

We also note that QHT circuit is not unique, and can be potentially optimized or recompiled for any quantum computing architecture to be used.



**Figure 2.** Quantum Hartley Transform. Quantum Hartley transform  $\hat{U}_{\text{QHT}}$  circuit maps computational basis states  $\{|0_a x_j\rangle\}_{j=0}^{2^N-1}$  into Hartley states  $\{|0_a h(x_j)\rangle\}_{j=0}^{2^N-1}$ , where  $x_j \in \mathbb{Z}_{2^N-1}$ . The QHT circuit involves a QFT circuit acting on  $N$ -qubit system beneath an ancilla qubit, followed by two sequential sets of CNOT ladder and permutation circuits with an inverse  $\sqrt{X}$  gate in the middle, sandwiched between two Hadamard gates applied to the ancilla. The standard single-qubit phase shift gate in QFT circuit is defined as  $P(\phi) = \text{diag}\{1, \exp(i\phi)\}$ .  $H$  and  $X$  are Hadamard and Pauli  $X$  gates, respectively, and  $\sqrt{X}^\dagger$  is equivalent to  $R_Z(\pi/2)P(-\pi/2)R_X(-\pi/2)$ .

### 2.3. Learning and Sampling via the Framework of Quantum Hartley-Based Generative Models

Next, we demonstrate examples of applying the quantum Hartley-based generative models (QHGMs) to two relevant distributions of stochastic models: Ornstein-Uhlenbeck (OU) and geometric Brownian motion (GBM) processes, which among others are widely used in financial analysis. For instance, the former is typically employed to model interest rates and currency exchange rates,<sup>[63]</sup> while the latter is usually utilized to model the log return of stock prices in the Black-Scholes model.<sup>[64]</sup> The stochastic process  $X_t$  at time  $t$  satisfies a stochastic differential equation (SDE)  $dX_t = \nu(\mu - X_t)dt + \sigma dW_t$  for OU and  $dX_t = \mu X_t dt + \sigma X_t dW_t$  for GBM process, where the constant parameters ( $\mu$ ,  $\sigma$  and  $\nu$ ) represent mean (drift), volatility and reversion speed, respectively, and  $W_t$  denotes the Wiener process. Using the Fokker-Planck equation to treat  $X_t$  as a deterministic variable  $x$ , we obtain the time evolution of the underlying probability density function  $p(x, t)$  for a given initial distribution of  $p(x, 0) = \delta(x - x_i)$ . The results show that  $X_t$  follows normal and log-normal distributions, respectively, as

$$p_{\text{OU}}(x, t) = \sqrt{\frac{\nu}{\pi(1 - e^{-2\nu t})\sigma^2}} \exp\left\{-\frac{\nu[x - \mu - (x_i - \mu)e^{-\nu t}]^2}{(1 - e^{-2\nu t})\sigma^2}\right\} \quad (4)$$

$$p_{\text{GBM}}(x, t) = \frac{1}{\sqrt{2\pi\sigma^2 t x}} \exp\left\{-\frac{[\ln(x/x_i) - (\mu - \sigma^2/2)t]^2}{2\sigma^2 t}\right\} \quad (5)$$

Setting  $p(x, t \rightarrow 1) = p_{\text{target}}(x)$  as a target distribution (ground truth), we make use of the Hartley feature map followed by an ansatz parameterized by a vector of variational parameters  $\theta$  that can be adjusted in a hybrid classical-quantum optimization scheme to learn these two distributions. To maintain the real amplitudes of the state vector, the variational ansatz  $\hat{V}_\theta$  comprises adjustable  $R_Y$  rotation and fixed CNOT/CZ gates. An exemplary  $\hat{V}_\theta$  used in this study is illustrated in Figure S1 (Supporting Informa-

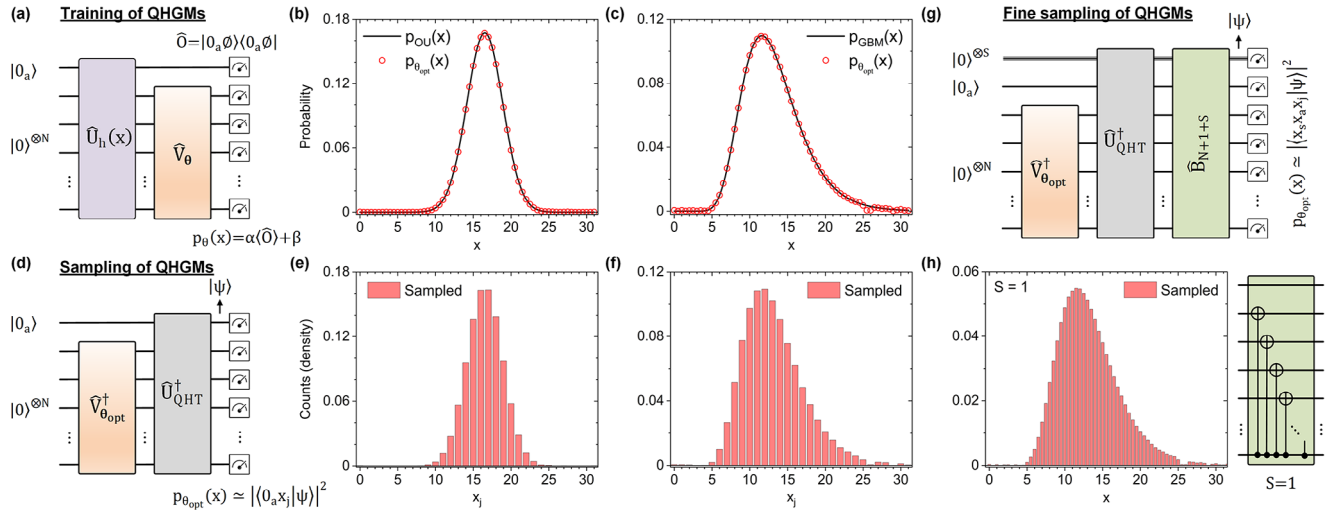
tion). Specifically, the  $N$ -qubit hardware efficient real-amplitude ansatz (HERA) consists of an initial layer of parameterized  $R_Y$  gates applied on each qubit, followed by depth- $d$  blocks presented in the alternating layered architecture. Each block is composed of a cascade of entangling (CNOT) gates and a layer of parameterized  $R_Y$  gates applied on each qubit. The total  $N(d+1)$  trainable parameters are randomly initialized. This ansatz prepares highly correlated quantum states with real-only amplitude while maximizing the expressivity of the model.<sup>[65,66]</sup> We begin with variationally training both normal and log-normal distributions in the latent space through feeding an initial product state  $|0_a \phi\rangle$  to the Hartley feature map  $\hat{U}_h(x)$  connected with a variational ansatz circuit  $\hat{I} \otimes \hat{V}_\theta$ , and then read out the quantum model as an expectation value of an observable  $\hat{O} = |0_a \phi\rangle\langle 0_a \phi|$ . Note that the latter can be substituted by a local proxy during the training stage.<sup>[67]</sup> In order to improve the trainability and expressivity of QML models, they are usually formulated as  $p_\theta(x) = \alpha\langle \hat{O} \rangle + \beta$  with variationally trainable scaling and (optional) bias parameters ( $\alpha$  and  $\beta$ ), as depicted in Figure 3a. The quantum model is trained to search for optimized  $\theta_{\text{opt}}$  to fit the target probability distribution by minimizing a mean squared error (MSE) loss function

$$\mathcal{L}(\theta) = \frac{1}{M} \sum_{m=1}^M [p_\theta(x_m) - p_{\text{target}}(x_m)]^2 \quad (6)$$

where  $M$  is a grid of training points consisting of the integers  $\{x_j\}$  and additional half-integers  $\{x_{(2j+1)/2}\} \forall j \in [0, 2^N - 1]$ . The loss minimization is performed via Adam optimizer for gradient-based training of variational parameters  $\theta$ . In Figure 3b,c we show the trained normal and log-normal distributions using  $N = 5$  qubits with  $\hat{V}_\theta$  of depth  $d = 4$  and 5, respectively. Trained models (red circles) tightly follow the target functions (black solid curves). Because qubits are entangled and rotated differently from the initial state to the target state in each learnable block, the number of depth blocks can significantly impact the learning accuracy of both financially-motivated models.

Since  $p_{\theta_{\text{opt}}}(x) = |\langle 0_a \phi | (\hat{I} \otimes \hat{V}_{\theta_{\text{opt}}}) | 0_a h(x) \rangle|^2 \simeq |\langle 0_a x_j | \hat{U}_{\text{QHT}}^\dagger (\hat{I} \otimes \hat{V}_{\theta_{\text{opt}}}) | 0_a \phi \rangle|^2$ , we can sample the probability distribution of the





**Figure 3.** Learning and sampling probability distributions with quantum Hartley-based generative models (QHGMs). a) Circuit used to train the model  $p_\theta(x)$  in the latent space. Measured observable is defined as  $\hat{O} = |0_a\rangle\langle 0_a|$ , where  $|\phi\rangle \equiv |0\rangle^{\otimes N}$ , with  $\alpha$  and  $\beta$  being trainable scaling and bias parameters. b) Trained  $p_{\theta_{\text{opt}}}(x)$  and target  $p_{\text{OU}}(x)$  distributions with parameters  $\mu = 5, \sigma = 0.3, x_i = 12$  and  $t = 1$  for  $x > 0$ . c) Trained  $p_{\theta_{\text{opt}}}(x)$  and target  $p_{\text{GBM}}(x)$  distributions with parameters  $\mu = 5, \sigma = 3, \nu = 0.5, x_i = 24$  and  $t = 1$ . d) Circuit used to sample the trained model in the computational basis  $|0_a x_i\rangle$ . The quantum state just prior to measurement is denoted as  $|\psi\rangle$ . e, f) Sampled probability distributions from the corresponding trained models, generated with  $10^5$  shots. The normalized histogram is plotted with respect to  $x_i \in \mathbb{Z}_{2N-1}$ . g) Circuit with  $S$ -qubit extended registers on the top line used to perform fine sampling in the computational basis  $|x_1 x_2 \dots x_{N+1+S}\rangle$ . h) Sampled probability distribution from the trained (GBM) model using the extended register of  $S = 1$ , generated with  $10^6$  shots. The normalized histogram is plotted as a function of  $x \in \mathbb{R}_{2N-1}$ . The bitstring network  $\hat{B}$  for  $S = 1$  is shown on the right.

trained model by applying the adjoint versions of the trained ansatz and QHT circuits to the zero product input state,  $|\psi\rangle = \hat{U}_{\text{QHT}}^\dagger (\hat{I} \otimes \hat{V}_{\theta_{\text{opt}}}^\dagger) |0_a\rangle$ , and then perform projective measurements in the computational basis to collect a batch of binary samples, as shown in Figure 3d. The histograms in Figure 3e,f show the resulting sampled probability distributions normalized with the total number of samples, and both of them are in good agreement with the corresponding target distributions (solid curves in Figure 3b,c). One can readily increase the sampling rate at the expense of decrease of maximum probability amplitude through  $S$ -qubit extended registers on the top line,  $x_s \in [0, \dots, 2^S - 1]$ . The relevant circuits associated with the extended registers are  $(N+1+S)$ -qubit inverse QHT and bitstring network  $\hat{B}$  circuits before projective measurements, as depicted in Figure 3g. The requirement of the inclusion of an extra bitstring network for  $S \geq 1$  originates from the use of  $x$ -dependent  $R_z$  gate acting on the ancilla in the Hartley feature map during the training procedure. The nature of  $x$ -dependent argument makes the training stable at the cost of having (bit-dependent) periodic signal on the ancilla. As a demonstration of the double-frequency sampling, the histogram of the sampled probability distribution from the trained (GBM) model with the extended register of  $S = 1$  is plotted as a function of  $x \in \mathbb{R}_{2N-1}$  and presented in Figure 3h, where the length- $(N+1+S)$  readout binary strings are linearly mapped to the domain  $x$  in the range of  $[0, 2^{N+1} - 1]$ . We note that the bitstring network is just a CNOT ladder for  $S = 1$ . Further fine sampling can be achieved by increasing extended register sizes together with different types of bitstring network. The other way to perform dense sampling is borrowed from the sampling procedure of Fourier model. One can transform the computational states to

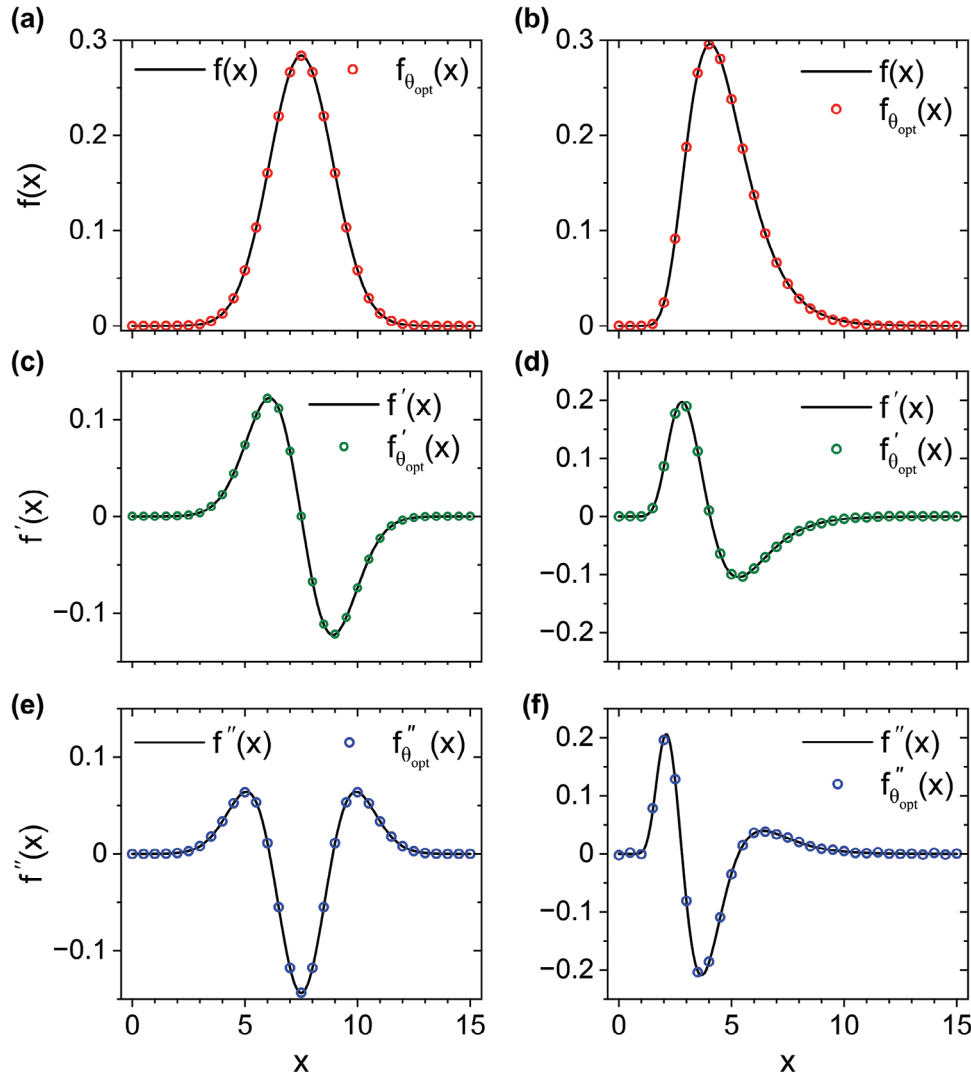
Fourier states by a  $(N+1)$ -qubit QFT circuit, followed by an extended inverse QFT circuit, as shown in Figure S2 (Supporting Information), where the histograms for double- and quadruple-frequency sampling corresponding to different sizes of the extended registers are demonstrated. More importantly, the resulting histograms from both sampling configurations qualitatively match the target distribution, providing a potential pathway to generate novel and unseen datasets that fall within the learnt probability distribution (i.e., draw samples from the trained probability distribution by means of the probabilistic nature of quantum measurements) for practical applications of quantum generative AI in machine learning tasks and quantum physics. In the following sections, we will employ the developed protocols on solving differential equations, making a comparison between Fourier and Hartley models, and exploring the possibility toward multidimensional quantum generative modeling, followed by brief discussions on the impact of noise on QHGMs and the circuit complexity of QHGMs in Appendices A.2 and A.3.

## 2.4. Solving Differential Equations

We first consider two exemplary second-order differential equations (DEs) with variable coefficients to be tackled. The first DE is of the form

$$\frac{d^2 f(x)}{dx^2} + \frac{(x - \mu)}{\sigma^2} \frac{df(x)}{dx} + \frac{f(x)}{\sigma^2} = 0 \quad (7)$$

for some real-valued parameters  $\mu, \sigma$  and the boundary conditions  $f(\mu) = 1/\sqrt{2\pi\sigma^2}$  and  $df(x)/dx|_{x=\mu} = 0$ .



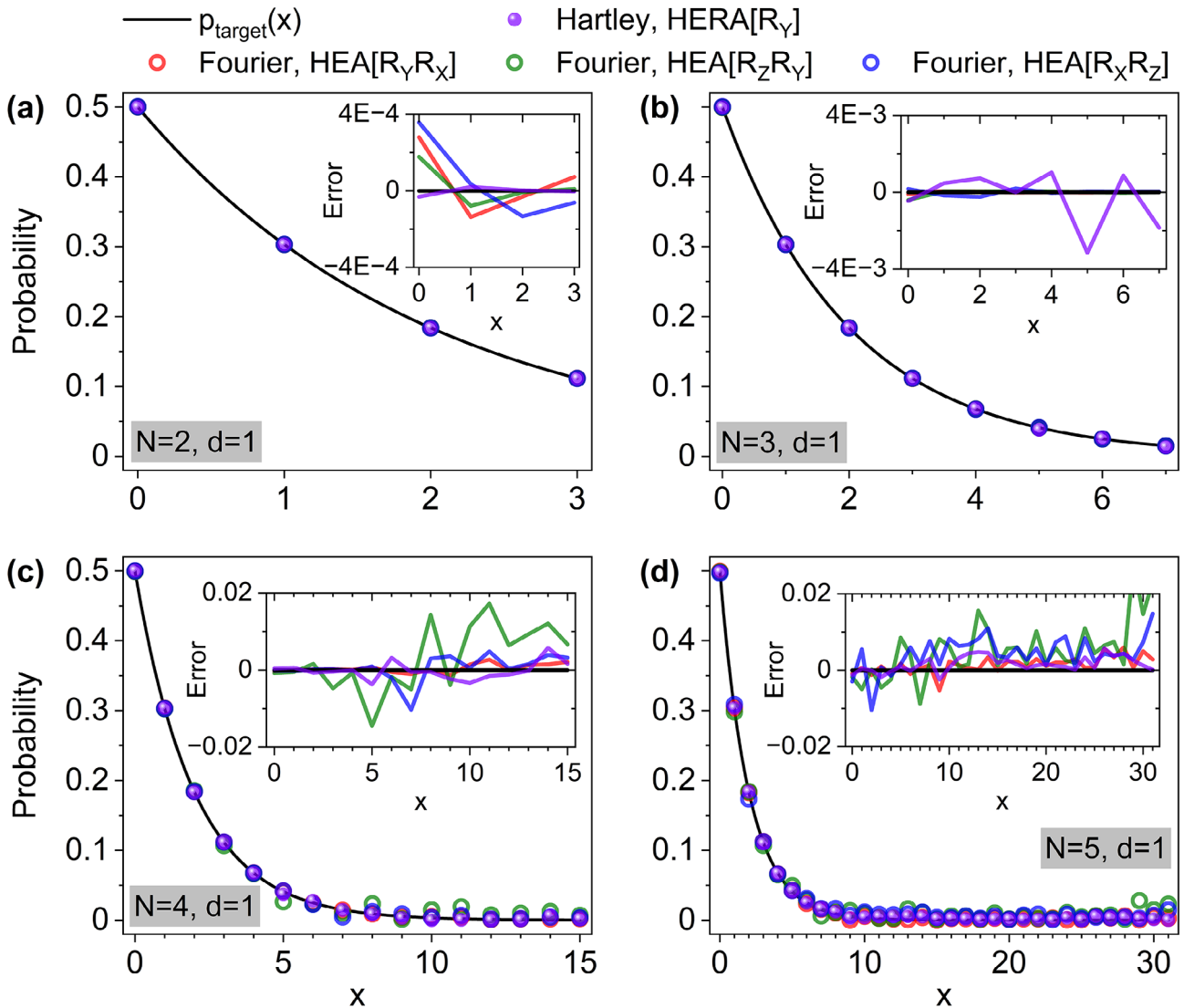
**Figure 4.** Results of solving differential equations. a) Plots of trained model  $f_{\theta_{\text{opt}}}(x)$  and analytic solution  $f(x)$  of Equation (7) with boundary conditions of  $f(\mu) = 1/\sqrt{2\pi\sigma^2}$  and  $df(x)/dx|_{x=\mu} = 0$ , where parameters are  $\mu = 7.5$  and  $\sigma = 1.406$ . b) Plots of trained model  $f_{\theta_{\text{opt}}}(x)$  and analytic solutions  $f(x)$  of Equation (8) with boundary conditions of  $f(e^{\mu-\sigma^2}) = e^{(\sigma^2/2-\mu)}/\sqrt{2\pi\sigma^2}$  and  $df(x)/dx|_{x=e^{\mu-\sigma^2}} = 0$ , where parameters are  $\mu = 1.5$  and  $\sigma = 0.316$ . The corresponding c,d) first and e,f) second derivatives of the analytic solutions and trained models.

This differential equation has a known analytical solution  $f(x) = \exp\left\{-0.5\left[(x-\mu)/\sigma\right]^2\right\}/\sqrt{2\pi\sigma^2}$ . The second DE reads

$$\frac{d^2f(x)}{dx^2} + \frac{[2\sigma^2 - \mu + \ln(x)]}{\sigma^2 x} \frac{df(x)}{dx} + \frac{f(x)}{\sigma^2 x^2} = 0 \quad (8)$$

with the boundary conditions  $f(e^{\mu-\sigma^2}) = e^{(\sigma^2/2-\mu)}/\sqrt{2\pi\sigma^2}$  and  $df(x)/dx|_{x=e^{\mu-\sigma^2}} = 0$ . The analytical solution of this DE is  $f(x) = \exp\left\{-0.5\left[(\ln x - \mu)/\sigma\right]^2\right\}/(\sqrt{2\pi\sigma^2} x)$  for  $x > 0$ . Solving DEs requires evaluations of the first and second derivatives of the quantum model  $f_{\theta}(x) = \alpha\langle 0_a\phi|\hat{U}_h^\dagger(x)(\hat{I} \otimes \hat{V}_\theta^\dagger)\hat{\mathcal{O}}(\hat{I} \otimes \hat{V}_\theta)\hat{U}_h(x)|0_a\phi\rangle + \beta$  with respect to  $x$ , which relies on a differentiable Hartley feature map  $\hat{U}_h(x)$  circuit. Gradient calculations are implemented with automatic differentiation techniques

(backpropagation) or with the application of the parameter shift rule.<sup>[68–70]</sup> For the latter case, controlled gates  $\hat{P}_l^{-2}(x)$  need to be decomposed using CNOT conjugations, and differentiation of  $\hat{U}_h(x)$  requires  $4N + 2$  shifts in total. We then variationally minimize an overall MSE loss function. Specifically, the total loss function is the equal-weighted sum of the contributions that match the differential equation and satisfy the boundary conditions. The loss is minimized with the gradient-based Adam optimizer with a small learning rate of 0.01 and generally the loss error converges to the level of  $10^{-6}$  in a few thousand epochs. We use the Python programming language (PennyLane from Xanadu<sup>[71]</sup> and Qadence from Pasqal<sup>[72]</sup>) together with machine learning packages (NumPy, JAX and PyTorch) for the full statevector simulation. In particular, we utilize Hartley encoding over  $N = 4$  qubits and the HERA of depth 3 to solve Equations (7) and (8). The results are presented in **Figure 4a–f**. The red, green,



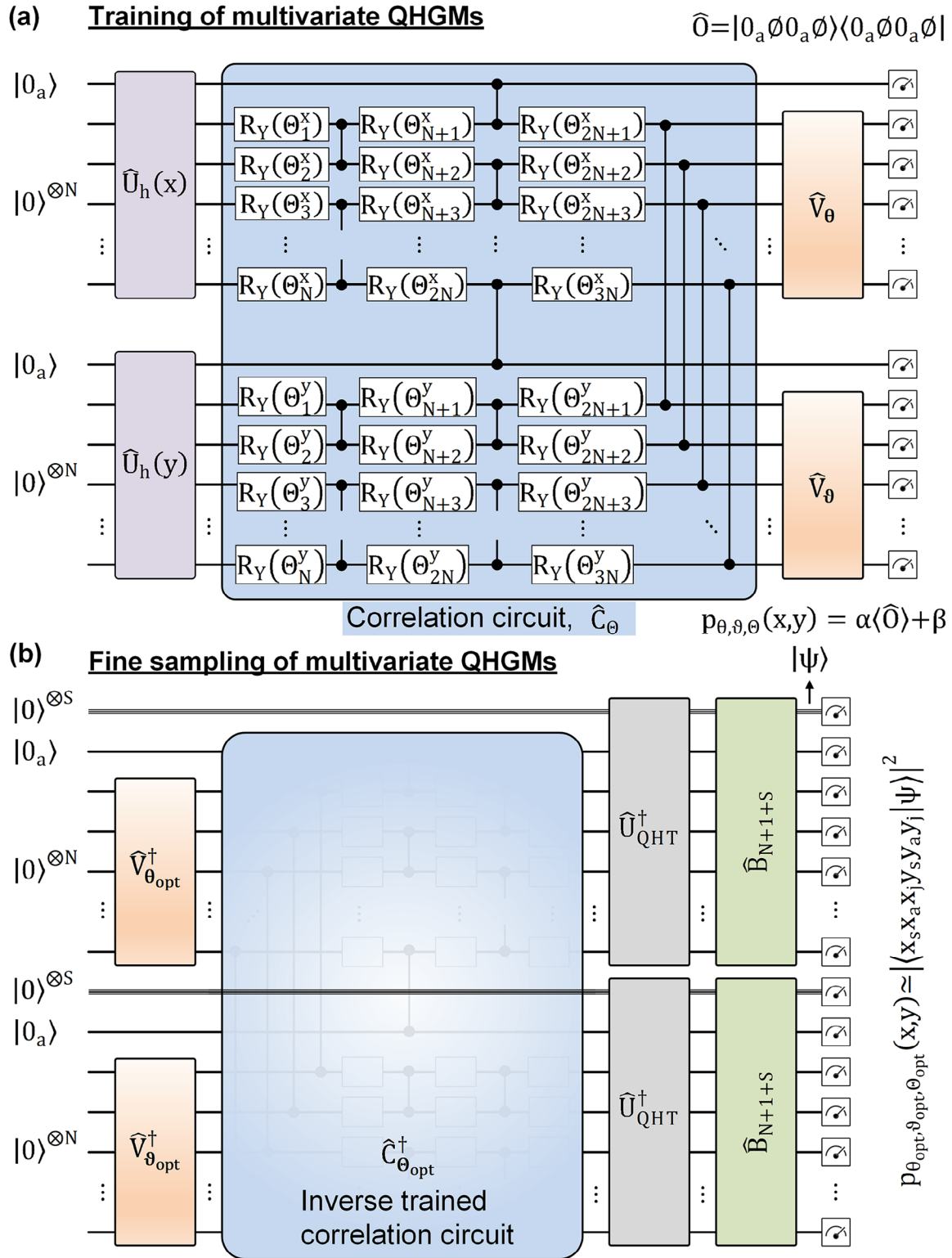
**Figure 5.** Comparison between the quantum Fourier- and Hartley-based models. Training a probability exponential distribution with rate parameter  $\lambda = 0.5$ ,  $p_{\text{target}}(x) = \lambda e^{-\lambda x}$  for  $x \geq 0$ , using Fourier and Hartley feature map circuits over  $N = 2$  a), 3 b), 4 c), and 5 qubits d) with the shallow ansatz circuit depth of  $d = 1$ . For Fourier models, the parameterized gate element in the HEA is composed of two adjustable single-qubit rotation gates per qubit ( $R_Y R_X$ ,  $R_Z R_Y$ ,  $R_X R_Z$ , expressed in matrix multiplication order). For Hartley models, only the  $R_Y$  gate is required in the HERA. The insets show the corresponding relative errors.

and blue hollow-circle curves represent functions and their first and second derivatives, respectively,  $f_{\theta_{\text{opt}}}(x)$ ,  $df_{\theta_{\text{opt}}}(x)/dx$  and  $d^2f_{\theta_{\text{opt}}}(x)/dx^2$  evaluated at optimal angles  $\theta_{\text{opt}}$  retrieved after the optimization procedure. Overall, the trained models are consistent with the analytic solutions (solid black curves).

## 2.5. Complex Fourier vs Real Hartley Models

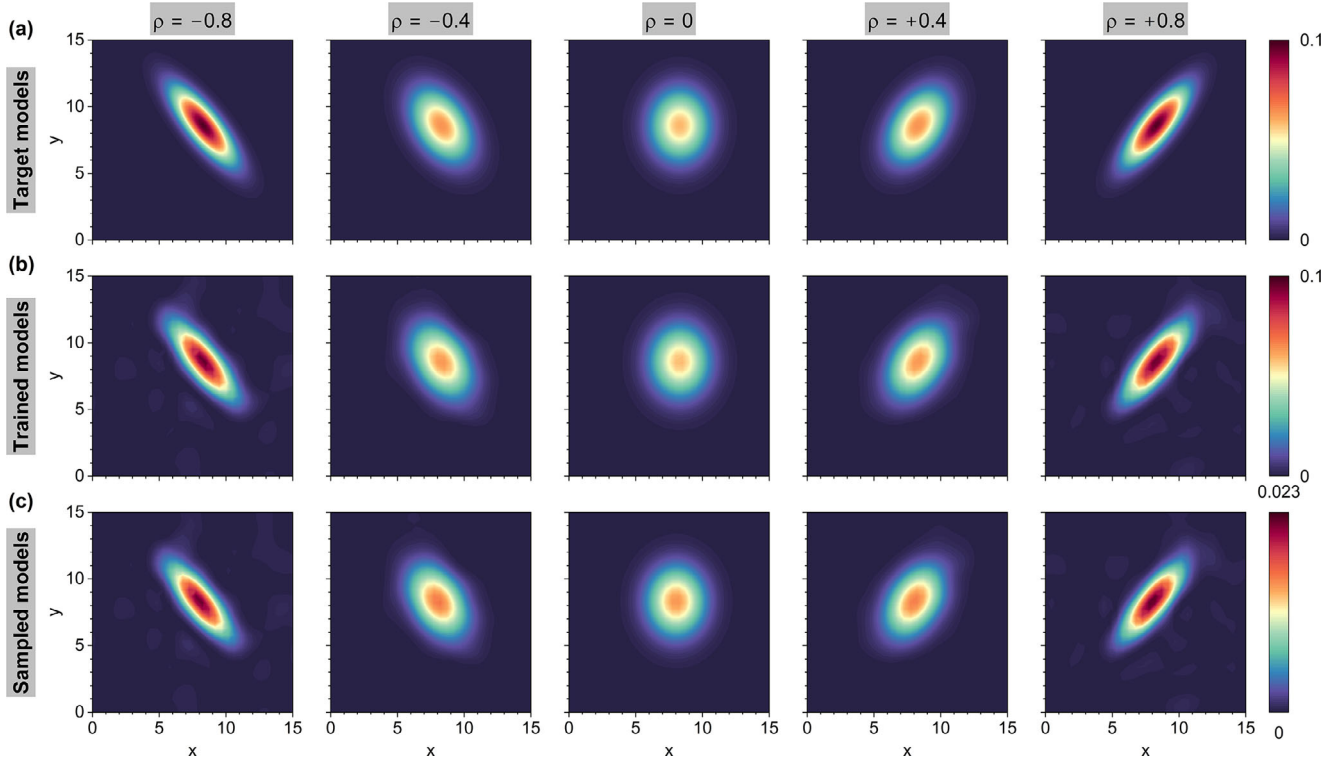
To evaluate the potential of using the Hartley models, we consider an exponential distribution with probability density function  $p(x) = \lambda e^{-\lambda x}$  as a ground truth. This  $p(x)$  describes the distances between successive events in homogeneous Poisson processes. As being the only continuous memoryless probability distribution, it is widely used in the calculation of various systems

in queuing theory and reliability theory.<sup>[73]</sup> We employ  $N$ -qubit Fourier encodings followed by different configurations of hardware efficient ansatz (HEA) to learn this exponential distribution on a training grid of  $2^N$  integer points and compare the results with Hartley models. Note that the ancillary register is not required for Fourier models (see Figure A1c). Due to the complex nature of the Fourier models, the parameterized gate element inside each depth block of the HEA is composed of two single-qubit rotation gates per qubit (marked as  $R_Y R_X$ ,  $R_Z R_Y$ ,  $R_X R_Z$  in Figure 5), each gate parameterized by a given angle, and the CNOT-based entangling layers are the same as the HERA (see Figure S3, Supporting Information, for other cases of parameterized gates). Therefore, there are  $2N(d + 1)$  trainable parameters for HEA of depth  $d$ . These HEAs allow easy access to the solution space by preparing quantum states with complex amplitude and



**Figure 6.** Multivariate quantum Hartley-based generative models. a) Circuit used to train the multivariate distribution in the latent space, where a parameterized correlation circuit  $\hat{C}_\theta$  is sandwiched between feature maps ( $\hat{U}_h(x)$  and  $\hat{U}_h(y)$ ) and variational ansatzes ( $\hat{V}_\theta$  and  $\hat{V}_\theta$ ). Measured observable is defined as  $\hat{O} = |0_a \phi 0_a \phi\rangle\langle 0_a \phi 0_a \phi|$  is utilized, where  $|\phi\rangle \equiv |0\rangle^{\otimes N}$ . Here,  $\alpha$  and  $\beta$  are trainable scaling and bias parameters. b) Circuit used to sample the multivariate distribution from the trained model, where  $\theta_{\text{opt}}$ ,  $\theta_{\text{opt}}$  and  $\theta_{\text{opt}}$  are retrieved after the optimization procedure in (a). Two identical sets of inverse QHT ( $\hat{U}_{\text{QHT}}^+$ ) and bitstring network ( $\hat{B}$ ) circuits associated with extended registers of  $S$  qubits are then applied in parallel for fine sampling in the computational basis  $|x_S x_a x_j y_S y_a y_j\rangle$ . The quantum state just prior to measurement is denoted as  $|\psi\rangle$ .





**Figure 7.** Learning and sampling bivariate probability distributions with multivariate quantum Hartley-based generative models. a) Density plots of target  $p_{\text{BN}}(x, y)$  distributions with parameters  $\mu_x = 8.3$ ,  $\mu_y = 8.6$  and  $\sigma_x = 1.5$ ,  $\sigma_y = 1.8$  for different correlation coefficients ranging from  $\rho = -0.8$  to  $0.8$  in a step of  $0.4$  (left to right). b) Density plots of corresponding trained  $p_{\theta_{\text{opt}}, \theta_{\text{opt}}, \theta_{\text{opt}}}(x, y)$  models. c) Normalized density plots of sampled distributions from the corresponding trained models in (b) using the parallel extended registers of  $S = 1$ , where  $10^7$  shots are measured at the readout. All plots share the same color bar on the right in each row with the same correlation coefficient in each column.

avoid the training issues such as barren plateaus and local minima for complex Fourier models. As shown in Figure 5a–d, both Fourier and Hartley models follow  $p_{\text{target}}(x)$  closely with small relative errors for varying number of qubits  $N$  under the same ansatz depth  $d = 1$ . However, the number of variational parameters in complex Fourier models is double compared to those in real Hartley models. Our results indicate for a  $N$ -qubit quantum system ( $N \geq 2$ ) that one can simply employ the Hartley model with the minimal number of parameterized gates needed to efficiently reach the target solution space compared to the Fourier model. This becomes important when considering practical implementations with limited quantum resources and relevant applications.

## 2.6. Multivariate Quantum Hartley-Based Generative Models

We proceed to extend the proposed strategies from univariate to multivariate distributions. As an example, we consider a bivariate/binormal distribution with probability density function in the form

$$p_{\text{BN}}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y} \exp\left[\frac{-(z_x^2 + z_y^2 - 2\rho z_x z_y)}{2(1-\rho^2)}\right] \quad (9)$$

where  $z_x = (x - \mu_x)/\sigma_x$  and  $z_y = (y - \mu_y)/\sigma_y$  with parameters  $\mu_x$ ,  $\mu_y$  and  $\sigma_x$ ,  $\sigma_y$  representing mean and standard deviation values for each (either stochastic or deterministic) variables  $x$  and  $y$ . Here,  $\rho$  is a correlation coefficient,  $-1 < \rho < 1$ . We start with Hartley encoding of two independent variables  $x$  and  $y$  in parallel registers, followed by a parameterized correlation circuit  $\hat{C}_\Theta$  and two separate variational ansatzes  $\hat{V}_\theta$  and  $\hat{V}_\phi$ , as illustrated in Figure 6a. The quantum model  $p_{\theta, \theta, \theta}(x, y)$  is trained to represent the target  $p_{\text{BN}}(x, y)$  through searching for optimal angles,  $\theta_{\text{opt}}$ ,  $\phi_{\text{opt}}$  and  $\Theta_{\text{opt}}$ , the same procedure in the previous univariate case. The correlation circuit plays a crucial role in making two otherwise independent latent variables correlated for efficient training, while keeping the sampling procedure the same as previously discussed in the univariate case. As usual,  $\hat{C}_\Theta$  comprises alternating layers of adjustable  $R_Y$  rotation and fixed CZ gates to ensure the real amplitudes of the state vector. Specifically, the  $(2N + 2)$ -qubit correlation circuit consists of three layers of parameterized  $R_Y$  gates applied on each qubit (except for two ancilla), two layers of CZ gates applied to odd and even subsequent pairs of qubits, respectively, and a final layer of CZ gates applied to the same qubit index between separate registers. Therefore, the overall number of trainable parameters includes  $6N$  by the correlation circuit and  $2N(d + 1)$  contributed by two variational ansatzes, all of which are randomly initialized. For the circuit used to densely sample multivariate distributions illustrated in Figure 6(b), we run the trained circuit with inverted parameters in reverse order and replace

the Hartley feature maps with extended inverse QHT circuits, a procedure accounting for the basis transformation from the latent to bitbasis space. Two identical bitstring networks are then applied in parallel, followed by projective measurements in the computational basis to collect a batch of binary samples, as described in the previous section. Similarly, its variant model via QFTs is shown in Figure S4 (Supporting Information). Figure 7a shows density plots of analytical bivariate (binormal) distribution based on Equations (9) with five different correlation coefficients, set as 2D target distributions (ground truths). The corresponding density plots of the trained quantum models with  $N = 4$  qubits and the variational ansatz of depth  $d = 2$  are displayed in each column of Figure 7b, respectively. As expected, they all quantitatively follow target distributions presented in Figure 7a. Typically the loss error reaches to the level  $10^{-6}$  after a few hundred iterations for each scenario. Successful bivariate learning over a wide range of correlation coefficients can be attributed to the well-designed architecture of the latent-space training of the quantum model with a problem-specific correlation circuit included. The proposed correlation circuit structure is designed to simultaneously support the successful training of uncorrelated, partially and highly (positive/negative) correlated distributions with  $|\rho|$  smaller than 0.9. For the cases of  $|\rho|$  higher than 0.9, increasing the depth of the variational ansatz circuit or a modification of the correlation circuit is essential. After projective measurements in the computational basis, we have to perform classical post-processing tasks on a batch of measured binary datasets. This involves with periodically dropping out those bits with zero probability, concatenating rest of the bits in a sequential way and reshaping the resulting bitstring in a 2D array for data plotting. For the case of the extended register of  $S = 1$ , the corresponding normalized density plots of sampled distributions are shown in Figure 7c, which are in excellent agreement with the target distributions. In the cases of  $|\rho| = 0.8$ , some defects (light-purple spots) emerge from the backgrounds of both trained and sampled models, but they do not affect the identification of appearance of the binormal distributions. Finally, the proposed framework enables us to systemically build multidimensional quantum models and easily scale to larger system sizes without the need of a huge amount of change. The size of each modular unit such as Hartley feature map, variational ansatz, quantum Hartley transform and bitstring network, is essentially determined by “local” Hilbert space independent of neighboring subsystem and each individual module has the same circuit architecture as that in the univariate case. That is to say, we only need to focus on the modification or redesign of layers of CZ gates responsible for “global” correlations among three separate registers when tackling  $p(x, y, z)$  associated with three correlated independent variables. As a consequence, we envision a future research of quantum generative modeling going beyond simple univariate towards complex multivariate diffusion models using the developed techniques presented in this work.

### 3. Conclusion

In this study, we developed protocols for building quantum generative models based on Hartley kernels where embedding is represented by a quantum state with amplitudes being  $x$ -dependent functions scaling as  $\cos(2\pi kx/2^N)$ . Being real-valued, Hartley

models are suitable for solving symmetry-related tasks and provide an advantage over Fourier basis states with inherently complex coefficients. We proposed a data-dependent embedding circuit to generate the exponentially expressive orthonormal Hartley basis in the latent space, enabling the differentiation of quantum models, and designed a real-amplitude ansatz for efficient training of quantum models. We constructed the quantum Hartley transform circuit for mappings between Hartley and computational bases. With these tools, we performed generative modeling from the learnt probability distributions being solutions of stochastic differential equations that arise in financial modeling, and demonstrated efficient sampling of these distributions in computational basis, revealing the consistent profiles between the learnt and sampled probability distributions. We then solved the differential equations and demonstrated favorable generalization properties of QHGMs. Finally, we showcased the capability of multivariate QHGMs, where we developed a problem-specific correlation circuit and used parallel extended registers for both correlated and uncorrelated probability distribution. This opens a way to multivariate learning and sampling enabled by differentiable physics-informed quantum generative models.

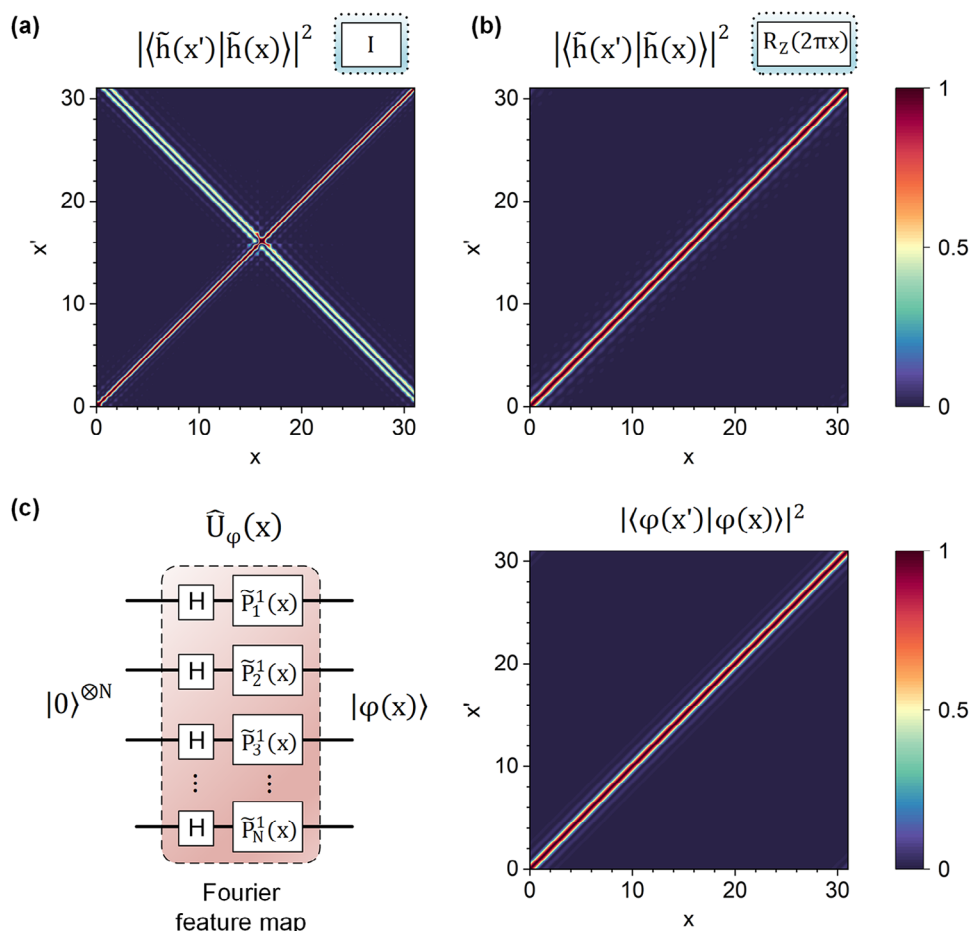
## Appendix A

### A.1. 2D Squared Overlaps

Similar to Equation (1), the squared overlap is employed to examine the orthonormal behavior of different states of continuous variables. As illustrated in Figure A1a, the squared overlap between mutual normalized Hartley states,  $|\langle \hat{h}(x') | \hat{h}(x) \rangle|^2$ , shows unity along the diagonal 45 degree line, meaning that the diagonal states are orthonormal as usual. However, those off-diagonal states, i.e., states along 135 degree line, reveal an unfavored behavior as they disrupt orthogonality in between nodes, contributing at least 50% of the maximum value. This is quite different from the Fourier (phase) feature map<sup>[48]</sup> written in the Fourier basis (Figure A1c). This is mainly caused by the abrupt change in amplitude of the opposite states, making the model potentially difficult to train when using points between Hartley nodes. This can be fixed by introducing an additional  $x$ -dependent rotation that fixes overlap in between nodes. Specifically, by appending a  $R_Z(2\pi x)$  gate to the ancilla register (see Figure 1), we ensure a smooth transition between continuous states while retaining the real-valued amplitudes of  $|\hat{h}(x)\rangle$  over the entire domain with an off-diagonal contribution less than 5% (Figure A1b). With  $R_Z(2\pi x)$  gate included, the resulting squared overlap looks similar to that in Fourier case.

### A.2. Impact of Noise

As the noise-induced error can be compensated by the variational ansatz during the training stage, the presence of low-level noise is not a major obstacle to the training of QHGMs (i.e., its level is smaller than that of sampling noise and there are no noise-induced barren plateaus). However, the influence of noise on the sampling of QHGMs is not negligible. To examine the performance of QHGMs operating under noise, for illustration purposes we consider coherent noise and the technical limitations in setting up correct phases. Practically, at the sampling stage this corresponds to running  $\hat{V}_{\theta_{\text{opt}} \pm \theta_{\text{noise}}}^\dagger$  with a vector  $\theta_{\text{noise}} = p\epsilon$  such that  $\epsilon$  is drawn randomly uniformly on  $[-1, 1]$  and  $p$  represents a percentage error factor. By varying  $p$ , this noise-incorporated model mimics noise-corrupted QHGMs. Similarly, other sources of noise can be modeled using wavefunction Monte Carlo, again leading to partially modified distributions. As shown in Figure A2, the sampled OU distributions still follow the target



**Figure A1.** 2D squared overlap between mutual normalized states. a,b) Representative 2D squared overlap between normalized Hartley states  $|\langle \tilde{h}(x') | \tilde{h}(x) \rangle|^2$  for  $\hat{U}_h(x)$  ( $N=5$ ) without and with the  $R_z(2\pi x)$  gate, respectively. Both cases show that Hartley states are orthonormal when  $x = x'$ , but the nonzero overlap residues in (a) are way too high, inhibiting model trainability. c)  $N$ -qubit Fourier feature map and representative 2D squared overlap between Fourier states  $|\langle \varphi(x') | \varphi(x) \rangle|^2$  for  $\hat{U}_\varphi(x)$  ( $N=5$ ). Notably, the Hartley overlaps are purely real,  $\langle \tilde{h}(x') | \tilde{h}(x) \rangle \in \mathbb{R}$ , unlike the complex overlaps between Fourier states,  $\langle \varphi(x') | \varphi(x) \rangle \in \mathbb{C}$ .

(noiseless) distribution for  $p$  less than 10% and deviate from the ideal case for  $p$  around 20%. Ultimately, the utility of the results under noisy sampling depends on the sensitivity of our task.

### A.3. Circuit Complexity

The total number of single-qubit and two-qubit gates for Hartley feature map circuit is  $3N + 4$ . For quantum Hartley transform circuit, the total number is  $N(N+2)/2$  (contributed by QFT) +  $2N$  (by two sets of CNOT ladder) + 3 (single-qubit gates on the ancilla) + the contribution from two sets of permutation elements. The Toffoli-based decompositions described in Barenco et al.<sup>[74]</sup> and He et al.<sup>[75]</sup> can be used to estimate the number of quantum gates contributed by permutation elements. The circuit complexity of QHGMs can be attributed to quantum Hartley transform circuit at the sampling stage. QHGMs can be readily differentiated but QCBMs is not automatically differentiable because its variable  $x$  is intrinsically length- $N$  binary strings. This differentiation is sometimes infeasible with QGAN architecture as well due to its ill-defined loss landscape.<sup>[76]</sup> Therefore, the proposed QHGMs is appropriate for SDE-based generative modeling when considering to generate samples from SDE-relevant distributions. In addition, QHGMs is easily trained and allow for systemically building large-scale generative models, particularly well-

suited for scaling as the complexity of problems increases, while QCBMs and QGANs are typically difficult to train at increasing scale.

## Supporting Information

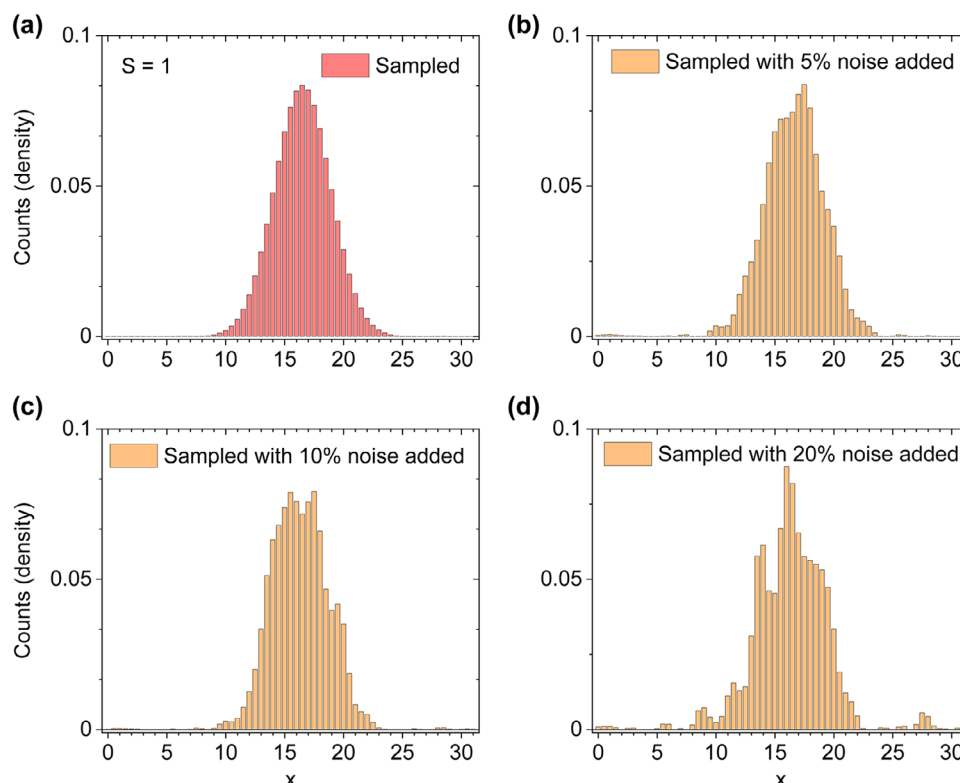
Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

The authors thank Annie E. Paine, Chelsea A. Williams, and Antonio Andrea Gentile for useful discussions on the project. The authors acknowledge the funding from UK EPSRC award under the Agreement EP/Y005090/1 and this work was supported by PASQAL.

## Conflict of Interest

The authors declare no conflict of interest. A patent application for the method described in this manuscript has been submitted by PASQAL with H.-Y.W., O.K., and V.E.E. as inventors.



**Figure A2.** The effect of noise on the sampled probability distributions. Before sampling, the trained noise-free model  $p_{\theta_{\text{opt}}}(x)$  is consistent with the analytic solution  $p_{\text{OU}}(x)$  with parameters  $\mu = 5$ ,  $\sigma = 3$ ,  $\nu = 0.5$ ,  $x_i = 24$  and  $t = 1$ . The sampled probability distributions from a) noise-free and b–d) noise-incorporated trained (OU) models, using the extended register of  $S = 1$ , generated with  $10^6$  shots. The normalized histogram is plotted as a function of  $x \in \mathbb{R}_{2N-1}$ .

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

quantum algorithms, quantum computation, quantum computers, quantum generative modeling, quantum Hartley transform, quantum machine learning

Received: July 17, 2024

Revised: October 11, 2024

Published online: November 11, 2024

- [1] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, Cambridge **2011**.
- [2] P. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, **1994**, pp. 124–134.
- [3] S. Lloyd, *Science* **1996**, 273, 1073.
- [4] I. Buluta, F. Nori, *Science* **2009**, 326, 108.
- [5] A. W. Harrow, A. Hassidim, S. Lloyd, *Phys. Rev. Lett.* **2009**, 103, 150502.
- [6] A. Ambainis, *arXiv:1010.4458* **2010**.
- [7] D. W. Berry, *J. Phys. A: Math. Theor.* **2014**, 47, 105301.
- [8] E. Farhi, J. Goldstone, S. Gutmann, *arXiv:1411.4028* **2014**.

- [9] S. Lloyd, *arXiv:1812.11075* **2018**.
- [10] S. Hadfield, Z. Wang, B. O’Gorman, E. G. Rieffel, D. Venturelli, R. Biswas, *Algorithms* **2019**, 12, 2.
- [11] A. Y. Kitaev, *arXiv:quant-ph/9511026* **1995**.
- [12] R. Cleve, A. Ekert, C. Macchiavello, M. Mosca, *Proc. Royal Soc. London. Series A: Math., Phys. Eng. Sci.* **1998**, 454, 339.
- [13] L. Hales, S. Hallgren, in *Proceedings 41st Annual Symposium on Foundations of Computer Science*, **2000**, pp. 515–525.
- [14] D. Coppersmith, *arXiv:quant-ph/0201067* **2002**.
- [15] J. W. Cooley, J. W. Tukey, *Math. Comput.* **1965**, 19, 297.
- [16] W. T. V. William H. Press, Austin Saul A. Teukolsky, B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge **2007**.
- [17] S. S. Aghaian, A. Klappenecker, in *Image Processing: Algorithms and Systems*, vol. 4667. International Society for Optics and Photonics, SPIE, **2002**, pp. 1–11.
- [18] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O’Brien, *Nat. Commun.* **2014**, 5, 1.
- [19] I. S. Maria Schuld, F. Petruccione, *Contemp. Phys.* **2015**, 56, 172.
- [20] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, *Nature* **2017**, 549, 242.
- [21] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, S. Lloyd, *Nature* **2017**, 549, 195.
- [22] M. Benedetti, E. Lloyd, S. Sack, M. Fiorentini, *Quantum Sci. Technol.* **2019**, 4, 043001.
- [23] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, D. Jaksch, *Phys. Rev. A* **2020**, 101, 010301.
- [24] J. Biamonte, *Phys. Rev. A* **2021**, 103, L030401.



- [25] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, P. J. Coles, *Nat. Rev. Phys.* **2021**, 3, 625.
- [26] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, P. J. Coles, *Quantum* **2023**, 7, 1188.
- [27] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, S. Ganguli, *arXiv:1503.03585* **2015**.
- [28] S. Cheng, J. Chen, L. Wang, *Entropy* **2018**, 20, 8.
- [29] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Commun. ACM* **2020**, 63, 139.
- [30] J. Ho, A. Jain, P. Abbeel, *arXiv:2006.11239* **2020**.
- [31] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, B. Poole, *arXiv:2011.13456* **2021**.
- [32] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, R. Melko, *Phys. Rev. X* **2018**, 8, 021050.
- [33] C. Zoufal, A. Lucchi, S. Woerner, *Quantum Mach. Intell.* **2021**, 3, 7.
- [34] S. Lloyd, C. Weedbrook, *Phys. Rev. Lett.* **2018**, 121, 040502.
- [35] P.-L. Dallaire-Demers, N. Killoran, *Phys. Rev. A* **2018**, 98, 012324.
- [36] C. Zoufal, A. Lucchi, S. Woerner, *npj Quantum Inf.* **2019**, 5, 103.
- [37] B. Zhang, P. Xu, X. Chen, Q. Zhuang, *Phys. Rev. Lett.* **2024**, 132, 100602.
- [38] M. Parigi, S. Martina, F. Caruso, *Adv. Quantum Technol.* *n/a*, 2300401.
- [39] C. Chen, Q. Zhao, M. C. Zhou, Z. He, Z. Sun, H. Situ, *arXiv:2401.07039* **2024**.
- [40] J.-G. Liu, L. Wang, *Phys. Rev. A* **2018**, 98, 062324.
- [41] M. S. Rudolph, S. Lerch, S. Thanasilp, O. Kiss, S. Vallecorsa, M. Grossi, *arXiv:2305.02881* **2023**.
- [42] A. Perdomo-Ortiz, M. Benedetti, J. Realpe-Gómez, R. Biswas, *Quantum Sci. Technol.* **2018**, 3, 030502.
- [43] M. Benedetti, D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, A. Perdomo-Ortiz, *npj Quantum Inf.* **2019**, 5, 45.
- [44] B. Coyle, D. Mills, V. Danos, E. Kashefi, *npj Quantum Inf.* **2020**, 6, 60.
- [45] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, J. M. Gambetta, *Nature* **2019**, 567, 209.
- [46] M. Schuld, N. Killoran, *Phys. Rev. Lett.* **2019**, 122, 040504.
- [47] S. Lloyd, G. D. Palma, C. Gokler, B. Kiani, Z.-W. Liu, M. Marvian, F. Tennie, T. Palmer, *arXiv:2011.06571* **2020**.
- [48] O. Kyriienko, A. E. Paine, V. E. Elfving, *Phys. Rev. Res.* **2024**, 6, 03329.
- [49] S. Kature, O. Kyriienko, V. E. Elfving, *Phys. Rev. A* **2023**, 108, 042406.
- [50] O. Kyriienko, A. E. Paine, V. E. Elfving, *Phys. Rev. A* **2021**, 103, 052416.
- [51] M. Schuld, F. Petruccione, *Supervised Learning with Quantum Computers*, Springer Cham, Switzerland **2018**.
- [52] S. Lloyd, M. Schuld, A. Ijaz, J. Izaac, N. Killoran, *arXiv:2001.03622* **2020**.
- [53] C. A. Williams, A. E. Paine, H.-Y. Wu, V. E. Elfving, O. Kyriienko, *arXiv:2306.17026* **2023**.
- [54] A. E. Paine, V. E. Elfving, O. Kyriienko, *arXiv:2308.01827* **2023**.
- [55] G. Mazzola, *Phys. Rev. A* **2021**, 104, 022431.
- [56] A patent application for the method described in this manuscript has been submitted by PASQAL, **2024**.
- [57] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd edition, McGraw-Hill, USA **1999**.
- [58] A. M. G. Alexander D. Poularikas, R. C. Dorf, *Transforms and Applications Handbook*, 3rd edition, CRC Press, USA **2010**.
- [59] K. J. Jones, *The Regularized Fast Hartley Transform: Low-Complexity Parallel Computation of the FHT in One and Multiple Dimensions*, 2nd edition, Springer, UK **2022**.
- [60] M. Schuld, *arXiv:2101.11020* **2021**.
- [61] A. M. Childs, R. Kothari, R. D. Somma, *SIAM Journal on Computing* **2017**, 46, 1920.
- [62] A. Klappenecker, M. Roetteler, *arXiv:quantph/0111039* **2001**.
- [63] O. Vasicek, *J. Financial Economics* **1977**, 5, 177.
- [64] F. Black, M. Scholes, *J. Political Economy* **1973**, 81, 637.
- [65] M. Schuld, A. Bocharov, K. M. Svore, N. Wiebe, *Phys. Rev. A* **2020**, 101, 032308.
- [66] M. Schuld, R. Sweke, J. J. Meyer, *Phys. Rev. A* **2021**, 103, 032430.
- [67] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, P. J. Coles, *Nat. Commun.* **2021**, 12, 1791.
- [68] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, N. Killoran, *Phys. Rev. A* **2019**, 99, 032331.
- [69] K. Mitarai, M. Negoro, M. Kitagawa, K. Fujii, *Phys. Rev. A* **2018**, 98, 032309.
- [70] O. Kyriienko, V. E. Elfving, *Phys. Rev. A* **2021**, 104, 052417.
- [71] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linares, B. AkashNarayanan, A. Asadi, J. M. Arrazola, U. Azad, S. Banning, C. Blank, T. R. Bromley, B. A. Cordier, J. Ceroni, A. Delgado, O. Di Matteo, A. Dusko, T. Garg, D. Guala, A. Hayes, R. Hill, A. Ijaz, T. Isacsson, D. Ittah, S. Jahangiri, P. Jain, E. Jiang, et al., *arXiv:1811.04968* **2022**.
- [72] D. Seitz, N. Heim, J. P. Moutinho, R. Guichard, V. Abramavicius, A. Wennersteen, G.-J. Both, A. Quelle, C. de Groot, G. V. Velikova, V. E. Elfving, M. Dagrada, *arXiv:2401.09915* **2024**.
- [73] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd Edition, Vol. 2, Wiley, London **1971**.
- [74] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, H. Weinfurter, *Phys. Rev. A* **1995**, 52, 3457.
- [75] H. Yong, M. Luo, E. Zhang, H. Wang, X.-F. Wang, *Int. J. Theor. Phys.* **2017**, 56, 2350.
- [76] A. E. Paine, V. E. Elfving, O. Kyriienko, *Adv. Quantum Technol.* **2023**, 6, 2300065.