

Variational Optical Phase Learning on a Continuous-Variable Quantum Compiler

Matthew A. Feldman^{1,2,*} Tyler Volkoff^{3,†} Seongjin Hong² Claire E. Marvinney^{1,2} Zoë Holmes⁴ Raphael C. Pooser^{1,2} Andrew Sornborger^{5,‡} and Alberto M. Marino^{1,2,‡}

¹*Quantum Information Science Section, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA*

²*Quantum Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA*

³*Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA*

⁴*Laboratory of Quantum Information and Computation, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*

⁵*Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico, USA*



(Received 13 February 2025; revised 20 November 2025; accepted 26 January 2026; published 23 March 2026)

Quantum process learning is a fundamental primitive that draws inspiration from machine learning with the goal of better studying the dynamics of quantum systems. One approach to quantum process learning is quantum compilation, whereby an analog quantum operation is digitized by compiling it into a series of basic gates. While there has been significant focus on quantum compiling for discrete-variable systems, the continuous-variable (CV) framework has received comparatively less attention. We present an experimental implementation of a CV quantum compiler that uses two-mode squeezed light to learn a Gaussian unitary operation. We demonstrate the compiler by learning a parameterized linear phase unitary through the use of target and control phase unitaries to demonstrate a factor of 5.4 increase in the precision of the phase estimation and a 3.6-fold acceleration in the time-to-solution metric when leveraging quantum resources. We further show how our approach can be extended to higher-dimensional compilation tasks. Our results are enabled by the tunable control of our cost landscape via variable squeezing, thus providing a critical framework to simultaneously increase precision and reduce time-to-solution.

DOI: [10.1103/7l8z-zn2m](https://doi.org/10.1103/7l8z-zn2m)

I. INTRODUCTION

There is growing awareness of the power of machine learning to process data from quantum experiments. However, to channel the spirit of Richard Feynman, if the underlying systems are inherently *quantum*, perhaps *quantum* machine learning is a more suitable general framework [1,2]. Branching from fundamental results that extended the framework of neural networks to the discrete-variable (DV) quantum [3–6] and continuous-variable (CV) quantum settings [7], this line of thought motivates the burgeoning field of quantum process learning [8–16], where quantum training data are used to learn a quantum

model for an unknown quantum process. Although quantum learning algorithms, unlike quantum process tomography or quantum sensing algorithms, rely on quantum training data, techniques from these approaches can often be applied in complementary ways toward characterizing quantum dynamics [17–23].

One key application of quantum process learning is quantum compilation, whereby a complex, potentially analog, quantum operation is digitized by compiling it into a series of basic gates that are implementable on the available hardware forming a digital quantum twin. Such an approach can be used with the aim of learning shorter and more noise resilient circuits to reduce the resources required to implement large scale quantum algorithms. Quantum compilation has been proposed for quantum state tomography [18], quantum entangled state preparation [24], quantum circuit optimization [25], and to learn quantum dynamics [26–29].

While receiving less attention, CV platforms provide fertile ground for almost all quantum-based applications, spanning from quantum computation [30] and quantum machine learning [7,31] to quantum sensing [32–34]. The

*Contact author: feldmanma@ornl.gov

†Contact author: volkoff@lanl.gov

‡Contact author: marinoa@ornl.gov

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

allure of these systems is rooted in their operational viability at ambient temperatures, their capacity for deterministic quantum entanglement [35,36], and the effective management of quantum noise through the implementation of quantum error correcting codes [37–40]. Strides have been made in the integration of such algorithms within CV quantum optical systems, notably for tasks like variational eigenvalue determination [41]. Variational quantum algorithms have recently been implemented in the CV domain using single-mode squeezed light to optimize quantum sensing of an optical phase [42] and a measurement-based quantum alternative operator ansatz algorithm using a nonvariational quantum circuit [43].

Recent theoretical work has shown that CV quantum processes can be learned using a two-mode squeezed state (TMSS) as a quantum entanglement resource [26,44]. Two key features are made possible by the use of TMSS: (1) Only a single input-output pair is necessary for learning a single-mode unitary operation, and (2) the tunability of the degree of squeezing in a TMSS provides a mechanism to adaptively narrow or broaden the minimum of the cost landscape to aid training. Employing this universal method enables optimization of all parameters of a CV unitary operation, thus overcoming limitations of compiling in smaller dimensional subspaces, which may include reduced expressiveness and a constrained capacity for accurate representation of complex unitary transformations [45].

Here, we implement a CV quantum compiling algorithm (QCA) that learns the phase of an optical phase gate using a TMSS resource. Unlike standard squeezing-enhanced phase estimation—which infers a single unknown phase via interferometry using techniques such as Bayesian estimation of first and second moments [46–51]—in our experimental QCA the readout matches a variational phase shift to a target phase by minimizing a cost function based

on homodyne detection (HD). If viewed purely as a phase-estimation technique, our QCA belongs to an adaptive homodyne strategy. As a result, it has distinct advantages over direct quantum-enhanced estimation protocols: (i) it requires no analytic form of the generator beyond the ability to apply the unitary and (ii) it integrates seamlessly into larger circuit-compilation workflows. While direct estimation excels for a single, well-characterized generator, our variational approach provides a flexible, model-agnostic toolkit both for quantum compiling and parameter estimation in CV architectures. We show this in Appendix A, where we generalize the framework of our QCA to demonstrate that it can be extended to compiling multiparameter, multimode operations.

In this work we demonstrate the first experimental implementation of a squeezing-based CV variational quantum compiler capable of learning a target unitary parameter. We demonstrate that tuning the squeezing level of the quantum resource state of light can be used to control the width of gorges containing cost function minima. In particular, we show that at low squeezing, the QCA cost function exhibits a larger gradient farther from the minimum, enabling the QCA to more easily find the minimum about the phase of interest. As the squeezing increases, the gradient near the minimum becomes steeper, allowing the algorithm to train more quickly and learn the phase parameter to a higher degree of precision. Operation of the QCA in these squeezing domains allows us to circumvent both noise-induced and high-energy barren plateau phenomena in practice.

II. METHODS

A. QCA overview

In our implementation of a QCA for learning the phase of an optical phase gate, we employ a spatially truncated

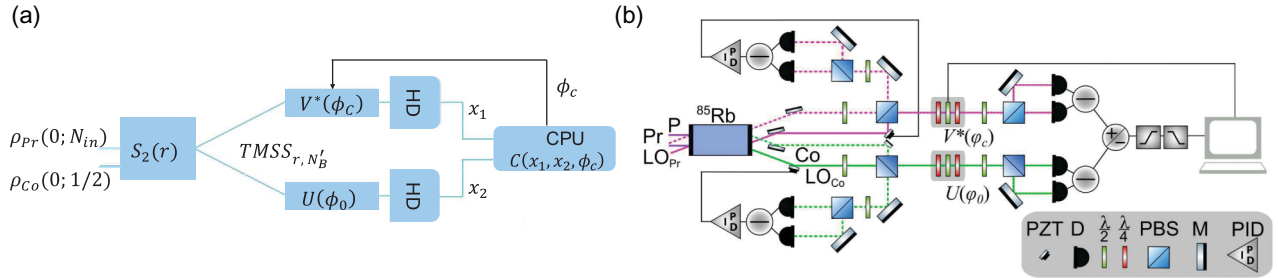


FIG. 1. (a) Schematic of a continuous-variable quantum compiler. A parametric amplifier is weakly seeded with a noisy probe beam modeled as a thermal state, ρ_{Pr} , with no displacement, i.e., $\rho_{Pr}(0; N'_{in}) = D(0)\rho_{Pr}(N'_{in})D^\dagger(0)$ where $D(\alpha)$ is the displacement operator and N'_{in} is the noise of the probe. The modes of the two-mode squeezed state interact with either the target, $U(\phi_0)$, or the control, $V^*(\phi_c)$, unitaries and are subsequently measured with homodyne detectors. These measurements are used to evaluate the cost function C , which is iteratively minimized using classical computing resources by adjusting the phase of the control unitary. (b) Experimental setup based on a truncated SU(1,1) interferometer used for the implementation of the CV quantum compiler. Pr: probe; Rb: Rubidium vapor cell; Co: conjugate; LO_{Pr}: local oscillator for the probe; LO_{Co}: local oscillator for the conjugate; PZT: piezoelectric transducer; D: detector; $\lambda/2$: half-wave plate; $\lambda/4$: quarter-wave plate; PBS: polarizing beam splitter; M: mirror; PID: proportional, integral, differential control.

version of the quantum compiler suggested in Ref. [26], as shown schematically in Fig. 1(a). Here, a noisy TMSS, with squeezing parameter r , is used as the quantum resource for the compilation. Such a noisy state allows us to take into account experimental imperfections and is obtained by taking one of the input modes (labeled probe) to be a thermal state with noise $\langle \Delta x_1^2 \rangle = \langle \Delta y_1^2 \rangle = N'_{in}$ for the amplitude quadrature, $q_1 = (a + a^\dagger)/\sqrt{2}$ with eigenvalue x_1 , and for the phase quadrature, $p_1 = i(a^\dagger - a)/\sqrt{2}$ with eigenvalue y_1 . The other input mode (labeled conjugate) is taken to be in the vacuum, such that it is shot noise limited. The modes of the TMSS interact with either a target, $U(\phi_0)$, or a control, $V^*(\phi_c)$, optical phase unitary with phases ϕ_0 and ϕ_c , respectively, before being measured. Given the focus on phase gates, a phase sensitive measurement, such as HD, is needed to measure the quadratures. The measurement results are then fed to a classical processor to evaluate a cost function, C , that needs to be minimized to perform the compilation.

During the QCA, the parameter of the target unitary, ϕ_0 , is held constant. Through successive iterations the QCA evaluates the cost function and performs a gradient descent to update the phase, ϕ_c , of the control unitary until the cost function has been minimized, at which point the QCA has learned the arbitrary phase shift, ϕ_0 , of the target unitary, $U(\phi_0)$. As we will see, the squeezing parameter, r , can be tuned to speed up the quantum compilation and increase the precision with which the parameter of the target unitary can be estimated.

B. QCA cost function structure

For the QCA to be successful, a proper cost function that depends on ϕ_0 and ϕ_c must be identified. Because our experiment measures homodyne data, we will define a cost function by considering the propagation of the Wigner function for a vacuum TMSS, $W(x_1, y_1, x_2, y_2)$, through the system shown in Fig. 1(a). We consider the two modes of the TMSS, with respective ladder operators $(a, a^\dagger)$ and $(b, b^\dagger)$, propagating through the corresponding phase gates, which introduce the local optical phase shifts $e^{i\phi_0 a^\dagger a} \otimes e^{-i\phi_c b^\dagger b}$. We model the seeding of the TMSS as a noisy state with input noise $N'_{in} = N'_B + 1/2$, where N'_B is the variance of the excess noise of the seed for the probe in our experiment, whereas the conjugate is not seeded $N_B = 0$. After propagation through these elements, the state can be described as a Gaussian state with zero mean and a

covariance matrix of the form

$$\Sigma = \frac{1 + 2\varepsilon N(1 + N'_B)}{2} I_2 \oplus \frac{1 + 2\varepsilon'(N(1 + N'_B) + N'_B)}{2} I_2 + \sqrt{\varepsilon\varepsilon'}\sqrt{N(N+1)}(1 + N'_B) \cos \Delta\phi X \otimes Z + \sqrt{\varepsilon\varepsilon'}\sqrt{N(N+1)}(1 + N'_B) \sin \Delta\phi X \otimes X, \quad (1)$$

where the quadratures are ordered (q_1, p_1, q_2, p_2) so that the correlations between the conjugate and probe modes are in the off-diagonal blocks, X and Z are Pauli matrices, $\Delta\phi := \phi_c - \phi_0$ is the phase difference between the control and target phase, $N = \sinh^2 r$ is the photon number of each of the modes of the vacuum TMSS, and the transmissivity parameters for the conjugate and probe are ε and ε' , respectively. Equation (1) can then be used to compute the Wigner function via the Fourier transform of the characteristic function [52].

Given the correlations that are present between the amplitude difference and phase sum quadratures in a TMSS, to obtain a cost function that can be efficiently evaluated through HD, we choose to project the Wigner function onto the amplitude difference quadrature, $(q_1 - q_2)/\sqrt{2}$. The distribution of such joint quadrature with eigenvalue $X_- = (x_1 - x_2)/\sqrt{2}$ is obtained via

$$\rho(X_-, r, \varepsilon, \varepsilon', N'_B, \Delta\phi) := \frac{1}{4} \int_{\mathbb{R}^3} dX_+ dy_1 dy_2 W(x_1, y_1, x_2, y_2), \quad (2)$$

where y_1 and y_2 are the eigenvalues of the phase quadratures of the respective modes and $X_+ = (x_1 + x_2)/\sqrt{2}$ is the eigenvalue of the quadrature sum, $(q_1 + q_2)/\sqrt{2}$.

Finally, to obtain a cost function that depends on $\Delta\phi$ and not on the measured eigenvalue, X_- , one can consider the zeroth order Taylor series expansion of $\rho(X_-, r, \varepsilon, \varepsilon', N'_B, \Delta\phi)$ in a small neighborhood around $X_- = 0$. We can then carry out the Gaussian integral in Eq. (2) to obtain the cost function

$$C(r, \varepsilon, \varepsilon', N'_B, \Delta\phi) := -\rho(0, r, \varepsilon, \varepsilon', N'_B, \Delta\phi) = -\sqrt{\frac{2}{\pi(\text{tr}A + 2A_{1,2})}}, \quad (3)$$

where

$$A(r, \varepsilon, \varepsilon', N'_B, \Delta\phi) := \begin{pmatrix} 1 + 2\varepsilon'(N'_B + (1 + N'_B)N) & -2\sqrt{\varepsilon\varepsilon'}(1 + N'_B)\sqrt{N(N+1)}\cos(\Delta\phi) \\ -2\sqrt{\varepsilon\varepsilon'}(1 + N'_B)\sqrt{N(N+1)}\cos(\Delta\phi) & 1 + 2\varepsilon(1 + N'_B)N \end{pmatrix} \quad (4)$$

is a matrix-valued function, with $A_{1,2}$ and $\text{tr}A$ representing the off-diagonal matrix element and the trace, respectively. See Supplemental Material [53], which includes Refs. [26, 54–68], for additional details. As we will see later, we find that the theoretical cost function closely approximates the experimentally obtained cost function.

Figure 2 shows the behavior of the cost function given in Eq. (3) as a function of $\Delta\phi$ for different squeezing parameters, r , for the ideal case of zero excess noise, $N'_B = 0$, and no optical losses, $\epsilon' = \epsilon = 1$ for the probe and conjugate. Crucially, the landscape of C depends on the squeezing parameter, r , with a narrowing of the cost minimum occurring around $\Delta\phi = 0$ as r increases. While the minimum of the cost function occurs at $\Delta\phi = 0$, in practice, there will be fluctuations around this minimum, which limit the precision with which the cost can be optimized and thus the precision with which the QCA compiles the parameter(s) of the target unitary. This precision is determined by the width of the cost function. To see why this is the case, one can again consider the ideal case of $\epsilon = \epsilon' = 1$ and $N'_B = 0$, i.e., the lossless noise-free case, and proceed by expanding the cost function (normalized by its global minimum value) about $\Delta\phi = 0$. This gives a quadratic scaling of the form

$$-1 + \frac{f(N)\Delta\phi^2}{2} + o(\Delta\phi^2), \quad (5)$$

$$f(N) := \sqrt{N(N+1)} \left[2\sqrt{N(N+1)} + (2N+1) \right], \quad (6)$$

which implies that success of the QCA (i.e., optimizing to within a neighborhood δ of the relative minimum value of the cost function) results in learning the phase with an error, as defined by the symbol $\Delta\phi^2$, that scales as the reciprocal of the energy squared (proportional to δ/N^2) of the squeezed light. This is similar to the precision obtained from direct phase detection with a Mach-Zehnder interferometer with squeezed light as input to one of its ports [69].

The general method of utilizing high-probability domains of a homodyne distribution to define a cost function for learning $\Delta\phi$ allows one to obtain an estimate of $\Delta\phi$ with a precision that depends on the nonclassicality of the probe state. To demonstrate this, we consider the phase precision achieved by such a learning algorithm when the probing state consists of a pure TMSS operation seeded with $|\alpha\rangle|\alpha\rangle$, coherent displacements of amplitude $\alpha \in \mathbb{R}$ in each mode. After the target and variational unitaries are applied to their respective mode, the (q_1, q_2) homodyne distribution is then given by

$$P(x_1, x_2) = \frac{1}{\pi \sqrt{\det A}} e^{-\frac{1}{\det A} (x-v)^T A (x-v)},$$

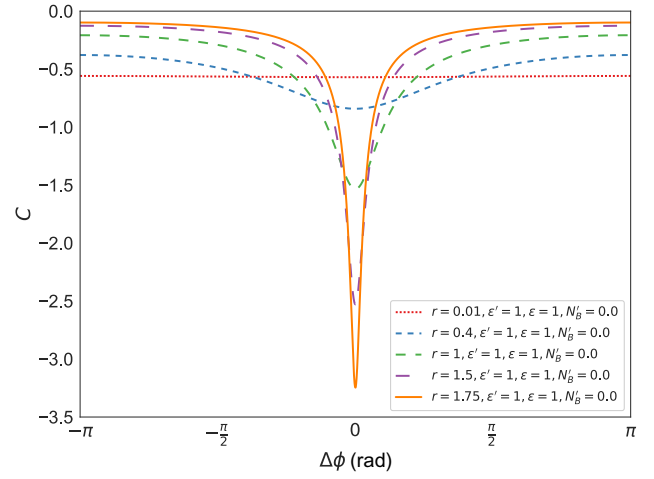


FIG. 2. Cost function landscape for different degrees of squeezing, for the ideal case of no excess noise ($N'_B = 0$) and no losses ($\epsilon' = \epsilon = 1$), as a function of the phase difference ($\Delta\phi = \phi_c - \phi_0$) between the target phase $U(\phi_0)$ and control phase $V^*(\phi_c)$ unitaries. The cost function directly measures how closely $V^*(\phi_c)$ approximates $U(\phi_0)$. We use the cost function to estimate $\Delta\phi$ in the QCA, which serves as a practical gauge of accuracy.

with

$$A = \begin{pmatrix} \cosh 2r & -\sinh(2r) \cos \Delta\phi \\ -\sinh(2r) \cos \Delta\phi & \cosh 2r \end{pmatrix}, \quad (7)$$

$$v = \begin{pmatrix} \sqrt{2}e^r \alpha \cos \phi_c \\ \sqrt{2}e^r \alpha \cos \phi_0 \end{pmatrix},$$

where we now consider the total energy per mode to be distributed between the amplified seed and squeezed vacuum according to $N = \alpha^2 e^{2r} + \sinh^2 r$. Defining a cost function by $-P(\sqrt{2}e^r \alpha, \sqrt{2}e^r \alpha)$, i.e., the most likely (q_1, q_2) -quadrature value occurring for the probe state, followed by normalizing to the global minimum value, gives a function of $\Delta\phi$ of the form

$$-1 + 2 \cosh^2 r \sinh^2 r \Delta\phi^2 + O(\alpha^2) e^{4r} \Delta\phi^4 + o(\Delta\phi^4), \quad (8)$$

where only the leading orders in α^2 and r are kept in the third term. The term of order $\Delta\phi^2$ results from the noise distribution of the TMSS, while the term of order $\Delta\phi^4$ is a result of the coherent seeding. In the weak-displacement limit ($\alpha^2 \rightarrow 0$), Eq. (8) reduces to the same quadratic scaling found in Eq. (5),

$$-1 + 2N^2 \Delta\phi^2 + o(\Delta\phi^4), \quad (9)$$

yielding $\Delta\phi^2 \propto \delta/N^2$ when the cost function is within δ of the minimum, thereby recovering Heisenberg scaling of

the precision, where here $N = \sinh^2 r$. In the opposite limit of negligible squeezing $r \rightarrow 0$, Eq. (8) simplifies to

$$-1 + \frac{N}{2} \Delta\phi^4 + o(\Delta\phi^4), \quad (10)$$

where $N = \alpha^2$ represents the energy of the coherent displacement. In this regime, the precision in $\Delta\phi$ obtained by calculating the cost to within δ of the global minimum scales as $\Delta\phi^2 \propto \sqrt{\delta/N}$, which is actually worse than standard quantum limit (SQL) scaling with respect to N . Achieving SQL scaling in $\Delta\phi$ would require more information from the homodyne distribution. We conclude that on the manifold of seeded TMSSs, the cost function-based readout of the phase difference interpolates between $O(N^{-1})$ scaling and $O(N^{-2})$ scaling of precision as the squeezing level increases. Note that the cost function given in Eq. (3) is defined in terms of the quadrature noise properties of the resource state and is not applicable to pure coherent states because it contains no information about $\Delta\phi$ in this case. In the setting of nonzero thermal noise and imperfect transmission, an analysis of the cost function is undertaken in Sec. VIII of the Supplemental Material [53].

Although in both Eqs. (5) and (8) the physical source of the precision advantage (namely quadrature squeezing) is the same as in homodyne readout of a Mach-Zehnder interferometer, the gradient-descent based algorithm learns ϕ_0 by tuning ϕ_c and calculating the cost function, as opposed to directly estimating the ϕ_0 phase shift. Our demonstration of this approach in Sec. II C shows that our algorithm is, in principle, capable of learning unknown target phase shift operations with precision comparable to Mach-Zehnder interferometry with squeezed light input. Furthermore, we show in Appendix A how this approach makes it possible to extend the QCA framework to compiling multiparameter, multimode operations.

One can additionally see from Fig. 2 that the expected cost function gradient in the neighborhood of the minimum increases concomitantly with the accuracy of the phase difference estimate as the squeezing parameter r increases, in agreement with the theoretical prediction in [26]. It follows that the choice of squeezing parameter, r (or equivalently the energy regime) plays a critical role in the time to solution of our variational QCA. For low squeezing values, e.g., $r = 0.01$ shown in red, the cost landscape is relatively flat for all $\Delta\phi$. In this case, given the inevitable reduced precision, as shown by Eq. (5), it will be challenging to discern the cost minimization direction, which will potentially prohibit training. For moderate squeezing values, e.g., $0.4 \leq r \leq 1.5$, the cost function exhibits substantial gradients over all phase difference values, thus making training easier. In the high-squeezing regime $r > 1.5$, once within the narrow gorge [70], the sharpness of the cost function will drive the optimization to a faster and more precise final estimate of the target unitary phase.

It is worth noting, however, that in the high-squeezing regime, the sharpness of the cost function makes it such that it approaches a constant exponentially fast for regions of the landscape away from the minimum. Thus, for a random initialization, this energy-dependent exponential flattening of the landscape would, with an exponentially high probability, increase the time to solution of the QCA, eventually precluding optimization altogether as the squeezing parameter r increases [26,71,72]. This is a manifestation of the *energy-dependent barren plateau* phenomenon.

Barren plateaus are a major barrier to the successful scaling of DV variational quantum algorithms [73], as their cost landscapes have been shown to flatten exponentially away from minima with the number of qubits [74–80]. The exponential flattening observed for our CV variational quantum compilation algorithm is crucially different in that it scales with squeezing parameter and not system size. Moreover, the potential offered by tuning the squeezing parameter, r , from low to high during training to avoid barren plateaus arguably makes the flattening and sharpening of the landscape with r a feature, rather than a barrier, to training.

C. Experimental implementation of a QCA

To demonstrate a CV quantum compiler based on the configuration shown in Fig. 1(a), we use a four-wave mixing (FWM) process in a double- Λ configuration in a ^{85}Rb vapor cell to generate a TMSS [81,82]. In this process two photons from a strong pump beam are absorbed to generate quantum correlated probe (Pr) and conjugate (Co) beams. We implement the FWM process with a continuous-wave Ti:Sapphire laser operating at 795 nm, which corresponds to the transition frequency of the D_1 line in ^{85}Rb ($5S_{1/2} \rightarrow 5P_{1/2}$). The pump beam is derived directly from the Ti:Sapphire laser and is tuned approximately 1 GHz to the blue of the D_1 transition to minimize losses while still preserving a strong nonlinear response. As described below, the exact value of this detuning is changed for different experimental runs to control the squeezing level of the TMSS generated by the FWM. The probe seed beam is generated by taking a small fraction of the pump beam and double passing it through a high frequency acousto-optic modulator to redshift it by 3.04 GHz, such that the pump and probe are close to two-photon resonance between the two ground state hyperfine levels. The pump beam, with a $1/e^2$ waist radius of $w_{0,\text{pump}} \approx 1.0$ mm and probe beam, with a $1/e^2$ waist radius of $w_{0,\text{probe}} \approx 0.50$ mm, intersect at a small angle $\theta \approx 0.4^\circ$ at the center of the vapor cell held at a temperature of 101 $^\circ\text{C}$.

As shown in Fig. 1(b), the FWM is part of a truncated $\text{SU}(1,1)$ interferometer [83] configuration with control and target unitaries placed in each of the arms of the interferometers to introduce the control (ϕ_c) and target (ϕ_0) phase shifts on the probe and conjugate, respectively.

Finally, homodyne detectors, one for each of the modes, are used to measure the quadratures required to evaluate the cost function of the quantum compiler circuit. For these measurements the phase of the respective local oscillator (LO) serves as a phase reference for the phase shift introduced by the control and target unitaries. We generate the required LOs with a second independent FWM process implemented in the same ^{85}Rb vapor cell [84].

The degree of squeezing of the TMSS, controlled by the squeezing parameter r , plays a vital role in ensuring the stability, precision, and efficiency of the quantum compiler. A critical aspect of our experiment is thus maintaining stable and precise control of r . To this end, we lock the different parameters of the FWM that determine r , specifically the frequency of the laser, the temperature of the ^{85}Rb vapor cell, and the power of the pump beam. The temperature of the cell and the power of the pump beam are stabilized through proportional-integral temperature and power locks, respectively. We stabilize the laser frequency with a side-band saturation spectroscopy lock using a reference ^{85}Rb vapor cell. The frequency lock allows us to control the one-photon detuning of the pump beam with respect to its corresponding transition. These locks make it possible to keep r constant during each QCA run. Likewise, stable power levels for the TMSS and LOs are essential for a reliable QCA run. We stabilize their powers using proportional-integral power locks applied to the TMSS and LO probe seed beams.

In our experiment, we adjust r by controlling the one-photon detuning of the pump beam, which also leads to a change in the gain of the FWM process. Given that the second FWM process used to generate the bright LOs is seeded with an input probe beam at a constant power of $145\text{ }\mu\text{W}$, the change in gain of the FWM leads to a change of the powers of the generated LOs. We consider three different values for the squeezing parameter, $r = 0.18, 0.35$, and 0.74 (see Sec. III), to study its impact on our CV QCA. At each degree of squeezing, the optical power of the local oscillator for the probe, LO_{Pr} , is attenuated to match the power of the generated conjugate local oscillator, LO_{Co} , with the power held constant at $17, 43$, and $330\text{ }\mu\text{W}$, respectively, during each QCA run, thereby ensuring that the overall amplification of the homodyne detector remains constant at each degree of squeezing.

After the FWM, we overlap the probe and conjugate with their respective LOs with a polarizing beam splitter (PBS), such that they are orthogonally polarized. We leverage the polarization degree of freedom to implement the unitary optical phase gates, $U(\phi_0)$ and $V^*(\phi_c)$, by transmitting the orthogonally polarized TMSS and LOs sequentially through a quarter-wave plate (QWP), a half-wave plate (HWP), and another QWP, as seen in Fig. 1(b). The initial QWP turns the linear orthogonally polarized fields into circular orthogonally polarized fields. The HWP

then introduces a relative phase shift between the left and right circularly polarized fields, with the degree of phase shift controlled by rotating the HWP. Finally, the last QWP converts the circularly polarized light back to linearly polarized light. This process can be visualized by a cyclic path on the Poincaré sphere that corresponds to the evolution of the light's polarization state [85,86]. This approach allows us to introduce a well-controlled and stable relative phase shift between the modes of the TMSS and their respective LOs.

After interacting with the target and control unitaries, the probe and conjugate are measured with HD. To implement the HD, the polarization of the probe (conjugate) and its local oscillator LO_{Pr} (LO_{Co}) are rotated by 45° with an HWP, such that they have orthogonal diagonal polarizations. After the HWP, a PBS is used to combine the fields, with the two outputs sent to a balanced photodetector system. This combination of HWP and PBS, along with the polarization rotation to orthogonal linear diagonal polarizations, effectively mimics combining fields with the same linear polarization using a 50:50 beam splitter.

In order to perform measurements that correspond to the amplitude difference quadrature (as defined by the phase reference set by the LOs), the relative phases between the TMSS and LOs need to be locked. To this end, we use a weak coherent state with a constant power of $3\text{ }\mu\text{W}$ to seed the input probe beam of the FWM process used to generate the TMSS. For this FWM process, the pump power was also held constant at 255 mW . To lock the relative phase between the TMSS and the LOs, we take advantage of the unused ports of the PBSs after the FWM process that are used to combine them. These PBSs reflect 1% of the TMSS and 50% of the LOs and are followed by a detection system with a configuration similar to the one used to implement the HD.

The output of the detection system directly provides an interference signal that serves as an error signal. A proportional-integral-differential (PID) controller is then used to lock the relative phase between the LO and the probe (conjugate) field to a chosen setpoint, by passing the control signal to a piezoelectric transducer (PZT) mounted on a mirror in the path of LO_{Pr} (LO_{Co}) before the LO is combined with the probe (conjugate) at the PBS, as seen in Fig. 1(b).

It is important to note that the implemented weak displacement of the TMSS by slightly seeding the FWM process does not affect the QCA or the evaluation of the cost function given that the cost function is evaluated based on fluctuations about the mean, as discussed in Sec. II B. For our implementation, the displacement will lead to the maximum of the X_- marginal of the Wigner function to be displaced from the origin; however, its profile, which determines the fluctuations about the mean, remains unchanged.

Once the initial relative phases between the TMSS and LOs are locked, the photocurrents of the measured quadratures after the transformations done by the target and control unitaries are combined with an RF hybrid junction to obtain the difference. The resulting signal is then passed through a bandpass filter with a full width at half maximum of 500 kHz and centered at 750 kHz to eliminate low-frequency laser noise and the displacement of the Wigner function introduced by weakly seeding the FWM. This is equivalent to performing a measurement with a spectrum analyzer with a 750 kHz analysis frequency and a resolution bandwidth of 500 kHz. Finally, the filtered time series data, $X_D(t)$, are digitized with an oscilloscope to accrue statistics to obtain the marginal of the Wigner function required to evaluate the cost function for the quantum compiler.

For the data analysis, the time series quadrature difference data, consisting of 1.2×10^8 samples, are processed by subtracting the mean voltage difference and scaling the result by the shot noise standard deviation corrected for electronic noise, thereby converting the raw voltage measurements into calibrated quadrature values and yielding a scaled dataset $X_{SD}(t) = X_D(t) / \sqrt{\sigma_{snl}^2 - \sigma_e^2}$, where σ_{snl}^2 and σ_e^2 represent the shot noise level and the electronic noise variances, respectively. To evaluate the cost function, a histogram of the rescaled quadrature data X_{SD} is fitted to a normal distribution to extract its variance, σ_{SD}^2 . Such a fit is justified given the Gaussian nature of the Wigner function. To properly account for the electronic noise, we subtract its contribution (using the same scaling as for the data) in quadrature, such that the variance of the marginal of the Wigner function is given by $\sigma^2 = \sigma_{SD}^2 - \sigma_e^2 / (\sigma_{snl}^2 - \sigma_e^2)$. The cost function C can then be easily evaluated by taking the maximum of the resulting normal distribution, which occurs at the origin. A full description of the data processing method is provided in Sec. I of the Supplemental Material [53].

III. RESULTS

In our quantum compiler methodology, we iteratively minimized the cost function derived from the marginal of the Wigner function, as outlined above. During each iteration, which is defined by a window of HD, we adjusted the phase of the control unitary to minimize the cost function, effectively learning the phase shift implemented by the target unitary. Our results show that in addition to the tunability of the cost function landscape, our QCA platform provides an advantage in speed and precision to efficiently and accurately learn a quantum unitary.

Figure 3(a) shows a comparison between the measured and theoretical cost functions as a function of $\Delta\phi$ for different squeezing parameters, r . The data show the narrowing of the cost function in the neighborhood around the

minimum as r increases, which will lead to the precision of the algorithm increasing with r . The agreement between the theoretical cost function and the experiment shows that our apparatus is faithful to the theoretical analysis of the cost function. The parameters in the theoretical cost function, given in Eqs. (3) and (4), were obtained using a maximum likelihood fit of C to the data. We reduce the number of fit parameters by directly measuring the degree of squeezing and losses after the FWM process. This leaves only two fitting parameters: the excess noise for the input probe seed, N'_B , and the probe transmissivity, ε' . Determining these parameters also allows us to extract the effective value of r for the FWM process. The details of our fitting procedure are discussed in detail in Sec. IV of the Supplemental Material [53].

We demonstrated the capabilities of our system by implementing the QCA for three different degrees of squeezing to examine the effect of narrowing the cost function. The QCA results are shown in Figs. 3(b) and 3(c). For all cases, the same initial value for the phase of the control unitary, ϕ_c , is used, such that the system is initialized with $\Delta\phi = 3$ rad. We ran the QCA over 550 iterations using gradient descent to determine the next required value of ϕ_c to optimize the cost function. In implementing the gradient descent, we used the same training rate ($\eta = 50$) for all degrees of squeezing to obtain a fair comparison. To determine the signed gradient of the cost function with respect to the control phase, we compute a finite-difference estimate of the gradient at each iteration using the two most recent control phase values, ϕ_{i-2} and ϕ_{i-1} , as outlined in Sec. VI of the Supplemental Material [53], thereby resolving the $\Delta\phi \leftrightarrow -\Delta\phi$ ambiguity inherent to the symmetric cost landscape.

Convergence in gradient-based methods occurs when small gradients near flat regions of a cost landscape near or at its minimum result in negligible parameter updates [87]. To ensure robustness against measurement noise near convergence, we adopt a noise-aware convergence and learning-rate strategy. Specifically, we declare convergence when the measured cost function reaches a value within one standard deviation of the noise-corrected theoretical minimum, i.e., $C(\Delta\phi) \leq C_{\min} + \sigma_{\text{noise}}$. The standard deviation of the cost, σ_{noise} , is computed from the scatter of the Fig. 3(a) cost values taken at phases $\Delta\phi$ very close to zero (for each r), i.e., the sample standard deviation of $C(\Delta\phi)$ within a small window around the minimum. Close to convergence, gradient estimates based on finite differences become unreliable, as the magnitude of the gradient approaches the noise floor $\sim \sigma_{\text{noise}}/\phi_\epsilon$, where ϕ_ϵ is the phase step between successive iterations. To prevent noise-driven excursions from the minimum, we reduce the learning rate η by a factor of 10 (to $\eta = 500$) once this threshold is reached, ensuring phase updates remain optimized so that the cost function is minimized. Similar noise-aware step-size annealing

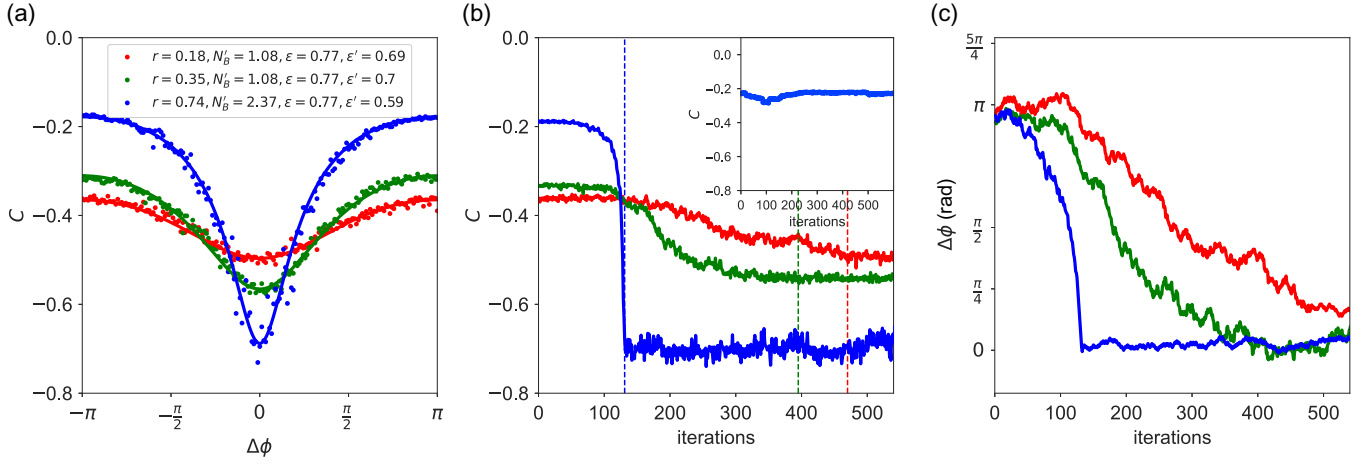


FIG. 3. (a) Theoretical (solid lines) and experimentally measured (dots) cost function based on the interaction of a two-mode squeezed state with the phase unitaries $U = e^{i\phi_0 a^\dagger a}$ and $V^* = e^{-i\phi_c b^\dagger b}$ for different values of r . (b) Measured cost function (C) as a function of the number of quantum compiling algorithm iterations. The dashed lines indicate the iteration at which the quantum compiling algorithm has converged for each r . The inset shows an instance of a barren plateau phenomenon occurring with the maximum degree of squeezing considered ($r = 0.74$). (c) Estimated phase difference ($\Delta\phi = \phi_c - \phi_0$) as a function of the number of quantum compiling algorithm iterations. For all figures red traces correspond to $r = 0.18$, green traces to $r = 0.35$, and blue traces to $r = 0.74$.

strategies are common in the stochastic optimization literature [88–90], where they are used to stabilize convergence near flat or noisy minima. In our experiment, once the cost function is within one standard deviation of the minimum, the stabilization of parameter estimates indicates practical convergence, with further progress limited by numerical precision.

We obtained a metric for the time-to-solution by evaluating the cost function as a function of the number of iterations of the QCA, as shown in Fig. 3(b). We observed a significant reduction in time-to-solution, t_{opt} , when comparing the QCA runs across increasing degrees of squeezing. The time-to-solution, t_{opt} , was defined as the number of iterations it took to reach convergence in Fig. 3(b) starting from the zeroth iteration, as indicated by the dashed vertical lines in the figure. The theoretical minimum used to establish convergence was empirically determined by taking an average of the cost function values about the center of the cost function where it was flat as shown in Fig. 3(a). We observed a factor of 3.6 speedup in time-to-solution to compile a stable parameterized unitary (i.e., to converge to a stable control phase parameter) measured as the ratio of t_{opt} between the $r = 0.74$ and $r = 0.18$ cases, and a factor of 4.55 reduction in the mean time-to-solution $\bar{t}_{\text{opt}}$ (i.e., the average of t_{opt} over the $\mathfrak{N} = 5$ runs; see Table I) between the same cases. Furthermore, the time-to-solution data shown in Table I indicate a nonlinear reduction in t_{opt} with respect to the squeezing parameter, tracking with the reduction in quadrature noise.

We next examine how the degree of squeezing influences the precision in estimating the phase parameter of the target unitary. To this end, we ran the QCA multiple times to determine the phase standard deviation in the estimation,

defined as the standard deviation, $\sigma_{\Delta\phi}$, of the mean phase differences over the $\mathfrak{N}$ runs. Here $\Delta\phi$ is the average over the phase differences within a single run after the QCA has converged. Due to the time it took to implement the 550 iterations of the QCA, which were necessary given the training rate for the gradient descent, we limited the number of runs to $\mathfrak{N} = 5$ for each of the different squeezing parameters. Figure 3(c) shows the estimated $\Delta\phi$ as a function of the number of QCA iterations for the different values of the squeezing parameter for one of the runs.

As shown by Eq. (5), the precision in the phase estimate for the compiled unitary, defined as $1/\sigma_{\Delta\phi}$, increased with increasing degree of squeezing. To obtain a measure of $\sigma_{\Delta\phi}$, for each of the five runs we obtained the mean over the phase differences after the QCA had converged, $\overline{\Delta\phi}$, using data such as that shown in Fig. 3(c). These five values were then used to calculate the phase difference mean

TABLE I. Measured time-to-solution (t_{opt}) for the quantum compiling algorithm runs reported in Fig. 3(b), the mean time-to-solution ($\bar{t}_{\text{opt}}$), phase difference mean ($\overline{\Delta\phi}$), and phase standard deviation $\sigma_{\Delta\phi}$ for different degrees of squeezing (r). The values of $\bar{t}_{\text{opt}}$, $\overline{\Delta\phi}$, and $\sigma_{\Delta\phi}$ are evaluated over the $\mathfrak{N} = 5$ runs. The numbers in parentheses are the corresponding standard deviations calculated using a Bayesian bootstrap approach, see Sec. VII of Supplemental Material [53].

r	t_{opt}	$\bar{t}_{\text{opt}}$	$\overline{\Delta\phi}$ (mrad)	$\sigma_{\Delta\phi}$ (mrad)
0.18	470	600 (54)	90	513.8 (76)
0.35	395	343 (18)	73	269.1 (35)
0.74	130	132 (1.5)	31	96 (12)

$\overline{\Delta\phi}$, which is the average across means of the phase differences, and the phase standard deviation $\sigma_{\Delta\phi}$, with the values reported in Table I. A detailed uncertainty analysis of the compiled phase precision and time-to-solution metrics is provided in Sec. VII of the Supplemental Material [53]. As can be seen, the precision increased in a nonlinear fashion consistent with the reciprocal of the reduction in quadrature noise. We show a factor of 5.4 improvement in the compiled phase precision between the highest and lowest values of r . We validated these results by performing a study using a larger number of runs ($\mathfrak{N} = 15$), albeit with a modified training rate for the gradient descent for a faster convergence of the QCA, as described in Sec. II of the Supplemental Material [53]. As shown in Table I, the phase difference mean decreased as the squeezing parameter, r , increased. The increase in r steepened the cost function and yielded larger gradients near the minimum, consistent with previous theoretical studies [26] that indicated that both the precision and the accuracy of the parameter estimation would increase because of the narrowing of the cost landscape [87]. Furthermore, the values obtained for $\overline{\Delta\phi}$ demonstrate that our measurements were unbiased as the difference between the expected phase difference mean, $\overline{\Delta\phi} = 0$, and the measured phase difference mean was negligible for all measured degrees of squeezing.

It is important to note that the enhancements obtained for the time-to-solution and precision for the implemented QCA were not dependent on the value of the phase for the target unitary, ϕ_0 , or the initial value used for the phase of the control unitary, ϕ_c . We showed that is the case by performing QCA runs where (i) the target unitary was set to arbitrary phases while the initial control unitary phase was fixed and (ii) the control unitary was set to an initial arbitrary phase while the target unitary phase was fixed (see Sec. V of the Supplemental Material [53]).

While for the data shown here the QCA converged for all degrees of squeezing considered, as pointed out in Sec. II B, the narrowing of the cost function occurs at the expense of the flattening of the landscape outside the gorge. Such a flattening can lead to barren plateaus, which have been shown theoretically to occur when squeezed resource states are used to define the cost function [26], a special case of the energy-dependent barren plateau in CV parametrized quantum circuits. In our experiments, we observed a few instances of failed convergence at the highest squeezing level tested ($r = 0.74$), where the gradient estimates remained near zero across multiple iterations. We empirically found that less than $\sim 1\%$ of the time our QCA failed to train completely. An example of one of the rare instances when our QCA failed to train is shown in the inset of Fig. 3(b). For the majority of other QCA runs with $r \leq 0.74$, including all the ones for lower squeezing values ($r = 0.18$ and 0.35), convergence was reliably

achieved. We conclude that, in practice, squeezing can be increased to the extent that it provides an enhanced QCA precision while remaining below an acceptable time-to-solution threshold.

The tunability of the cost function landscape of our CV QCA points to the ability to avoid barren plateaus through the implementation of an adaptive QCA in which the compiling is started with a low degree of squeezing to obtain an initial estimate of the parameter of the target unitary. Such an estimate would then serve to initialize the phase of the control unitary for a subsequent compiling with a higher degree of squeezing. Compared to random phase initialization for each compiling run, an adaptive approach is capable of counteracting the energy-dependent flattening of the cost landscape, allowing to maximize precision and minimize time-to-solution.

While our experimental implementation focuses on learning a single parameter unitary, it can be extended to higher-dimensional compilation tasks. To illustrate this, we present in Appendix A an extension of the same SU(1,1)-based variational ansatz, cost-function construction, and homodyne-based optimization protocol to compile a broader class of linear optical unitaries. In particular, we show that for a two-parameter generalized beamsplitter gate, a faithful cost function can still be obtained via HD with an extended two-copy TMSS resource state. This measurement-induced cost preserves key features of our method—experimental tractability, strong gradients near the minimum, and noise-aware tunability—while providing access to a higher-dimensional parameter space. We also analyze the structure of the cost landscape in the ideal lossless case and verify Heisenberg-like scaling near the minimum, thereby confirming that our experimental demonstration constitutes a scalable, resource-efficient proof-of-principle demonstration of CV quantum compiling.

IV. DISCUSSION

To the best of our knowledge, our work presents the first experimental implementation of a squeezing-based variational quantum algorithm for quantum compiling on noisy intermediate-scale quantum hardware in the CV setting. Our proof-of-principle experimental implementation demonstrates the ability of a CV quantum compiler to learn the phase parameter of a target optical phase gate using gradient descent optimization.

To implement the CV QCA, we leveraged an optical TMSS as a quantum resource in our quantum compiler and showed, both theoretically and experimentally, the narrowing of the cost function around $\Delta\phi = 0$ as the degree of squeezing increases. We also showed that such a narrowing and associated increased slope led to a reduction in time-to-solution and increased precision in the estimation of the target phase. In particular, we showed that our

CV quantum compiler framework gave a 3.6-fold speedup in the time-to-solution and increased the precision by a factor of 5.4. This tunability of our CV QCA can enable the implementation of an adaptive QCA to avoid barren plateaus while maximizing precision and minimizing time-to-solution.

An important question is whether these results signify a quantum enhancement over a classical system of equal energy (e.g., a coherent state). Increasing r raises the total photon number $N = \sinh^2(r)$, which narrows the neighborhood of the cost-function minimum. Our analysis in Eq. (5) demonstrates that squeezing dramatically steepens the cost function slope near the neighborhood of the minimum, allowing faster convergence to a more precise phase difference estimate. We also show that our cost function-based readout allows one to achieve phase difference estimation precision scaling as $O(E^{-a})$, $1 < a < 2$, which is enabled by the two-mode squeezing in our CV quantum compiler. A detailed analysis of the experimental scaling with respect to the total detected energy E is provided in Secs. VIII and IX of the Supplemental Material [53]. Although losses and excess noise weaken this behavior, our experimentally observed speedup and precision improvement remain robust under realistic conditions. These findings show that the quantum correlations in the squeezed resource—not merely the total photon number—drive the enhanced performance, thereby substantiating a quantum enhancement in our CV quantum compiler.

Our results represent the first step toward the implementation of a universal quantum compiler capable of compiling any CV circuit [26]. Such a universal quantum compiler requires an inverse squeezing operation after the target and control unitaries and enables the compilation of all CV unitaries on M -modes through a multimode SU(1,1) configuration. A nonuniversal quantum compiler for linear optical CV or, more generally, Gaussian CV circuits, is an important benchmark for optimization of quantum resources on near-term CV hardware, and provides a path forward for the minimization of quantum circuit size and depth in CV quantum algorithms. These developments are critical for the realization of both near-term and fault-tolerant CV quantum computing systems as well as for the characterization of dynamics in quantum systems, such as those of novel quantum materials.

ACKNOWLEDGMENTS

The authors extend their gratitude for the support provided by the U.S. Department of Energy Office of Science, along with the National Quantum Information Science Research Centers and the Quantum Science Center, under the identifier LA-UR-24-33284.

This manuscript has been authored, in part, by UT-Battelle LLC, under contract DE-AC05-00OR22725 with

the U.S. Department of Energy (DOE). The publisher acknowledges the U.S. government license to provide public access under the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX: COMPILING MULTIMODE LINEAR OPTICS WITH A COST FUNCTION BASED ON HOMODYNE READOUT

Analogous to the Bell measurements carried out in CV teleportation [91], the Bell measurements that define the universal cost function in Ref. [26] can be achieved by linear optics and homodyne measurement. For full generality, the homodyne measurements should involve two quadrature directions per TMSS resource used (i.e., per optical mode in the compilation problem). Although our cost function for learning an optical phase, which is constructed from high-probability outcomes of a single homodyne measurement, is not universal for compiling a general single-mode CV unitary, the basic strategy of defining cost functions based on a minimal number of readily implemented measurements on TMSS resource states has wide applicability even for multimode, multiparameter compiling problems. The reason for this is that such a strategy only differs from the universal CV compiling method of Ref. [26] by replacing the CV Bell measurement with a simpler measurement while maintaining faithfulness on the submanifold of interest for a given compiling problem. For certain submanifolds of unitaries that one may seek to compile (e.g., linear optical unitaries), an appropriately chosen measurement can still give a faithful signal.

In Fig. 4, we show a homodyne-measurement circuit that leads to a faithful cost function for target unitaries consisting of the two-parameter generalized beamsplitter,

$$U(\theta, \varphi) := e^{\theta(e^{i\varphi} a^\dagger b - e^{-i\varphi} b^\dagger a)}. \quad (\text{A1})$$

The variational ansatz is a depth-3 circuit consisting of the phase shift $P(\phi) = e^{-i\phi a^\dagger a}$ and a real beamsplitter. Note from Ref. [26] that in this configuration, the ansatz circuit should be implemented with complex conjugate gates due to the approximate Ricochet property of the two-mode squeezed states. We consider the resource state to be two copies of the TMSS vacuum state with squeezing parameter r ,

$$|\text{TMSS}_r\rangle_{13} |\text{TMSS}_r\rangle_{24}. \quad (\text{A2})$$

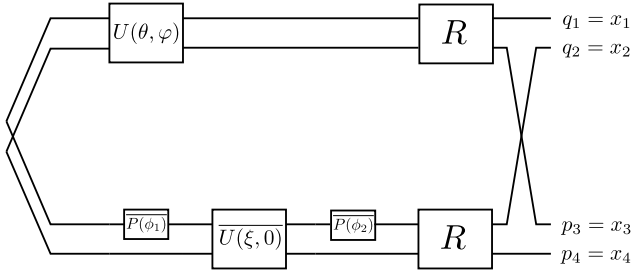


FIG. 4. Circuit for compiling a generalized beamsplitter with homodyne readout. The variational circuit on modes 3 and 4 has three real parameters ϕ_1 , ϕ_2 , and ξ .

For large squeezing r , the two-mode squeezed state resource converges to a zero eigenvector of the operators

$$\frac{q_1 - q_3}{\sqrt{2}}, \frac{p_2 + p_4}{\sqrt{2}}, \frac{q_2 - q_4}{\sqrt{2}}, \frac{p_1 + p_3}{\sqrt{2}}, \quad (\text{A3})$$

which can be considered as approximate nullifiers of Eq. (A2). The operators

$$\frac{q_3 - q_1}{\sqrt{2}}, \frac{p_2 + p_4}{\sqrt{2}}, \frac{q_4 - q_2}{\sqrt{2}}, \frac{p_3 + p_1}{\sqrt{2}} \quad (\text{A4})$$

are also valid approximate nullifiers due to the permutation symmetry of the modes of Eq. (A2) under $1 \leftrightarrow 3$, $2 \leftrightarrow 4$. By selecting either of these homodyne measurements (depending on the selected form for the 50:50 beamsplitter R in Fig. 4), the outcome $x_1 = x_2 = x_3 = x_4 = 0$ mimics the full contraction with $\langle \text{TMSS}_r |_{13} \langle \text{TMSS}_r |_{24}$ carried out in the universal cost function of Ref. [26].

The full three parameter normalized cost function is obtained from Fig. 4 as follows: 1. the target unitary and variational circuit are applied to their respective modes to the state defined in Eq. (A2), resulting in a Gaussian state with covariance matrix Σ , 2. the homodyne measurement given in Eq. (A3) is equivalent to a local homodyne measurement of q_1 , q_2 , p_3 , and p_4 under a linear optical transformation given by the last circuit elements in Fig. 4, which can be implemented by covariance matrix transformation $\Sigma \mapsto O^\top \Sigma O$ for orthogonal symplectic O , 3. the $\{1, 3, 6, 8\}$ principal minor of $O^\top \Sigma O$ defines the normally distributed homodyne outcomes, and the probability density at zero is taken. For the sake of obtaining a cost function with range $[-1, 0]$, we normalize this function by its value at parameters corresponding to the target unitary. The resulting three-variable function is difficult to visualize, so in Fig. 5 we show the two-dimensional submanifold defined by $\phi := \phi_1 = -\phi_2$ and ξ for a target unitary defined by $\theta = \pi/4$ and $\varphi = 0$. The faithful compiling points $(\xi, \phi) = (\pi/4, 0)$ and $(\xi, \phi) = (-\pi/4, \pi)$ are clearly visible. For larger squeezing levels, the energy-dependent barren plateau becomes evident. One can verify by expanding ξ about $\pi/4$ and ϕ about 0 that the quadratic

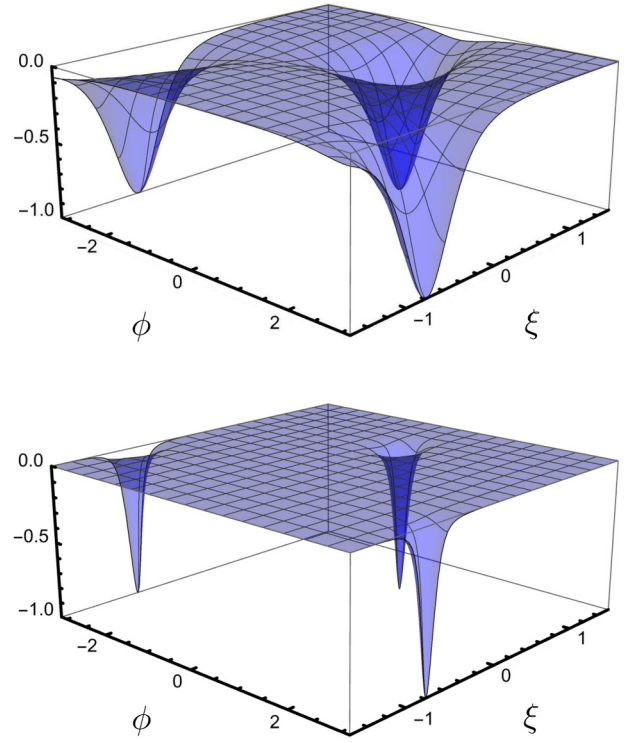


FIG. 5. Normalized cost function on the submanifold parametrized by ξ and ϕ . The squeezing levels are $r = 3/4$ (top) and $r = \sqrt{2}$ (bottom).

behavior in a neighborhood of the global minima has the form

$$-1 + O(N^2) \left(\xi - \frac{\pi}{4} \right)^2 + O(N^2) \phi^2 + o(\phi^2) + o\left(\left(\xi - \frac{\pi}{4} \right)^2 \right), \quad (\text{A5})$$

indicating Heisenberg scaling for both parameters.

-
- [1] Maria Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, *Contemp. Phys.* **56**, 172 (2015).
 - [2] Y. Zhang and Q. Ni, Recent advances in quantum machine learning, *Quantum Eng.* **2**, e34 (2020).
 - [3] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, Quantum-assisted quantum compiling, *Quantum* **3**, 140 (2019).
 - [4] K. Heya, Y. Suzuki, Y. Nakamura, and K. Fujii, Variational quantum gate optimization, [arXiv:1810.12745](https://arxiv.org/abs/1810.12745).
 - [5] K. Poland, K. Beer, and T. J. Osborne, No free lunch for quantum machine learning, [arXiv:2003.14103](https://arxiv.org/abs/2003.14103).
 - [6] K. Sharma, M. Cerezo, Z. Holmes, L. Cincio, A. Sornborger, and P. J. Coles, Reformulation of the no-free-lunch theorem for entangled datasets, *Phys. Rev. Lett.* **128**, 070501 (2022).

- [7] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, Continuous-variable quantum neural networks, *Phys. Rev. Res.* **1**, 033063 (2019).
- [8] C. Cirstoiu, Z. Holmes, J. Iosue, L. Cincio, P. J. Coles, and A. Sornborger, Variational fast forwarding for quantum simulation beyond the coherence time, *Npj Quantum Inf.* **6**, 1 (2020).
- [9] J. Gibbs, K. Gili, Z. Holmes, B. Commeau, A. Arrasmith, L. Cincio, P. J. Coles, and A. Sornborger, Long-time simulations for fixed input states on quantum hardware, *Npj Quantum Inf.* **8**, 135 (2022).
- [10] J. Gibbs, Z. Holmes, M. C. Caro, N. Ezzell, H.-Y. Huang, L. Cincio, A. T. Sornborger, and P. J. Coles, Dynamical simulation via quantum machine learning with provable generalization, *Phys. Rev. Res.* **6**, 013241 (2024).
- [11] M. R. Geller, Z. Holmes, P. J. Coles, and A. Sornborger, Experimental quantum learning of a spectral decomposition, *Phys. Rev. Res.* **3**, 033200 (2021).
- [12] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, Quantum advantage in learning from experiments, *Science* **376**, 1182 (2022).
- [13] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, Generalization in quantum machine learning from few training data, *Nat. Commun.* **13**, 4919 (2022).
- [14] M. C. Caro, H.-Y. Huang, N. Ezzell, J. Gibbs, A. T. Sornborger, L. Cincio, P. J. Coles, and Z. Holmes, Out-of-distribution generalization for learning quantum dynamics, *Nat. Commun.* **14**, 3751 (2023).
- [15] H.-Y. Huang, S. Chen, and J. Preskill, Learning to predict arbitrary quantum processes, *PRX Quantum* **4**, 040337 (2023).
- [16] M. C. Caro, Learning quantum processes and Hamiltonians via the Pauli transfer matrix, *ACM Trans. Quantum Comput.* **5**, 1 (2024).
- [17] S. Jerbi, J. Gibbs, M. S. Rudolph, M. C. Caro, P. J. Coles, H.-Y. Huang, and Z. Holmes, The power and limitations of learning quantum dynamics incoherently, *arXiv:2303.12834*.
- [18] V. T. Hai and L. B. Ho, Universal compilation for quantum state tomography, *Sci. Rep.* **13**, 3750 (2023).
- [19] T. Haug and M. S. Kim, Generalization of quantum machine learning models using quantum fisher information metric, *Phys. Rev. Lett.* **133**, 050603 (2024).
- [20] R. Kautbruegger, P. Silvi, C. Kokail, R. van Bijnen, A. M. Rey, J. Ye, A. M. Kaufman, and P. Zoller, Variational spin-squeezing algorithms on programmable quantum sensors, *Phys. Rev. Lett.* **123**, 260505 (2019).
- [21] R. Kautbruegger, A. Shankar, D. V. Vasilyev, and P. Zoller, Optimal and variational multiparameter quantum metrology and vector-field sensing, *PRX Quantum* **4**, 020333 (2023).
- [22] C. Huerta Alderete, M. H. Gordon, F. Sauvage, A. Sone, A. T. Sornborger, P. J. Coles, and M. Cerezo, Inference-based quantum sensing, *Phys. Rev. Lett.* **129**, 190501 (2022).
- [23] F. A. Mele, A. A. Mele, L. Bittel, J. Eisert, V. Giovannetti, L. Lami, L. Leone, and S. F. E. Oliviero, Learning quantum states of continuous variable systems, *Nat. Phys.* **21**, 2002 (2025).
- [24] V. T. Hai, N. T. Viet, and L. B. Ho, Variational preparation of entangled states on quantum computers, *arXiv:2306.17422*.
- [25] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, Learning the quantum algorithm for state overlap, *New J. Phys.* **20**, 113022 (2018).
- [26] T. Volkoff, Z. Holmes, and A. Sornborger, Universal compiling and (No-)Free-Lunch theorems for continuous-variable quantum learning, *PRX Quantum* **2**, 040327 (2021).
- [27] T. Volkoff and A. T. Sornborger, Learning linear optical circuits with coherent states, *J. Phys. A: Math. Theor.* **57**, 305302 (2024).
- [28] T. J. Volkoff, Strategies for variational quantum compiling of a zero-phase beam splitter on the Xanadu X8 processor, *J. Russ. Laser Res.* **43**, 448 (2022).
- [29] T. Jones and S. C. Benjamin, Robust quantum compilation and circuit optimisation via energy minimisation, *Quantum* **6**, 628 (2022).
- [30] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya, M. Menotti, I. Dhand, Z. Vernon, N. Quesada, and J. Lavoie, Quantum computational advantage with a programmable photonic processor, *Nature* **606**, 75 (2022).
- [31] M. Schuld and N. Killoran, Quantum machine learning in feature Hilbert spaces, *Phys. Rev. Lett.* **122**, 040504 (2019).
- [32] B. J. Lawrie, P. D. Lett, A. M. Marino, and R. C. Pooser, Quantum sensing with squeezed light, *ACS Photonics* **6**, 1307 (2019).
- [33] C. Hua, C. E. Marvinney, S. Hong, M. Feldman, Y.-Y. Pai, M. Chilcote, J. Rabinowitz, R. C. Pooser, A. M. Marino, and B. J. Lawrie, Quantum enhanced probes of magnetic circular dichroism, *Adv. Quantum Technol.* **8**, 2300126 (2025).
- [34] X. Guo, C. R. Breum, J. Borregaard, S. Izumi, M. V. Larsen, T. Gehring, M. Christandl, J. S. Neergaard-Nielsen, and U. L. Andersen, Distributed quantum sensing in a continuous-variable entangled network, *Nat. Phys.* **16**, 281 (2020).
- [35] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Deterministic generation of a two-dimensional cluster state, *Science* **366**, 369 (2019).
- [36] W. Asavanant, Y. Shiozawa, S. Yokoyama, B. Charoen-sombutamon, H. Emura, R. N. Alexander, S. Takeda, J. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, Generation of time-domain-multiplexed two-dimensional cluster state, *Science* **366**, 373 (2019).
- [37] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, *Phys. Rev. A* **64**, 012310 (2001).
- [38] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, Quantum computing with rotation-symmetric bosonic codes, *Phys. Rev. X* **10**, 011058 (2020).
- [39] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Performance and structure of single-mode bosonic codes, *Phys. Rev. A* **97**, 032346 (2018).
- [40] T. J. Volkoff, Nonclassical properties and quantum resources of hierarchical photonic superposition states, *J. Exp. Theor. Phys.* **121**, 770 (2015).

- [41] H.-K. Lau, R. Pooser, G. Siopsis, and C. Weedbrook, Quantum machine learning over infinite dimensions, *Phys. Rev. Lett.* **118**, 080501 (2017).
- [42] J. A. H. Nielsen, M. J. Kicinski, T. N. Arge, K. Vijayadharan, J. Foldager, J. Borregaard, J. J. Meyer, J. S. Neergaard-Nielsen, T. Gehring, and U. L. Andersen, Variational quantum algorithm for enhanced continuous variable optical phase sensing, *Npj Quantum Inf.* **11**, 70 (2025).
- [43] Y. Enomoto, K. Anai, K. Udagawa, and S. Takeda, Continuous-variable quantum approximate optimization on a programmable photonic quantum processor, *Phys. Rev. Res.* **5**, 043005 (2023).
- [44] C. Oh, S. Chen, Y. Wong, S. Zhou, H. Y. Huang, Jens A. H. Nielsen, Z. H. Liu, J. S. Neergaard-Nielsen, U. L. Andersen, L. Jiang, and J. Preskill, Entanglement-enabled advantage for learning a bosonic random displacement channel, *Phys. Rev. Lett.* **133**, 230604 (2024).
- [45] J. M. Arrazola, T. R. Bromley, J. Izaac, C. R. Myers, K. Bradler, and N. Killoran, Machine learning method for state preparation and gate synthesis on photonic quantum computers, *Quantum Sci. Technol.* **4**, 024004 (2019).
- [46] H. M. Wiseman, Adaptive phase measurements of optical modes: Going beyond the marginal q distribution, *Phys. Rev. Lett.* **75**, 4587 (1995).
- [47] M. A. Rodrıguez-Garcıa and F. E. Becerra, Adaptive phase estimation with squeezed vacuum approaching the quantum limit, *Quantum* **8**, 1480 (2024).
- [48] S. Ataman, Optimal Mach-Zehnder phase sensitivity with Gaussian states, *Phys. Rev. A* **100**, 063821 (2019).
- [49] L. Pezze and A. Smerzi, Mach-Zehnder interferometry at the Heisenberg limit with coherent and squeezed-vacuum light, *Phys. Rev. Lett.* **100**, 073601 (2008).
- [50] J. Yu, Y. Qin, J.-L. Qin, Z.-H. Yan, and X.-J. Jia, Quantum phase estimation with a stable squeezed state, *Eur. Phys. J. D* **74**, 76 (2020).
- [51] J. Yu, Y. Qin, J. Qin, H. Wang, Z. Yan, X. Jia, and K. Peng, Quantum enhanced optical phase estimation with a squeezed thermal state, *Phys. Rev. Appl.* **13**, 024037 (2020).
- [52] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (CRC Press, Boca Raton, Florida, 2017).
- [53] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/718z-zn2m> which includes additional theoretical derivations of the cost function, analysis of the cost landscape under thermal noise and loss, details of the experimental model fitting procedure, numerical/algorithmic details of gradient estimation and optimization, and a Bayesian bootstrap uncertainty analysis.
- [54] K. Zhang, S. Liu, Y. Chen, X. Wang, and J. Jing, Optical quantum states based on hot atomic ensembles and their applications, *Photonics Insights* **1**, R06 (2022).
- [55] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).
- [56] S. Becker, N. Datta, L. Lami, and C. RouzLe, Energy-constrained discrimination of unitaries, quantum speed limits, and a Gaussian Solovay-Kitaev theorem, *Phys. Rev. Lett.* **126**, 190504 (2021).
- [57] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 2007), 3rd ed., pp. 807–817.
- [58] M. Jasperse, L. D. Turner, and R. E. Scholten, Relative intensity squeezing by four-wave mixing with loss: An analytic model and experimental diagnostic, *Opt. Express* **19**, 3765 (2011).
- [59] M. T. Turnbull, P. G. Petrov, C. S. Embrey, A. M. Marino, and V. Boyer, Role of the phase-matching condition in nondegenerate four-wave mixing in hot vapors for the generation of squeezed states of light, *Phys. Rev. A* **88**, 033845 (2013).
- [60] P. Gupta, R. W. Speirs, K. M. Jones, and P. D. Lett, Effect of imperfect homodyne visibility on multi-spatial-mode two-mode squeezing measurements, *Opt. Express* **28**, 652 (2020).
- [61] K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto, and K. Nakamura, Accuracy of quantum-state estimation utilizing Akaike’s information criterion, *Phys. Rev. A* **68**, 022314 (2003).
- [62] H. Yano and N. Yamamoto, Quantum information criteria for model selection in quantum state estimation, *J. Phys. A: Math. Theor.* **56**, 405301 (2023).
- [63] P. Lougovski and S. J. van Enk, Characterizing entanglement sources, *Phys. Rev. A* **80**, 052324 (2009).
- [64] J. O. S. Yin and S. J. van Enk, Information criteria for efficient quantum state estimation, *Phys. Rev. A* **83**, 062110 (2011).
- [65] A. M. Vadiraj, A. Ask, T. G. McConkey, I. Nsanzineza, C. W. S. Chang, A. F. Kockum, and C. M. Wilson, Engineering the level structure of a giant artificial atom in waveguide quantum electrodynamics, *Phys. Rev. A* **103**, 023710 (2021).
- [66] S. J. van Enk and R. Blume-Kohout, When quantum tomography goes wrong: drift of quantum sources and other errors, *New J. Phys.* **15**, 025024 (2013).
- [67] Q.-C. Liu, T.-F. Li, X.-Q. Luo, H. Zhao, W. Xiong, Y.-S. Zhang, Z. Chen, J. S. Liu, W. Chen, F. Nori, J. S. Tsai, and J. Q. You, Method for identifying electromagnetically induced transparency in a tunable circuit quantum electrodynamics system, *Phys. Rev. A* **93**, 053838 (2016).
- [68] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer Science & Business Media, New York, NY, 2002), 2nd ed.
- [69] C. M. Caves, Quantum-mechanical noise in an interferometer, *Phys. Rev. D* **23**, 1693 (1981).
- [70] A. Arrasmith, Z. Holmes, M. Cerezo, and P. J. Coles, Equivalence of quantum barren plateaus to cost concentration and narrow gorges, *Quantum Sci. Technol.* **7**, 045015 (2022).
- [71] T. J. Volkoff, Efficient trainability of linear optical modules in quantum optical neural networks, *J. Russ. Laser Res.* **42**, 250 (2021).
- [72] B. Zhang and Q. Zhuang, Energy-dependent barren plateau in bosonic variational quantum circuits, *Quantum Sci. Technol.* **10**, 015009 (2024).
- [73] M. Cerezo, M. Larocca, D. Garcıa-Martı, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability? or,

- why we need to rethink variational quantum computing, *Nat. Commun.* **16**, 7907 (2025).
- [74] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [75] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, *PRX Quantum* **3**, 010313 (2022).
- [76] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, *Nat. Commun.* **12**, 1 (2021).
- [77] C. Ortiz Marrero, M. Kieferová, and N. Wiebe, Entanglement-induced barren plateaus, *PRX Quantum* **2**, 040316 (2021).
- [78] Z. Holmes, A. Arrasmith, B. Yan, P. J. Coles, A. Albrecht, and A. T. Sornborger, Barren plateaus preclude learning scramblers, *Phys. Rev. Lett.* **126**, 190501 (2021).
- [79] E. Fontana, D. Herman, S. Chakrabarti, N. Kumar, R. Yalovetzky, J. Heredge, S. H. Sureshbabu, and M. Pistola, Characterizing barren plateaus in quantum ansätze with the adjoint representation, *Nat. Commun.* **15**, 7171 (2024).
- [80] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. Ortiz Marrero, M. Larocca, and M. Cerezo, A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits, *Nat. Commun.* **15**, 7172 (2024).
- [81] C. F. McCormick, V. Boyer, E. Arimondo, and P. D. Lett, Strong relative intensity squeezing by four-wave mixing in rubidium vapor, *Opt. Lett.* **32**, 178 (2007).
- [82] C. F. McCormick, A. M. Marino, V. Boyer, and P. D. Lett, Strong low-frequency quantum correlations from a four-wave-mixing amplifier, *Phys. Rev. A* **78**, 043816 (2008).
- [83] B. E. Anderson, P. Gupta, B. L. Schmittberger, T. Horrom, C. Hermann-Avigliano, K. M. Jones, and P. D. Lett, Phase sensing beyond the standard quantum limit with a variation on the SU(1,1) interferometer, *Optica* **4**, 752 (2017).
- [84] V. Boyer, A. M. Marino, R. C. Pooser, and P. D. Lett, Entangled images from four-wave mixing, *Science* **321**, 544 (2008).
- [85] P. Hariharan and M. Roy, A geometric-phase interferometer, *J. Mod. Opt.* **39**, 1811 (1992).
- [86] J. Brendel, W. Dultz, and W. Martienssen, Geometric phases in two-photon interference experiments, *Phys. Rev. A* **52**, 2551 (1995).
- [87] J. Nocedal and S. Wright, *Numerical Optimization* (Springer, New York, New York, 2006).
- [88] L. Armijo, Minimization of functions having Lipschitz continuous first partial derivatives, *Pacific J. Math.* **16**, 1 (1966).
- [89] J. Duchi, E. Hazan, and Y. Singer, Adaptive subgradient methods for online learning and stochastic optimization, *J. Mach. Learn. Res.* **12**, 2121 (2011).
- [90] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework, *Phys. Rev. A* **84**, 022305 (2011).
- [91] S. L. Braunstein and H. J. Kimble, Teleportation of continuous quantum variables, *Phys. Rev. Lett.* **80**, 869 (1998).