Non-convex Quadratic Programming Using Coherent Optical Networks

Farhad Khosravi, Uğur Yıldız, Artur Scherer, and Pooya Ronagh

1QB Information Technologies (1QBit), Vancouver, BC, Canada
2Institute for Quantum Computing, University of Waterloo, Waterloo, ON, Canada
3Department of Physics & Astronomy, University of Waterloo, Waterloo, ON, Canada
4Perimeter Institute for Theoretical Physics, Waterloo, ON, Canada

(Dated: March 10, 2023)

We investigate the possibility of solving continuous non-convex optimization problems using a network of interacting quantum optical oscillators. We propose a native encoding of continuous variables in analog signals associated with the quadrature operators of a set of quantum optical modes. Optical coupling of the modes and noise introduced by vacuum fluctuations from external reservoirs or by weak measurements of the modes are used to optically simulate a diffusion process on a set of continuous random variables. The process is run sufficiently long for it to relax into the steady state of an energy potential defined on a continuous domain. As a first demonstration, we numerically benchmark solving box-constrained quadratic programming (BoxQP) problems using these settings. We consider delay-line and measurement-feedback variants of the experiment. Our benchmarking results demonstrate that in both cases the optical network is capable of solving BoxQP problems over three orders of magnitude faster than a state-of-the-art classical heuristic.

I. INTRODUCTION

Real-world optimization problems frequently involve decision variables that are naturally continuous. This is the case for continuous non-convex global optimization problems and the more general class of mixed-integer programming (MIP) models [1] designed to solve a variety of NP-hard optimization problems of high practical relevance, including, for example, scheduling problems, flow problems, resource allocation problems, and routing problems. Even the special class of quadratic non-convex global optimization problems captures extremely challenging computational tasks that are NP-hard [2, 3]. Classical optimization techniques lack satisfactory performance in solving continuous non-convex optimization problems at scale [3–5].

In recent years, the design of special-purpose computing architectures for solving optimization problems has been gaining increasing attention [6]. Common discrete optimization engines include physics-inspired heuristics such as thermal (i.e., simulated) annealing [7, 8] and its variants [9, 10], quantum annealing [11–14], coherent Ising machines (CIM) [15, 16], local heuristic search techniques [17, 18], and various spin-based Monte Carlo methods. All these approaches focus on solving discrete optimization problems natively, and even in that class, the attention has been limited to solving Ising problems. This means that more-general types of problems (e.g., those comprising continuous variables) require processing at the software level. Algorithmic problem reformulation and reduction is not only costly in terms of the computing resources it requires, but it can result in ill-behaved energy landscapes that are difficult to optimize. Such schemes hinder the practical applicability of the resultant technologies, as evidenced by over a decade’s worth of unsuccessful attempts to find practical applications for the technologies.

As demonstrated by several previous studies [19–23], coherent-network computing has appeared as a promising approach for solving NP-hard combinatorial optimization problems. The well-known time-multiplexed CIMs utilize a network of degenerate optical parametric oscillators (degenerate OPO or DOPO) that are injected into a ring cavity and gradually pumped at a rate well above the bifurcation threshold [15, 16]. These CIMs can be realized through either fully optical experiments involving optical delay lines (DL-CIM) [24, 25], or using the assistance of a digital processing device (e.g., an FPGA) in a measurement-feedback procedure (MF-CIM) [20, 26]. While the digital processor may limit the performance of the measurement-feedback variant, in principle, by operating at optical frequencies, a CIM operates at a much faster clock speed and far lower power consumption compared to digital computers. With optical pulses that are only a few picoseconds apart, a CIM can solve binary optimization problems involving thousands of variables in a fraction of a second [27].

Pumping DOPOs above their bifurcation threshold forces their amplitudes to attain certain extreme positive or negative values. In this paper, we show that if the DOPOs are pumped less aggressively, their mean-field amplitudes can converge to stable fractional values. When solving binary optimization problems, conventional CIMs begin an optimization process as quantum analog devices and finish it as classical digital devices [28]. In the case of a device that exhibits very low signal loss, a product of coherent cat states of the form $|\alpha\rangle + |\alpha\rangle$ is generated at slightly above the bifurcation threshold [29]. Iterative weak measurements then result in a collapse into a product of classical bits representing a low-energy state of the optimization problem. While this quantum-to-classical

* Corresponding author: pooya.ronagh@1qbit.com
transition is widely considered to be a key characteristic of the CIM [28], a clear computational mechanism for why low-energy states are favoured by this process is not apparent.

In our work, we focus on the Langevin dynamics in the CIM, as opposed to the quantum-to-classical transition, as the mechanism responsible for solving optimization problems, with the diffusion process generated by the iterative noise injection using the nonlinear crystal and the measurement process. Langevin dynamics is known to mix into low-energy states of the dynamical system’s potential [30, 31]. However, in the CIM, the drift term deviates from the gradient of the objective function. This discrepancy, along with the non-convex landscape of the optimization problem, prevent the CIM from always succeeding in finding the global optimum. In order to test the hypothesis that Langevin dynamics in the CIM is the mechanism responsible for solving optimization problems, we demonstrate its effectiveness in solving non-convex continuous optimization problems without the need to prepare cat states. Indeed, being intrinsically continuous, the analog amplitudes of the DOPO pulses can be used to represent continuous variables of a continuous optimization problem below the saturation threshold. In order to encourage the realization of experiments beyond Ising optimization, we refer to our setup as a coherent continuous-variable machine (CCVM).

At the pump rates at which a CCVM is operated, the entire time evolution is carried out in the coherent regime of the device. While this itself opens the door to investigating the signatures of a quantum parallel search, we do not focus on such aspects of the operation of the machine, and instead view the machine as an optical device intended to optically perform a Brownian motion process which is then used to optically integrate a diffusion process of the form

\[ dc = b(t, c)\, dt - \nabla f(c)\, dt + \sigma(t)\, dW. \tag{1} \]

Here, \( c \) represents the mean amplitude and phase of the DOPO pulses, \( f \) is a continuous function we are interested in minimizing, and \( b(t, c) \) is a component-wise force driving the dynamics of each DOPO, which are controlled by the DOPO’s pump rate. The variable \( W \) is a standard Wiener process realized by weak measurements of the optical modes either via photon loss in a highly dissipative cavity, particularly in a delay-line CCVM (DL-CCVM), or via out-coupling and homodyne detection in a measurement-feedback CCVM (MF-CCVM).

Finally, \( \sigma(t) \) represents the diffusion rate, and is controlled via the usage of cavities with a certain amount of finesse in the DL-CCVM, or by controlling the out-coupling rate of beamsplitters in the MF-CCVM.

Monte Carlo simulations of such stochastic differential equations (SDE) have provided a breath of techniques in digital computing for solving global optimization problems [31, 32]. Indeed, the convergence rate of diffusion processes of the form presented above are well-studied. For instance, if the drift is a conservative force satisfying Poincaré or log-Sobolov inequalities, then the process mixes exponentially fast into the Gibbs distribution of the potential, but for highly non-convex potentials the mixing time is exponentially long. We do not propose employing coherent computing as a way to avoid exponentially long mixing times. Instead, we are interested in proposing power-efficient computing architectures that naturally integrate SDEs of the form given above by exploiting the quantum mechanical properties of light.

Our long-term goal is to explore the capabilities of CCVMs in solving continuous and mixed-integer non-convex optimization problems natively, thereby avoiding the aforementioned overhead. As a first step, we benchmark the performance of our proposed technology by predicting the time-to-solution (TTS) scaling of CCVMs for solving a simple example of continuous non-convex optimization problems, that is, the box-constrained quadratic programming (BoxQP) problem [33]. This is the simplest non-convex optimization problem one can attempt to solve, yet it is NP-hard [34] and no reasonably efficient classical algorithms for solving it exist [3–5].

This paper is organized as follows. In Section II, we introduce the various SDEs used to solve the BoxQP problem in our benchmarking study, including the systems of SDEs that represent the dynamics of the DL-CCVM and the MF-CCVM. In Section III, we propose methods by which these algorithms for solving BoxQP problems can be experimentally realized. Our benchmarking results are presented in Section IV. In Section V, we analyze the time evolution of the optical pulses in the DL-CCVM scheme when solving BoxQP problems. We conclude with a brief discussion and prospects for future research in Section VI.

II. SDEs GOVERNING CCVM DYNAMICS

Let \( f: \mathbb{R}^N \to \mathbb{R} \) be a real-valued differentiable function of \( N \) variables. To explore the dynamics of the proposed CCVM as a heuristic for the global optimization of \( f \), we consider six systems of SDEs. The first system of SDEs is typical Langevin dynamics, the convergence properties of which have been well-studied in mathematics and computer science [35]:

\[ dc_i = -\partial_i f(c)\, dt + \sigma_i\, dW_i \quad \forall i \in \{1, \ldots, N\}. \tag{2} \]

Here, \( c_i \) are \( N \) continuous random variables, and \( dW_i \sim \sqrt{dt}\, N(0, 1) \), hence injecting Gaussian noise of with a mean of zero and variance of \( dt \) into the dynamics. We use \( \partial_i \) to denote the \( i \)-th partial derivative \( \partial_i = \partial / \partial c_i \). For example, if \( f(c) = \sum_{i,j=1}^N \xi_{ij} c_i c_j \) is a quadratic function, then the \( i \)-th drift term is

\[ \partial_i f(c) = \sum_{j=1}^N \xi_{ij} c_j. \]

1 Our simulators of these CCVMs are publicly available as an open source Python package: https://github.com/1QB-Information-Technologies/ccvm/
\[-\partial_t f(c) = -\sum_{j=1}^N \xi_{ij} c_j,\] which may be realized physically via two-mode interactions.

The second system of SDEs is a modification of Langevin dynamics of the form

\[
dc_i = (-1 + p - c_i^2) c_i \, dt - \partial_t f(c) \, dt + \sigma dW_i. \tag{3}
\]

Here, the equation of motion [Eq. (2)] is augmented with an additional drift term representing three physical processes: the constant \(-1\) represents a photon loss rate that is normalized to that of the optical cavity; the parameter \(p\) represents the strength of the external pump field (e.g., a laser), normalized to the photon decay rate at the signal frequency \(\gamma_s\), inducing an amplification of field strengths in the cavity during the process; and the term \(c_i^2\) results from the nonlinear self-interaction induced by a nonlinear crystal (e.g., a periodically poled lithium niobate crystal (PPLN)) [15]. We refer to this system of SDEs as pumped Langevin dynamics.

The third system of SDEs is akin to the continuous-time model for the optical DL-CIM [24, 25], a fully quantum model for describing the trajectories of the amplitudes of the DOPO pulses inside the cavity via the positive P-representation expressed in terms of pairs of the random variables \(c_i\) and \(s_i\), representing the in-phase and quadrature-phase components of the signal field [36]

\[
dc_i = \left[\left(-1 + p - c_i^2 - s_i^2\right) c_i - \partial_t f(c)\right] \, dt \\
+ \frac{\kappa(t)}{\kappa} \sqrt{c_i^2 + s_i^2 + \frac{1}{2}} \, dW_{i1},
\]

\[
ds_i = \left[\left(-1 - p - c_i^2 - s_i^2\right) s_i - \partial_t f(s)\right] \, dt \\
+ \frac{1}{\kappa(t)} \frac{\gamma_s}{\kappa} \sqrt{c_i^2 + s_i^2 + \frac{1}{2}} \, dW_{i2}, \tag{4}
\]

where the Wiener increments \(dW_{i1}\) and \(dW_{i2}\) are independently sampled from an identical distribution. Here, \(\kappa = (\gamma_s \gamma_p / 2 \kappa^2)^{1/2}\), with \(\gamma_s\) and \(\gamma_p\) being the signal and pump decay rates, respectively. The parameter \(\kappa(t)\) is a factor for controlling the variance of the noise during the time evolution. It can be generated by continually injecting a squeezed vacuum state into the open port of the beamsplitter that out-couples the optical pulses from the cavity into the delay-line network [27].

We note that the drift term in Eq. (4) may not be realizable for arbitrary differentiable functions \(f\). Even in the case of polynomials, this requires the experimental realization of higher-order optical interactions, which is a challenge being tackled in current research [37]. Therefore, we restrict our study to the case of quadratic functions, as two-mode interactions of programmable strength can be realized using optical delay lines and beamsplitters. With this in mind, we refer to the system of SDEs (4) as DL-CCVM dynamics.

The final system of SDEs studied follows the continuous-time model of the MF-CIM [20, 26]. For the conventional MF-CIM, under the assumption that the DOPO pulses maintain a Gaussian distribution throughout the time evolution, their mean values and variances, \(\mu_i\) and \(\sigma_i\), follow the SDEs [38]

\[
d\mu_i = \left[-(1 + j) + p - g^2 \mu_i^2\right] \mu_i \, dt \\
- \lambda \partial_t f(\mu) \, dt + \sqrt{\lambda} dW_i,
\]

\[
d\sigma_i = 2 \left[-(1 + j) + p - 3g^2 \mu_i^2\right] \sigma_i \, dt \\
- 2j (\sigma_i - 1/2)^2 \, dt + \left[(1 + j) + 2g^2 \mu_i^2\right] \, dt. \tag{5}
\]

Here, \(j\) is the normalized continuous measurement strength, \(\lambda\) is a hyperparameter, and \(g\) is the normalized second-order nonlinearity coefficient. We refer to this system of SDEs as MF-CCVM dynamics. Note that, unlike in the previous systems of SDEs, the gradient of the objective function, forming the key part of the optimization threshold and may employ more-general mechanisms to generate the coupling term. The machines associated with these implementations use squeezed coherent states of light in solving non-convex continuous optimization problems. By using the quadrature-field properties of the optical pulses to represent the variables of a given problem, and by controlling the noise injected into the system, a stochastic process resembling Langevin dynamics is implemented, which typically mixes into a low-energy state of the objective function after a sufficiently long evolution period. The key element of the architecture is a DOPO, which involves an optical resonator in the form of a ring cavity, along with a nonlinear optical crystal pumped by a laser, for processing information that is encoded in time-multiplexed oscillations of the resonator. Using the process of optical parametric amplification (OPA), which is based on spontaneous parametric down-conversion, squeezed states are generated and continually amplified. The associated time-multiplexed pulses are coherently and iteratively coupled with the gradient of the objective function, forming the key part of the drift term of the dynamics. The gradient can be implemented by either a delay-line or measurement-feedback scheme. The diffusion process is realized by the injection of quantum noise during the OPA process, as
where the respective physical machines. The wall-clock TTS is computed from the round-trip time of the optical pulses. These solvers are used to predict the wall-clock TTS of standard digital hardware and analog quantum-optical devices. In our simulations, the success probabilities of these solvers are used to predict the wall-clock TTS of the respective physical machines. The wall-clock TTS is computed from the round-trip time of the optical pulses in the cavity, the pulse durations, and the number of pulses; for more details, see Section IV.

III. SOLVING BoxQP PROBLEMS

The BoxQP problem can be formulated as follows:

\[
\begin{align*}
\text{maximize} & \quad f(x) = \frac{1}{2} \sum_{i,j=1}^{N} Q_{ij} x_i x_j + \sum_{i=1}^{N} V_i x_i, \\
\text{subject to} & \quad \ell_i \leq x_i \leq u_i \quad \forall i \in \{1, \ldots, N\},
\end{align*}
\]

where \( Q \in \mathbb{R}^{N \times N} \) is a symmetric matrix, \( V \in \mathbb{R}^{N} \) is a real \( N \)-dimensional vector, and the lower and upper bounds \( \ell_i \in \mathbb{R} \) and \( u_i \in \mathbb{R} \) specify the box constraints.

In both the DL-CIM and MF-CIM, a nonlinear crystal placed in an optical ring cavity fed by a sequence of pump laser pulses plays a key role. A PPLN crystal is commonly used to generate a degenerate nonlinear optical parametric process: some pump photons with a frequency of \( \omega_p \) are down-converted to two degenerate signal photons that have the frequency \( \omega_s = \omega_p/2 \). These signal pulses then propagate inside the optical ring cavity, and are coherently amplified in each round trip when passing through the PPLN crystal. We note that, in the DL-CCVM, the nonlinear crystal is the only active element in the system and amplifies the vacuum states toward coherent states representative of the solution.

In CIMs, the pump rate of DOPOs is increased from below a bifurcation threshold to a pump rate above the threshold. This creates a transition from squeezed vacuum states below the threshold to coherent states with a binary phase oscillation (0 or \( \pi \)). Meanwhile, the pulse amplitudes saturate to extreme minimum and maximum values that depend on the properties of the nonlinear crystal and the pump field strength [15]. Operating in this regime allows a CIM to solve binary optimization problems. In contrast, we are interested in performing the same experiment below the saturation threshold, where instead of a binary spin, we obtain an analog degree of freedom encoded in the linear superposition of quadrature-amplitude eigenstates forming the squeezed coherent state.

During each round trip, a portion of each optical pulse is coupled out of the optical ring cavity to calculate and re-inject the drift term \( -\nabla f \). In the MF-CCVM scheme, each out-coupled pulse is measured using a homodyne measurement. This measurement serves two key roles: it creates the Brownian motion \( \sqrt{\sigma} \xi(W_i) \) as required for the diffusion process and it provides the measured mean-field amplitudes \( \mu \), which are then converted to digital information and used to compute \( -\nabla f(\mu) \) via a digital processor. The result is then converted back to an analog signal. Intensity and phase modulators clocked with the external pump are used to achieve amplitudes equal to the partial derivatives \( \partial_i f(\mu) \), and to match the phase of the newly generated pulses with that of the original pulse. The result is injected back into the main optical ring cavity at the appropriate time for each optical pulse [27].

In contrast, in the DL-CCVM scheme, the drift terms \( -\partial_i f(c) = \sum_{j=1}^{N} \xi_{ij} c_j \) are linear functions of pulse amplitudes. As with the DL-CIM, this coupling term may be calculated through a network of delay lines. Each pulse is split between optical lines of different lengths using an array of beamsplitters. The amplitude and the phase of the pulses within each delay line are appropriately manipulated using amplitude and phase modulators, before the pulses join back together to create a coupling term. The round-trip time for both architectures can be estimated as the pulse repetition time multiplied by the number of pulses [25, 27].

The continuous variables of the BoxQP problem can be
natively encoded into analog pulse amplitudes. The box constraints can be implemented either by exploiting the saturation thresholds of the pulse amplitudes in the case of both the DL-CCVM and the MF-CCVM, described by Eq. (4) and Eq. (5), respectively, or, in the case of the MF-CCVM, also by clipping the amplitudes during each round trip using the digital processor. Similarly, for the Langevin and pumped Langevin dynamics, the box constraints can be implemented by simply clamping each amplitude $c_i$ between the respective lower and upper bounds $\ell_i$ and $u_i$ at each iteration of Monte Carlo simulation.

More specifically, in the DL-CCVM scheme, by performing the change of variables

$$x_i \mapsto \frac{1}{2} \left( \frac{c_i}{s} + 1 \right) (u_i - \ell_i) + \ell_i,$$  \hspace{1cm} (8)

we can enforce the box constraint for each DOPO pulse. Here, $s$ is the saturation amplitude for all DOPO pulses, assuming that they all saturate to the same value. Assuming the coupling strength between the pulses is small, and the quadrature amplitudes $s_i$ are negligible, the saturation amplitude is approximately equal to $s \approx \sqrt{1 - p_0}$, where $p_0$ is the final pump field strength in a linear pump schedule $p(t) = \frac{t}{T}p_0$. Therefore, the drift terms in the SDEs of the DL-CCVM described by Eq. (4) are

$$-\partial_t f(c) = -\sum_{j=1}^{N} Q_{ij} \frac{u_i - \ell_i}{2s} \left[ \frac{1}{2} \left( \frac{c_j}{s} + 1 \right) (u_j - \ell_j) + \ell_j \right] + \frac{V_i(u_i - \ell_i)}{2s},$$  \hspace{1cm} (9)

where we have substituted the variables in Eq. (7) by the expression on the right-hand side of Eq. (8) and taken the gradient of the objective function with respect to $c_i$. This expression is linear in the pulse amplitudes $c_i$ and thus able to be implemented using the DL-CCVM. The coefficients of the pulse amplitudes in the above equation are equal to $-Q_{ij} \frac{(u_i - \ell_i) + (u_j - \ell_j)}{4s}$. They can be implemented using the intensity and phase modulators on each delay line and for each pulse index $j$. The rest of the terms in Eq. (9) amount to the constant $\sum_j Q_{ij} \frac{(u_i - \ell_i)(u_j + \ell_j)}{4s} + \frac{V_i(u_i - \ell_i)}{2s}$, which can be calculated in advance and added to each pulse during each round trip. An offset value for each pulse can be introduced by an external source (e.g., a laser) that operates at the frequency of the DOPO pulses.

In the MF-CCVM scheme, the BoxQP problem’s variables may simply be encoded as $x_i := \tilde{\mu}_i$. However, if the lower bounds $\ell_i$ are non-negative, which is the case for the problems studied in Section IV, we again use the encoding

$$x_i \mapsto \frac{1}{2} \left( \frac{\tilde{\mu}_i}{s} + 1 \right) (u_i - \ell_i) + \ell_i,$$  \hspace{1cm} (10)

to increase the success probability of the optimizer. This encoding results in the drift term

$$-\partial_t f(\tilde{\mu}) = -\sum_{k=1}^{N} Q_{ik} \frac{u_i - \ell_i}{2s} \left[ \frac{1}{2} \left( \frac{\tilde{\mu}_k}{s} + 1 \right) (u_k - \ell_k) + \ell_k \right] + \frac{V_i(u_i - \ell_i)}{2s}.$$  \hspace{1cm} (11)

Here, $s$ is a hyperparameter that represents a saturation bound. As long as $s < \sqrt{p - (1+j)/g}$, it is tuned to obtain the best performance.

**IV. BENCHMARKING RESULTS**

We have tested the solvers introduced in the previous sections on randomly generated problem instances. We generated our instances using the same approach as in Ref. [39], where the elements of the $N \times N$ symmetric matrix $Q$ and the vector $V$ are random integers generated uniformly between $-50$ and $+50$. The generated instances are labelled $N$-D-$S$, where $N$ is the problem size, $D$ is the density of the nonzero entries of $Q$ and $V$, and $S$ is the seed number of the random number generator. For simplicity, all variables are assumed to be constrained to the domain $[0,1]$, that is, $\ell_i = 0$ and $u_i = 1$ for all $i \in \{1, \ldots, N\}$. The global maxima of the generated instances were found using Gurobi 9.5 [40].

Figure 2 shows the physical TTS for the two CCVM architectures introduced in Section II. The TTS is defined as the number of trials that must be performed to ensure a high probability (specified by some target probability of success) of observing an optimal or an acceptable approximate solution at least once, multiplied by the time required for the execution of a single trial. Here, we assume a target probability of success of $0.99$. Thus, we compute the TTS according to the equation $TTS = R_{90} \cdot T_{\text{max}}$, where $R_{90} = \log(0.01) / \log(1 - P_T^i)$ is the required number of trials to solve a given problem instance (i.e., observe an acceptable solution at least once) with a $99\%$ success rate, and

![FIG. 2: Estimated physical TTS for the two CCVM architectures studied: (a) DL-CCVM and (b) MF-CCVM. For each CCVM architecture, several TTS curves are displayed, corresponding to a solution within a target percentage gap from the optimal solution. The shaded regions show the IQR corresponding to the percentile of the solved instances that were found to be within the specified gap.](image)
$T_{\text{max}}$ is the estimated physical time required for a single trial at solving a given BoxQP instance on the physical device. In the definition of $R_{\text{gap}}$, the quantity $P_e$ denotes the success probability of a single trial; it is found by solving a given problem instance 1000 times and evaluating the number of times of the global maximum was found for that given instance. The time for a single trial is estimated as $T_{\text{max}} = n_{\text{iter}} \cdot N \cdot T_{\text{pulse}}$, where $T_{\text{pulse}}$ is the time delay between two individual pulses within the cavity; for our benchmarking study, we have used the value $T_{\text{pulse}} = 10$ picoseconds [23]. Figure 2 displays the TTS for finding an approximate solution that is correct up to a multiplicative error. This is because the solutions to the problem instances are intrinsically fractional values, and due to the noise inherent to the method. Figure 2 shows the TTS for different multiplicative factors of the optimal value. For instance, a 0.1% gap indicates that the approximate objective value found is within 0.1% of the global optimal value of the problem. The shaded regions show the interquartile range (IQR), that is, between the 25th and 75th percentiles of the solved problem instances, while the solid curves represent the median.

Figure 3 shows the wall-clock TTS (i.e., the TTS for solving the instances on a GPU) for all four solvers introduced in Section II. Table I shows the parameters of the solvers used for generating Figs. 2 and 3. For the Langevin, pumped Langevin, and DL-CCVM solvers, the pump field $p(t) = \frac{1}{T}p_0$ follows a linear schedule, while for the MF-CCVM it is set to $p(t) = \frac{1}{T}p_0 + 1 + j(t)$ to compensate for the measurement loss and background loss. On the other hand, the measurement strength $j(t) = j_0 \exp(-\alpha \frac{t}{T})$, where $\alpha$ is an arbitrary parameter. In this scheme, the measurement strength is programmed according to the required amount of noise in the stochastic process. For the DL-CCVM solver, the injected noise parameter $r(t) = r_0 \exp(-\beta \frac{t}{T})$, where $\beta$ is an arbitrary parameter.

Figure 4 shows our benchmarking comparison of the estimated physical TTS for the optics-based implementations of DL-CCVM and MF-CCVM, the TTS for all four solvers introduced in Section II implemented on conventional digital devices (i.e., those based on electronics, not optics), and the TTS for the quadratically constrained quadratic programming (QCQP) solver [41], a general heuristic solver for quadratic programming problems. Optics-based CCVMs can be several orders of magnitude more performant than their conventional digital counterparts, and, especially in the case of the DL-CCVM, be much more power efficient.

Figure 5 demonstrates a lack of dependency of the performance of the solvers on the density of $Q$ and $V$. Specifically, it shows the wall-clock TTS of the six solvers defined in Section II for finding a solution within a 0.1% gap from the optimal solution, individually calculated for various density values of the problem instances. While for exact solvers, such as CPLEX and Gurobi, the TTS heavily depends on the density [39, 40], for the solvers analyzed here, the TTS does not show any significant dependency on it. This is because, unlike the exact solvers, the heuristic solvers analyzed here do not exploit the sparsity of problem instances.

V. TIME-EVOLUTION OF THE DL-CCVM

Figure 6 shows the time evolution of the DOPO pulses’ amplitudes for the DL-CCVM dynamics (see Eq. (4)). A randomly generated problem instance of size $N = 20$ was used for the results displayed in each panel. The heavy dashed lines indicate the values $-s$ and $s$ at which the amplitudes are clamped at the end of the process. The

| Parameter | Langevin | Pumped Langevin | DL-CCVM | MF-CCVM |
|-----------|----------|-----------------|---------|---------|
| $n_{\text{iter}}$ | 15000 | 15000 | 15000 | 15000 |
| $p_0$ | 1.5–2.0 | 1.75–2.5 | 0.1–1.0 | |
| $dt$ | 0.005–0.01 | 0.005–0.01 | 0.005–0.05 | 0.0025 |
| $\sigma$ | 0.02–0.5 | 0.02–0.5 | – | – |
| $(r_0, \beta)$ | – | – | (10,3) | – |
| $A_j$ | – | – | 10 | – |
| $s$ | – | – | $\sqrt{1-p_0}$ | 0.2 |
| $(j_0, \alpha)$ | – | – | – | (20,3) |
| $\lambda$ | – | – | – | 10–20 |
| $g$ | – | – | – | 0.01 |

TABLE I: Parameters for the Langevin, pumped Langevin, DL-CCVM, and MF-CCVM solvers.
system parameters, including the pump value, are optimized for the highest success probability through parameter tuning. Their ranges are shown in Table I. Since the clamping is applied at the end of the evolution, it can be considered a post-processing step that is performed on the corresponding physical device. The estimated TTS values for optics-based CCVM solvers are more than three orders of magnitude lower than the CPU wall-clock TTS values for the conventional digital solvers.

By exploiting the fact that degenerate optical parametric oscillator (DOPO) pulses’ amplitudes are intrinsically continuous, the continuous variables of a BoxQP problem can be directly encoded and processed as analog information. This allows us to avoid the associated large resource overheads of digital (i.e., binary) information processing and discretization errors. Furthermore, the box constraints can be implemented by exploiting the amplitude saturation of the DOPO pulses, or simply by clamping the pulse amplitudes when using a measurement-feedback scheme. Our numerical results motivate using coherent continuous-variable machines (CCVM) as analog heuristic solvers for non-convex continuous optimization.

Our proposed CCVMs are intrinsically analog devices, that is, information is encoded in analog signals and processed by manipulating these signals continuously in time. Note that this is different from the notion of analog computing often associated with some qubit-based quantum devices that do not rely on the circuit model of quantum computation (e.g., quantum annealers), which use a binary encoding in qubits, but process information in an analog manner, that is, by evolving the corresponding quantum states continuously in time.

Our benchmarking study demonstrates that CCVMs can solve non-convex continuous optimization problems over three orders of magnitude faster than classical heuristic solvers. We do not expect CCVMs to offer a scaling advantage in the time to solution. Instead, our

VI. CONCLUSION

We have proposed an approach for solving problems in the class of NP-hard box-constrained quadratic programming (BoxQP) problems using modifications to the dynamics of conventional coherent Ising machines (CIM).
FIG. 6: Time evolution of the DOPO pulses’ amplitudes for four randomly generated BoxQP problem instances solved using the DL-CCVM: (a) tuning020-50-2, (b) tuning020-90-8, (c) tuning020-100-3, and (d) tuning020-100-5, where the format tuning$N$-$D$-$S$ represents a problem instance of size $N$, density $D$, and seed number $S$. The heavy black dashed lines indicate the values $-s$ and $s$, at which the amplitudes are clamped prior to computing the objective function. The objective function is calculated after the substitution $x_i = \frac{1}{2}(c_i/s + 1)$ to satisfy the box constraint in Eq. (7), with $l_i = 0$ and $u_i = 1$ for all $i \in \{1, \ldots, N\}$. Note that, for each of the problem instances, the majority of the amplitudes have an optimal value at the boundary of the box constraint, while only some of the amplitudes converge to fractional values.

our goal is to dramatically reduce the energy consumption of simulating Langevin dynamics compared to conventional digital computers.

Our study serves to stimulate research into experimental realizations of CCVMs. Implementing more-complicated constraints than box constraints is an avenue for further exploration. In addition, more-general objective functions may be implemented using CCVMs. Measurement-feedback mechanisms can be used to generate the gradient of an arbitrary differentiable objective function acting as a drift term. However, generating the gradient of such an arbitrary function is more challenging in the case of CCVM schemes based on optical delay lines, further motivating research into the experimental realization of higher-order optical interactions.

ACKNOWLEDGEMENT

The authors thank Yoshihisa Yamamoto, Edwin Ng, Satoshi Kako, and Sam Reifenstein for helpful discussions. We thank our editor, Marko Bucyk, for his careful review and editing of the manuscript. The authors acknowledge the financial support received through the NSF’s CIM Expeditions award (CCF-1918549). P. R. ack-
[1] S. Burer and A. N. Letchford, Non-convex mixed-integer nonlinear programming: A survey, Surveys in Operations Research and Management Science **17**, 97 (2012).

[2] P. Pardalos and G. Schnitger, Checking local optimality in constrained quadratic programming is NP-hard, Operations Research Letters **7**, 33 (1988).

[3] S. Burer and D. Vandenberghe, Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound, Computational Optimization and Applications **43**, 181 (2009).

[4] P. Bonami, O. Günlük, and J. Linderoth, Globally solving nonconvex quadratic programming problems with box constraints via integer programming methods, Mathematical Programming Computation **10**, 333 (2018).

[5] S. de Vries and B. Perscheid, Tight compact extended relaxations for nonconvex quadratic programming problems with box constraints, Journal of Global Optimization (2022).

[6] N. Mohseni, P. L. McMahon, and T. Byrnes, Ising machines as hardware solvers of combinatorial optimization problems, Nature Reviews Physics **4**, 363 (2022).

[7] S. Kirkpatrick, Optimization by simulated annealing, Science **220**, 671 (1983).

[8] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon, Optimization by simulated annealing: An experimental evaluation; part I, graph partitioning, Operations Research **37**, 865 (1989).

[9] M. Aramon, G. Rosenberg, E. Valiante, T. Miyazawa, H. Tamura, and H. G. Katzgraber, Physics-inspired optimization for quadratic unconstrained problems using a digital annealer, Frontiers in Physics **7**, 48 (2019).

[10] T. Takemoto, M. Hayashi, C. Yoshimura, and M. Yamakoa, 2.6 A 2×30k-spin Multichip Scalable Annealing Processor based on a Processing-In-Memory Approach for Solving Large-Scale Combinatorial Optimization Problems, in *2019 IEEE International Solid-State Circuits Conference-(ISSCC)* (IEEE, 2019) pp. 52–54.

[11] D. de Falco, B. Apolloni, and N. Cesa-Bianchi, A numerical implementation of quantum annealing (1988).

[12] B. Apolloni, C. Carvalho, and D. de Falco, Quantum stochastic optimization, Stochastic Processes and their Applications **33**, 233 (1989).

[13] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, Phys. Rev. E **58**, 5355 (1998).

[14] J. Brooke, D. Bitko, G. Aeppli, *et al.*, Quantum annealing of a disordered magnet, Science **284**, 779 (1999).

[15] Z. Wang, A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto, Coherent Ising machine based on degenerate optical parametric oscillators, Phys. Rev. A **88**, 063853 (2013).

[16] Y. Yamamoto, K. Aihara, T. Leleu, K.-i. Kawarabayashi, S. Kako, M. Fejer, K. Inoue, and H. Takesue, Coherent Ising machines—optical neural networks operating at the quantum limit, npj Quantum Information **3**, 49 (2017).

[17] U. Benlic, E. K. Burke, and J. R. Woodward, Breakout local search for the multi-objective gate allocation problem, Computers & Operations Research **78**, 80 (2017).

[18] V. K. Prajapati, M. Jain, and L. Chouhan, Tabu search algorithm (TSA): A comprehensive survey, in *2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE)* (IEEE, 2020) pp. 1–8.

[19] R. Hamerly, T. Inagaki, P. L. McMahon, D. Venturelli, A. Marandi, T. Onodera, E. Ng, C. Langrock, K. Inaba, T. Honjo, *et al.*, Experimental investigation of performance differences between coherent Ising machines and a quantum annealer, Science Advances **5**, eaau0823 (2019).

[20] T. Leleu, Y. Yamamoto, P. L. McMahon, and K. Aihara, Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity, Physical Review Letters **122**, 040607 (2019).

[21] T. Leleu, F. Khoyratee, T. Levi, R. Hamerly, T. Kohno, and K. Aihara, Scaling advantage of chaotic amplitude control for high-performance combinatorial optimization, Communications Physics **4**, 266 (2021).

[22] S. Reifenstein, S. Kako, F. Khoyratee, T. Leleu, and Y. Yamamoto, Coherent Ising Machines with optical Error Correction Circuits, Adv. Quantum Technol. **4**, 2100077 (2021).

[23] K. Sankar, A. Scherer, S. Kako, S. Reifenstein, N. Ghaedermarzy, W. B. Krayenhoff, Y. Inui, E. Ng, T. Onodera, P. Ronagh, *et al.*, Benchmark study of quantum algorithms for combinatorial optimization: unitary versus dissipative, arXiv:2105.03528 (2021).

[24] K. Takata, A. Marandi, and Y. Yamamoto, Quantum correlation in degenerate optical parametric oscillators with mutual injections, Phys. Rev. A **92**, 043821 (2015).

[25] D. Maruo, S. Utsunomiya, and Y. Yamamoto, Truncated Wigner theory of coherent Ising machines based on degenerate optical parametric oscillator network, Physica Scripta **91**, 083010 (2016).

[26] S. Kako, T. Leleu, Y. Inui, F. Khoyratee, S. Reifenstein, and Y. Yamamoto, Coherent ising machines with error correction feedback, Advanced Quantum Technologies **3**, 2000045 (2020).

[27] P. L. McMahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, R. L. Byer, M. M. Fejer, H. Mabuchi, and Y. Yamamoto, A fully programmable 100-spin coherent Ising machine with all-to-all connections, Science **354**, 614 (2016), https://www.science.org/doi/pdf/10.1126/science.aah5178.

[28] Y. Yamamoto, T. Leleu, S. Ganguli, and H. Mabuchi, Coherent Ising machines—quantum optics and neural network perspectives, Applied Physics Letters **117**, 160501 (2020), https://doi.org/10.1063/5.0016140.

[29] S. Kiesewetter and P. D. Drummond, Coherent ising machine with quantum feedback: The total and conditional master equation methods, Phys. Rev. A **106**, 022409 (2022).

[30] T.-S. Chiang, C.-R. Hwang, and S. J. Sheu, Diffusion for global optimization in $\mathbb{R}^n$, SIAM Journal on Control and Optimization **25**, 737 (1987), https://doi.org/10.1137/0325042.

[31] P. Xu, J. Chen, D. Zou, and Q. Gu, Global convergence of Langevin dynamics based algorithms for nonconvex optimization, arXiv:1707.06618 (2017).

[32] M. Raginsky, A. Rakhlin, and M. Telgarsky, Non-convex learning via stochastic gradient Langevin dynamics: A nonasymptotic analysis, in *Conference on Learning Theory (FMLR, 2017)* pp. 1674–1703.
[33] P. L. D. Angelis, P. M. Pardalos, and G. Toraldo, Quadratic programming with box constraints, in *Developments in global optimization* (Springer, 1997) pp. 73–93.

[34] K. G. Murty and S. N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Tech. Rep. (1985).

[35] G. O. Roberts and R. L. Tweedie, Exponential convergence of Langevin distributions and their discrete approximations, Bernoulli (1996).

[36] A. Marandi, Z. Wang, K. Takata, R. L. Byer, and Y. Yamamoto, Network of time-multiplexed optical parametric oscillators as a coherent Ising machine, Nature Photonics 8, 937 (2014).

[37] R. Yanagimoto, T. Onodera, E. Ng, L. G. Wright, P. L. McMahon, and H. Mabuchi, Engineering a Kerr-based deterministic cubic phase gate via Gaussian operations, Physical Review Letters 124, 240503 (2020).

[38] T. Shoji, K. Aihara, and Y. Yamamoto, Quantum model for coherent Ising machines: Stochastic differential equations with replicator dynamics, Phys. Rev. A 96, 053833 (2017).

[39] D. Vandenbussche and G. L. Nemhauser, A branch-and-cut algorithm for nonconvex quadratic programs with box constraints, Mathematical Programming 102, 559 (2005).

[40] Gurobi Optimization, LLC, Gurobi Optimizer Reference Manual (2022).

[41] J. Park and S. Boyd, General heuristics for nonconvex quadratically constrained quadratic programming, arXiv:1703.07870 (2017).