Gain-dissipative simulators for large-scale hard classical optimisation

Kirill P. Kalinin[1] and Natalia G. Berloff[2,1*]

[1]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom and*

[2]*Skolkovo Institute of Science and Technology Novaya St., 100, Skolkovo 143025, Russian Federation*

(Dated: April 22, 2018)

Abstract

Recently several gain-dissipative platforms based on the networks of optical parametric oscillators, lasers and various non-equilibrium Bose-Einstein condensates have been proposed as analogue Hamiltonian simulators for solving large-scale hard optimisation problems. The parameters of such problems depend on the node occupancies that are not *a priory* known, which limits the applicability of the gain-dissipative simulators to the classes of problems easily solvable by classical computations. We show how to overcome this difficulty and formulate the principles of operation of such simulators for solving the NP-hard large-scale optimisation problems such as constant modulus continuous quadratic optimisation and quadratic binary optimisation for any general matrix. To solve such problems any gain-dissipative simulator has to implement a feedback mechanism for the dynamical adjustment of the gain and coupling strengths, so that occupancy of each node is the same. The estimates of the time operation of the physical implementation of the gain-dissipative simulators for large matrices show the speed-up of the several orders of magnitude in comparison with classical computations.

*correspondence address: N.G.Berloff@damtp.cam.ac.uk

Constructing models to simulate the physical world – analog simulators – has been instrumental in advancing scientific knowledge starting with orreries for the description of the motions of celestial objects [1] and differential analysers or mechanical integrators capable of modelling heat flow and transmission lines [2, 3]. With the invention of the classical universal computer based on von Neumann architecture the need to use analog simulators seemed to disappear. Their revival took place a couple of decades ago with Richard Feynmann proposing to use one well-tuneable quantum system to simulate another less accessible quantum system [4]. A large number of various quantum simulators has been proposed and elucidated with a vast number of quantum hardware developed and types of problems they aim to solve analysed. Among these are superconducting circuits, trapped ions, cold atoms in optical lattices, photons, liquid and solid-state NMR, quantum dots, and NV centres [5]. In addition to proposals to simulate quantum systems, there are proposals for quantum devices to simulate classical NP-hard problems, where the number of operations grows exponentially fast with the size of the problem [6, 7].

Solving efficiently an NP-complete problem means solving all NP problems, therefore, it is important to find one subclass of problems within the NP-complete problems class that starting with some size can be solved faster by an analogue simulator than a classical computer. For instance, D-Wave built from superconducting qubits is a so-called quantum annealer, which is a special purpose analogue device that solves quadratic unconstrained binary optimisation problems (QUBO) [8], which can be reformulated as the problem of finding the ground state of the famous model in statistical mechanics – the Ising Hamiltonian. Other platforms were proposed and demonstrated a proof-of-principle for finding the global minimum of the spin Hamiltonians such as the Ising and XY models: the networks of optical parametric oscillators, [9–13], coupled lasers [14], polariton condensates [15], and photon condensates [16]. The main principle of such simulators is based on a gain process that raises the system above the threshold for a phase transition to a coherent state. Accordingly, we refer to these platforms as gain-dissipative (GD) analog Hamiltonian simulators [17] that, in spite of having different quantum hardwares, share the basic principle that suggests the convergence to the global minimum of the Ising or XY Hamiltonian.

Such Hamiltonians are written for N classical "spins" $\mathbf{s}_j$, where for the Ising Model the "spins" take discrete values $s_j = \pm 1$ and for the XY Model are continuous and lie on a unit circle $\mathbf{s}_j = (\cos\theta_j, \sin\theta_j)$. The spins are coupled to each other with "strengths" J_{ij} that can

be positive (ferromagnetic) or negative (antiferromagnetic). At the ground state "spins" arrange their orientation as to minimise the Ising Hamiltonian $H_{\mathrm{I}} = -\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij}s_i s_j$ or the XY Hamiltonian

$$\min_{\boldsymbol{\theta}} \quad H_{XY}(\boldsymbol{\theta}) \equiv -\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij}\mathbf{s}_i \cdot \mathbf{s}_j, \tag{1}$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_i, ..., \theta_N)$.

Here we focus on the XY model, but the proposed approach can be generalised to the Ising or other spin models in a straightforward fashion as we also show below. The minimisation of the XY Hamiltonian can be alternatively written as the constant modulus quadratic continuous optimisation (QCO) problem

$$\min -\mathbf{z}^H \boldsymbol{J}\mathbf{z}, \quad \text{subject} \quad \text{to} \quad |z_j| = 1, \tag{2}$$

if we associate the "spins" $\mathbf{s}_j$ with the complex numbers $z_j = \cos\theta_j + \mathrm{i}\sin\theta_j$ and the coupling strengths J_{ij} between the spins with the matrix elements of $\boldsymbol{J}$. In general, these problems are known to be NP-hard [18] and finding even an approximate solution is hard as we discuss in the Supplementary Information.

In the gain-dissipative simulators the "spin" – the node (the "bit") of the simulators – is represented by the phase of a condensate at a particular spatial position [15, 16] or by the phase of a coherent state generated in a laser cavity [13, 14], so by the phase of the so-called *coherent center* (CC). All of the proposals for the gain-dissipative simulators that aim to find the absolute minimum of the spin Hamiltonians suffer from a serious limitation. As we show below, the coupling strengths J_{ij} in such systems are modified by the occupations (number densities) of CCs i and j. The densities, however, are not *a priory* known and, for a general matrix $\boldsymbol{J}$ of the coupling strengths J_{ij}, will be different from one CC to another. The variability in the densities may also lead to the system fragmentation when different subsystems acquire their own mode of oscillations, and the coherent steady state across all CCs will never be reached. In all the previous experimental realisations of the gain-dissipative simulators, an explicit or implicit assumption was made that the coupling terms are small so that each laser or condensate is stabilised independently at the same steady-state amplitude [9, 10, 13–15]. Such assumption is justified only for the simplest structures of the coupling matrix $\boldsymbol{J}$, where all CCs are almost equally connected with about the same

coupling strengths. The found solution for a more general matrix is bound to be either only approximate or invalid. The problem of unequal densities has been recognised before [19], but the proposed method to reduce the heterogeneity in densities was to drive the system using randomly modulated signals and can yield only very modest improvement and only if the densities are quite close for the unmodulated signals.

In this paper we show that by establishing a feedback connection between the gain mechanism and the density of the CC we can drive the system to the coherent ground state of the XY model, while the minimisers of this ground state will give the true minimum of the XY model precisely for the externally provided coupling strengths. We develop a general framework for the operation of the gain-dissipative analogue simulators based on the Fokker-Plank-Langevin gain-dissipative equations written for a set of CCs for two distinct systems: a system dominated by the dissipative coupling and the system dominated by the gain coupling, and describe the convergence to the global minimum of the XY model in both cases. This framework allows us to formulate the hardware requirements for a physical realisation of such simulator to achieve such a minimum and argue that such requirements are within the recent technological capabilities. We show that for large problem sizes the analogue simulator would outperform the classical computations by several orders of magnitude. Finally, motivated by the operation of such a simulator we develop a new class of global optimisers.

Gain-dissipative simulators dominated by dissipative coupling

The operation of a gain-dissipative optimiser consists of two stages: bosonic stimulation below the threshold and the coherence of operations at and above the threshold. As one ramps up the power of the gain mechanism (e.g. laser intensity) the gain overcomes the linear losses and is stabilised by the nonlinear gain saturation. The emergent coherent state maximises the total number of particles (minimises losses) and, therefore, minimises the Ising or XY Hamiltonian depending on whether the phases of the wavefunctions describing the CCs are constrained to 0 and π or not, respectively. To derive the governing equations for the system one can describe each CC at a position $\mathbf{r} = \mathbf{r}_i$ by a classical complex function $\Psi_i(t)$. Such description can be obtained by eliminating the spatial degrees of freedom or by using the Glauber-Sudarshan formalism of coherent states [20].

In a system dominated by the dissipative coupling [21] the state of the oscillator (CC) indexed by i is described by a wavefunction $\Psi_i(t)$ that obeys the rate equation

$$\frac{d\Psi_i}{dt} = (\gamma_i - \sigma|\Psi_i|^2)\Psi_i - iU|\Psi_i|^2\Psi_i + \sum_j J_{ij}\Psi_j, \tag{3}$$

where $\gamma_i = \gamma_i^{\rm inj} - \gamma_c$ is the effective gain at the CC indexed by i, where $\gamma_i^{\rm inj}$ is the rate of adding particles to the $i-$th CC and γ_c is the rate of losing the particles via linear dissipation, U is the strength (pseudo-potential) of the particle-particle interactions (that is negligible for laser systems, but not for the interacting condensates), σ is the strength of the nonlinear losses. The sum on the right-hand side corresponds to the dissipative coupling via the wave interference between the $i-$th and $j-$th CCs of the system [21]. Such systems of the coupled oscillators appear in many areas of dynamical research where it is used to describe the complex amplitude of the oscillators near the supercritical Hopf bifurcation [22] and various phenomena such as collective synchronisation transition, spatiotemporal chaos, the Benjamin-Feir instability, nonlinear wave propagation have been shown to occur in them [23]. To understand the nature of the function being minimised by Eq. (3) we use the Madelung transformation $\Psi_i = \sqrt{\rho_i}\exp i\theta_i$ to rewrite it in terms of the number densities and the phases of the CCs:

$$\frac{1}{2}\dot{\rho}_i(t) = (\gamma_i - \sigma\rho_i)\rho_i + \sum J_{ij}\sqrt{\rho_i\rho_j}\cos(\theta_i - \theta_j) \tag{4}$$

$$\dot{\theta}_i(t) = -U\rho_i - \sum J_{ij}\frac{\sqrt{\rho_j}}{\sqrt{\rho_i}}\sin(\theta_i - \theta_j). \tag{5}$$

The first term on the right-hand side of Eq. (5) tends to provide θ_i with its own frequency of oscillations whereas the second term tends to synchronise the phases. Phase synchronisation in such a system has been extensively studied especially in the context of semiconductor laser arrays [24, 25]. To guarantee that the steady state of the system is reached and coincides with the steady state of the QCO problem, however, one needs to make sure that the gain mechanism is chosen so that all densities ρ_i are the same. Only under this condition Eq. (5) describes the gradient decent to the minimum of the XY Model. Previously, any realisation of the XY model using the laser systems (or non-equilibrium condensates) was based on the assumption that all lasers (condensates) have the same steady-state photon

(particle) number [13, 24, 25]. This limits the QCO problems that can be addressed by such a framework to trivial ones where all CCs have an almost equal number of connections with almost the same pumping rate. To implement any couplings and connectivities, one needs to be able to control the pumping rate of the individual CCs and bring all the densities to the same value at the threshold. We schematically illustrate the operational principle of such control mechanism in Fig. 1. Starting from below the threshold, all CCs are equally pumped as shown at some $t = t_1$. Depending on the node connectivity, the non-zero densities emerge at different rates for each CC as the pumping intensity increases and takes some CCs above the specified threshold $\rho = \rho_{\rm th}$ as illustrated on Fig. 1 at some later time $t = t_2$. The pumping mechanism must be adjusted for each CC to enable the saturation at the same density: decreased for CCs with densities above the threshold and increased for the CCs below the threshold. This requires the introduction of a feedback mechanism for the pumping $\gamma_i^{\rm inj}$ that controls each CC separately

$$\frac{d\gamma_i^{\rm inj}}{dt} = \epsilon(\rho_{\rm th} - \rho_i), \tag{6}$$

where ϵ is a parameter that can be tuned to control the speed of approaching the threshold. The steady state of Eqs. (4-6) is

$$\rho_i = \rho_{\rm th} = -\frac{1}{U}\frac{H_{XY}(\boldsymbol{\theta})}{d\theta_i} = \frac{1}{\sigma}(\gamma_i + \sum_j J_{ij}\cos(\theta_i - \theta_j)) \tag{7}$$

for all i and describes the gradient decent to the minimum of the XY Hamiltonian. The minimum is approached from below by gradually raising the pumping above the threshold which facilitates the achievement of the global minimum as Fig. 1 at $t = t_3$ illustrates.

The actual physical system is driven by classical and quantum fluctuations that force phases to span various configuration before the coherence is established. In our classical model we can introduce white complex noise $\xi_i(t)$ ($\langle \xi_i(t)\xi_i^*(t')\rangle = \delta(t - t')$) which represents intrinsic vacuum fluctuations and classical noise with a diffusion coefficient D which disappears at the threshold:

$$\frac{d\Psi_i}{dt} = (\gamma_i - \sigma|\Psi_i|^2)\Psi_i - iU|\Psi_i|^2\Psi_i + \sum_j J_{ij}\Psi_j + D\xi_i(t). \tag{8}$$

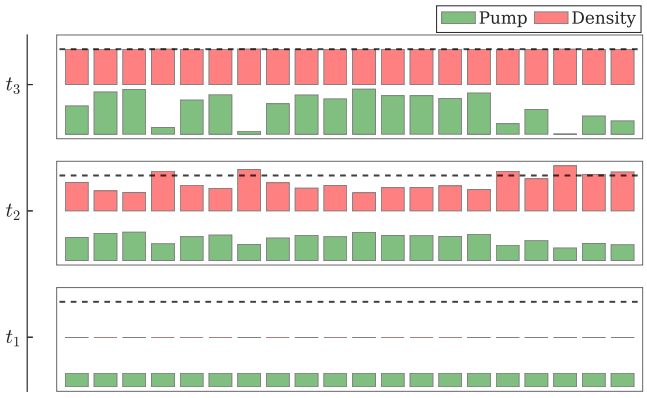

FIG. 1: The schematics of operation of the GD simulator. Initially, at $t = t_1$ all CCs are equally pumped (green bars) below the threshold, and all have negligible densities. As the pumping intensities increase and depending on the connectivity between CCs, the different densities emerge (red bars) as shown at some $t = t_2$. The individual control of the pumping intensity as described by Eq. (6) leads to the steady state with all the densities reaching the threshold at $t = t_3$.

Equations (6) and (8) can be viewed as the classical mean-field descriptions that can be efficiently simulated on classical computers, and indeed as we show below lead to a new class of global optimisation algorithms. We argue that the actual physical gain-dissipative simulator may outperform any classical computer algorithm due to the super-fast rates of its operation and inherent parallelism of spanning various phases before the configuration that enables coherence at the threshold is reached. To stand such a challenge any simulator has to meet several requirements for its operation such as the realisation of independent control of the couplings between two nodes, coupling beyond the next neighbour, and the feedback mechanism that allows one to adjust the pumping depending on whether or not the corresponding CC is above or below the threshold.

Gain-dissipative simulators dominated by gain coupling

The method for achieving the global minimum of the QCO problems presented in the last section has to be modified when the coupling strengths depend on the gain mechanism. The gain-dissipative simulator in this case is described by

$$\frac{d\Psi_i}{dt} = (\gamma_i - \sigma|\Psi_i|^2)\Psi_i - iU|\Psi_i|^2\Psi_i + \gamma_i^{\text{inj}}(t)\sum_j K_{ij}\Psi_j + D\xi_i(t). \tag{9}$$

The couplings K_{ij} in this system can have a different origin: they can be geometrically induced by the particle outflow from other CCs [15] or induced by the mutual injection rate between lasers [13, 26] or spatially separated condensates; they can be controlled by external potentials [16]. In Methods we derive Eq. (9) for the geometrically induced flow for the network of the polariton condensates starting from the mean–field description by the Ginzburg-Landau equations [27].

Unlike as in Eq.(8), the coupling terms $\gamma_i^{\text{inj}}(t)K_{ij}$ now depend on the particle injection rates γ_i^{inj} that would not be known *a priori* if one requested the equal densities at the threshold. Therefore, not only $\gamma_i^{\text{inj}}(t)$ has to be adjusted in time to equalise the densities using Eq. (6) but also the coupling parameters K_{ij} have to be modified in time to bring the required couplings J_{ij} of the QCO matrix to the steady state by

$$\frac{dK_{ij}}{dt} = \hat{\epsilon}(J_{ij} - \gamma_i^{\text{inj}}(t)K_{ij}), \tag{10}$$

where $\hat{\epsilon}$ controls the rate of the coupling strengths adjustment. Since $\hat{\epsilon} \gg \epsilon$ such adjustments do not significantly slow down the operation of the simulator as they have to be performed much more rarely then adjustments of the gain. Equation (10) indicates that the couplings need to be reconfigured depending on the injection rate: if the coupling strength scaled by the gain at time t is lower (higher) then the objective coupling J_{ij}, it has to be increased (decreased). Equations (6,9,10) replace Eqs. (6,8) for solving QCO problems using the gain-dissipative simulator with gain couplings and we refer to the corresponding algorithm as the GD-mod algorithm.

We note that the gain-dependent rate of nonlinear dissipation σ can be incorporated into Eq. (10) in a similar fashion. The generalisation of the GD algorithm for solving different classes of NP-hard problems with the analytically available Jacobians is also straightforward for both continuous and discrete problems. For instance, the Coherent Ising Machine based on the network of degenerate optical parametric oscillators [9–12] or other physical system implementing QUBO problems (the global minimisation of the Ising Hamiltonian), are modelled by incorporating the continuous spin mapping $f(x_i) = \delta \sum w_{ij}x_ix_j + \frac{1}{2}\alpha x_i^2 - \frac{1}{4}x_i^4$, where the last two terms represent double-well potentials that impose spins discreetness $s_i = x_i/|x_i|, x_i \neq 0$, and δ, α are the system dependent parameters [28]. To describe such a

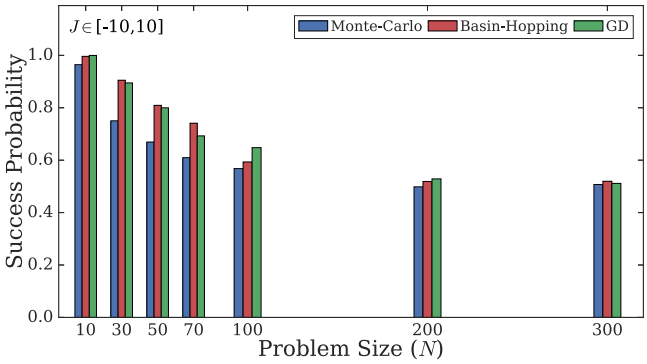

FIG. 2: The success probability of different methods. GD, BH and MC algorithms are used to solve QCO problems for dense matrices (all coupled with all) of sizes $N = \{10, 30, 50, 70, 100, 200, 300\}$ with the elements randomly distributed in $[-10, 10]$. The results are averaged over 30 random coupling matrices. The percentage of iterations that leads to the value of the objective function that is 99% close to the global minimum is shown for each method. The number of internal BH iterations was set to ten to bring about a similar performance to the GD algorithm.

simulator the equations Eqs. (4-5) have to be modified as

$$\dot{\rho}_i(t) = (\gamma_i - \sigma\rho_i)\rho_i + \delta \sum \sqrt{\rho_i \rho_j} w_{ij} x_i x_j + \rho_i \left(\frac{1}{2}\alpha x_i^2 - \frac{1}{4} x_i^4 \right), \tag{11}$$

$$\dot{x}_i(t) = -U\rho_i + 2\delta \sum \frac{\sqrt{\rho_j}}{\sqrt{\rho_i}} w_{ij} x_j + \alpha x_i - x_i^3, \tag{12}$$

where $\delta = \text{const}, w_{ij} = J_{ij}$ for the simulators with dissipative coupling and $\delta = \gamma_i^{\text{inj}}(t), w_{ij} = K_{ij}$ for the simulators with gain coupling. The system of Eqs. (11, 12) with white noise and together with the feedback mechanisms given by Eq. (6) or Eq. (6, 10) realises the global minimum of the QUBO problem at the steady state.

Large-Scale Non-linear Optimisation

To confirm that the GD simulators find the global minimum of the QCO problems we compared the solutions found by the GD algorithms with those found by two global classical optimisers such as the direct Monte Carlo sampling (MC) and the basin-hopping (BH) algorithm (see Methods) when solving QCO for various matrix sizes N using the parameters listed in [29]. For each N, we generate 30 real symmetric coupling matrices $\mathbf{J} = \{J_{ij}\}$ with elements that are randomly distributed in $[-10, 10]$. For each such matrix, we run the BH and MC algorithms starting from 500 random initial conditions and the GD algorithm

starting from zero initial conditions and 500 different noise seeds. For each method, the success probabilities to find a solution within 99% of the global minimum are shown in Fig. 2. The success probabilities of the GD algorithm are similar to both comparison methods for all considered matrix sizes. The spin configurations of the global minimum (the "minimisers") are the same as well, which we show for $N = 100$ in Fig. 3. Also shown on Fig. 3 are the values of the global minimum of the objective function found by GD and the comparison methods that match to ten significant digits.

The distribution of success probabilities over the various matrix instances is shown in Fig. 4(a-c) for $N = 50$ and suggests that the randomly generated matrices of couplings have very narrow spectral gap so the distributions are densely packed for probabilities over 93% for the MC, 96% for the BH, and 95% for GD algorithm. To separate the absolute minimum from the low energy excited states, we also consider sparse coupling matrices where each CC is connected to exactly three other CCs with the coupling strengths randomly generated from the interval with the bounds that are randomly taken from $\{-10, -3, 3, 10\}$. This structure makes the barriers between local and global minima higher, and, therefore, worsens the performance of BH and MC methods, but does not affect the GD algorithm that much as Figs. 4(d,e,f) illustrate. The further advantages of the GD algorithm over the best classical global optimisers for some special types of the coupling matrices are elucidated elsewhere [30].

The number of iterations one needs to reliably find the global minimum grows with the size of the problem N. This growth is expected to be exponential for any algorithm (if P≠NP). However, we can compare how time per iteration grows with the problem size for

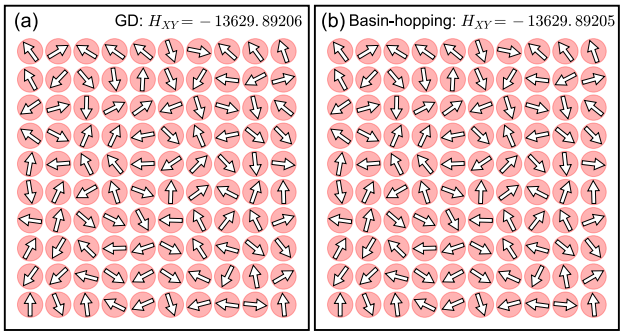

FIG. 3: The spin configurations of the absolute minimum of XY Hamiltonian with the system size $N = 100$ found by (a) GD and (b) BH algorithms. All spins are coupled with all the other spins with the coupling strengths being randomly distributed in $[-10, 10]$.

considered algorithms. We perform the complexity analysis per iteration on mid- and large-scale problems and summarise the results in Fig. 5. The GD algorithm demonstrates the consistent speedup over BH algorithm for all problem sizes N in Fig. 5(a). The log plot in Fig. 5(b) indicates that both algorithms show polynomial time per iteration with the complexity of GD algorithm being close to $\mathcal{O}(N^{2.29})$.

Projected performance of the GD simulators

So far we discussed the implementation of the GD and GD-mod simulators on a classical computer. An actual physical implementation of these simulators will enjoy a super-fast operation and parallelism in processing various phase configurations as the system approaches the global minimum from below even if the system behaves fully classically. Further acceleration could be expected if quantum fluctuations and quantum superpositions contribute to processing of the phase configurations. The times involved into the hardware operation of the GD simulators vary on the scale of pico- to milli-seconds. For instance, in the system of non-degenerate optical parametric oscillators (NOPO) the time-devision multiplexing

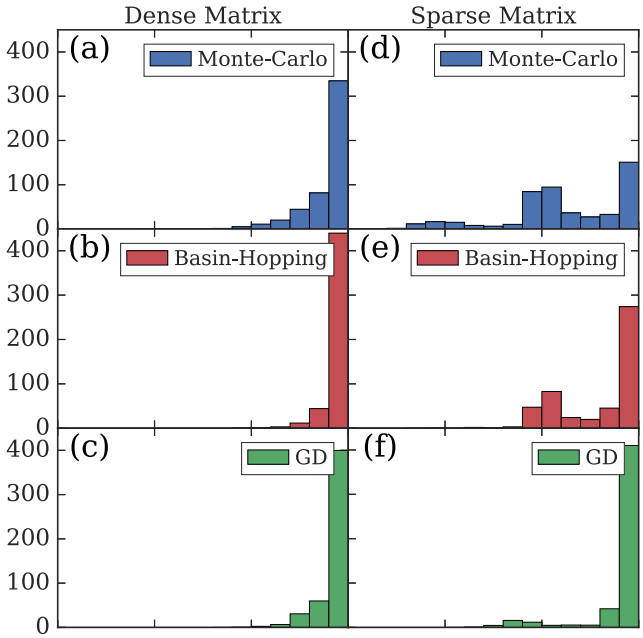

FIG. 4: The success probability of (a,d) MC, (b,e) BH, and (c,f) GD algorithms when solving the QCO problem for the matrix size $N = 50$. The results of 500 runs are averaged over 30 real symmetric coupling matrices $\mathbf{J}$ with the elements randomly distributed in $[-10, 10]$ for (a-c) dense and (d-f) sparse matrices as described in the main text.

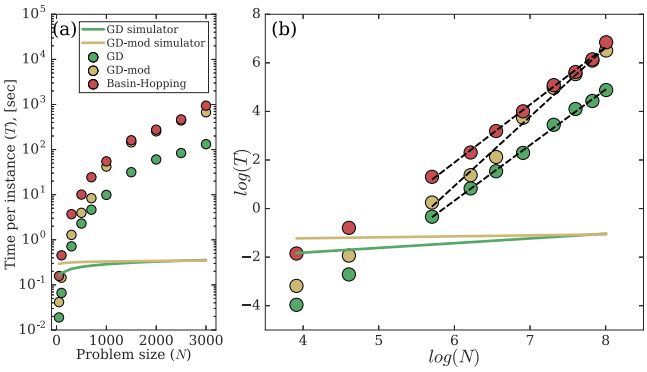

FIG. 5: The performance of the GD, GD-mod, and BH algorithms for solving the QCO problems of sizes N up to 3000. (a) The time per instance T as a function of the problem size N. In the case of GD and GD-mod algorithms, T is the time necessary to reach a stationary state averaged over 20 runs. For the BH algorithm, this time is per ten internal BH iterations necessary to have about the same success probabilities as with the GD algorithm. (b) T as a function of N in the logarithmic scale. The performance of the algorithms are fitted by the linear interpolation functions $2.29 \log N - 13.42$, $2.85 \log N - 16.2$, and $2.38 \log N - 12.39$, for the GD, GD-mod, and BH algorithms, respectively. The projected performance of the GD simulator dominated by the dissipative (gain) coupling is shown with solid green (yellow) lines with the linear asymptotics in (b) given by $0.2 \log N - 2.6$ ($0.04 \log N - 1.38$).

method is used to connect a large number of nodes and the couplings are realised by mutually injecting with optical delay lines with the cavity round trip time being of the order of μs [13], it takes an order of 100 picoseconds for the polariton graphs to condense [15] and 10 ps to 1 ns for photon condensates [16]. The feedback mechanism can be implemented via optical delay lines (in a NOPO system), by adjusting the injection via the spatial light modulator (in the polariton and photon condensates) or by electrical injection (e.g. in the polariton lattices [31]). As the feedback mechanism has to be implemented we need to factor in the time it takes and the necessary number of such adjustments which corresponds to the number of internal iterations for the GD and GD-mod algorithm per each run. By taking an upper limit on the feedback time as $0.1ms$ and using an average number of iterations of the GD and GD-mod algorithms we can estimate the upper bound of the time needed by the physical implementation of the GD simulator to find the global minimum. In Fig. 5 we show such estimates by the solid green (yellow) lines for the GD (GD-mod) simulators. For large N from Fig. 5 we estimate the speed-up of the GD simulators in comparison with the classical computations to be of the order of $10^{-5}N^2 - 10^{-7}N^3$. For N of the order of ten thousands this gives the speed-up of at least four orders of magnitude. Because of the adaptive setting of the coupling matrix in the GD-mod algorithm, the number of internal

iterations grows slower with the size of problem N than for the GD algorithm so that the performance of GD simulator with dissipative couplings can possibly be surpassed by the GD simulator with gain couplings for large N.

Conclusions

Motivated by a recent emergence of a new type of analog Hamiltonian optimisers – the gain-dissipative simulators – we formulate a novel gain-dissipative algorithm for solving NP-hard problems such as constant modulus quadratic continuous and quadratic binary optimisation that can be realised by gain-dissipative simulators of different nature. We show that the feedback mechanism between the CCs density and the pumping intensity is necessary for the operation of such simulators. When the run-time of the classical algorithm is interpreted in terms of the time of the actual operation of the physical system one would expect such simulators to greatly outperform the classical computer starting with thousands of nodes.

In this paper we discussed gain-dissipative systems such as optical cavities, lasers, and non-equilibrium condensates. The physics of the gain-dissipative oscillators may cover a much large variety of systems such as electronic circuits, voltage controlled oscillators, microelectromechanical systems, spin-torque nano-oscillators, oxide-based oscillators, etc. Our analysis is applicable to a broad family of oscillatory networks regardless of the nature of each oscillator as long as it is governed by the generic Eqs. (8) or (9), may be operated stably at the threshold and allows for the feedback mechanism of adjusting the pumping intensity of an individual CCs, similar to the operation of the considered systems.

Finally, we would like to comment on classical vs quantum operation of such simulators. When a condensate (a coherent state) is formed – the system behaves classically as many bosons are in the same single-particle mode and non-commutativity of the field operators can be neglected. However, the condensation process by which the global minimum of the XY model is found involves quantum effects. It was shown before, that the condensation process can be described by a fully classical evolution of the Nonlinear Schrödinger equation that takes into account only stimulated scattering effects and neglects spontaneous scattering [32]. The classical or quantum assignment to gain-dissipative simulators depends on whether quantum fluctuations and spontaneous scattering effects during the condensation provide a

speed-up in comparison with fully classical noise and stimulated scattering. This is an important question to address in the future research on such simulators and the comparison with the classical algorithm that we developed based on the gain-dissipative simulators architecture allows one to see if the time to find the solution scales better than with the best classical algorithms.

Acknowledgements

The authors acknowledge financial support from the NGP MIT-Skoltech. K.P.K. acknowledges the financial support from Cambridge Trust and EPSRC.

Methods

Performance of the GD algorithm

To characterise the performance of the GD algorithm, we compared it to Basin-Hopping and Monte-Carlo algorithms. Both methods depend on a local minimisation algorithm for the optimal decent to a local minimum at each iteration. We considered several local minimisation methods as applied to the QCO problems and determined that the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS-B) [33, 34] has shown the best performance (see Supplementary Information). The L-BFGS-B algorithm is a local minimisation solver which is designed for large-scale problems and shows a good performance even for non-smooth optimisation problems [33, 34]. At each iteration of the MC algorithm, we generate a random starting point and use L-BFGS-B algorithm to find the nearest local minimum. These minima are compared to find the global minimum. The BH algorithm is a global minimisation method that has been shown to be extremely efficient for a wide variety of problems in physics and chemistry [35] and to give a better performance on QCO than other heuristic methods such as simulated annealing [36]. It is an iterative stochastic algorithm that at each iteration uses a random perturbation of the coordinates with a local minimisation followed by the acceptance test of new coordinates based on the Metropolis criterion. Again L-BFGS-B algorithm has shown the best performance as a local optimiser at each step of the BH algorithm. Both BH and MC algorithms were supplied with the analytical Jacobian of the objective function for better performance results.

System of gain-dissipative coherent centers with gain couplings

Here we derive the equations for the coherent centers (CCs) starting with the mean-field Ginzburg-Landau equations. Such derivation has been done before using the infinite quantum well orthogonal basis [37] which is not natural for the localized condensates in the lattice. As the result, the phases of the time-dependent modes are not directly related to the phases of the condensates. Instead, we will use the natural basis associated with the wavefunction of an individual condensate.

The mean–field description of the gain-dissipative condensates is provided by the Ginzburg-Landau equations [27]

$$\mathrm{i}\frac{\partial\psi}{\partial t} = -\nabla^2\psi + (\tilde{U} - i\tilde{\sigma})|\psi|^2\psi + \mathrm{i}(\Gamma(\mathbf{r},t) - \gamma_c)\psi, \tag{13}$$

where $\psi(\mathbf{r},t)$ is the wavefunction of the system, $\tilde{U}$ is the strength of the δ – function interaction (pseudo) potential, $\Gamma(\mathbf{r},t)$ represents the rate of adding particles to the system (pumping), γ_c is the rate of linear losses and $\tilde{\sigma}$ is the rate of the density-dependent losses. In writing Eq. (13) we let $\hbar = 1, m = 1/2$. For the network of N CCs we introduce the pumping profile that creates particles in N spatial locations $\mathbf{r} = \mathbf{r}_i$, $i = 1, ..., N$

$$\Gamma(\mathbf{r},t) = \sum_{i=1}^{N} \gamma_i^{\mathrm{inj}}, \quad \gamma_i^{\mathrm{inj}} = P(|\mathbf{r} - \mathbf{r}_i|) f_i(t), \tag{14}$$

where $f_i(t)$ is the time-dependent part of the pumping and $P_i(\mathbf{r}) \equiv P(|\mathbf{r} - \mathbf{r}_i|)$ is a given spatially localised pumping profile. If the distances between CCs are larger than the width of $P(r)$, the wavefunction of the condensate can be written as

$$\psi(\mathbf{r},t) = \sum_{i=1}^{N} a_i(t)\phi_i(\mathbf{r}), \tag{15}$$

where $\phi_i(\mathbf{r}) = \phi(|\mathbf{r} - \mathbf{r}_i|)$ is the wavefunction of a single localised condensate for the steady localised pumping profile $P(r)$, normalised so that $2\pi \int |\phi(r)|^2 r\, dr = 1$. We substitute the expression for ψ into Eq. (13), multiply by ϕ_i^* for $i = 1, ..., N$, and integrate in space to yield

the rate equations on $a_i(t)$

$$i\dot{\boldsymbol{a}}^T = \left[-\boldsymbol{a}^T\boldsymbol{D} + i\boldsymbol{a}^T\boldsymbol{\Gamma} - i\gamma_c\boldsymbol{a}^T\boldsymbol{\Xi} + (\tilde{U} - i\tilde{\sigma})\boldsymbol{Q}\right]\boldsymbol{\Xi}^{-1}, \tag{16}$$

where $\boldsymbol{a} = \{a_i\}$, $\boldsymbol{\Xi} = \{\chi_{ij}\}$, $\boldsymbol{D} = \{d_{ij}\}$, $\boldsymbol{\Gamma} = \{\Gamma_{ij}\}$, $\boldsymbol{Q} = \{q_{ij}\}$, where $\chi_{ij} = \int \phi_j\phi_i^*\, d\mathbf{r}$ ($\chi_{ii} = 1$), $q_{ij} = a_j^2 a_i^* \int \phi_j^2(\phi_i^*)^2\, d\mathbf{r} + 2|a_i|^2 a_j \int |\phi_i|^2\phi_i\phi_j\, d\mathbf{r} + ...$ for $i \neq j$, $q_{ii} = |a_i|^2 a_i \int |\phi_i|^4\, d\mathbf{r}$, $\Gamma_{ij} = \sum_k f_k \int P_k\phi_j\phi_i^*\, d\mathbf{r}$, and $d_{ij} \equiv \int \phi_i^*\nabla^2\phi_j\, d\mathbf{r}$. An asymptotics of $\phi(r)$ for the Gaussian pumping profile have been developed in [38]. Function ϕ can be approximated by

$$\phi(r) = \beta^2 \exp[-\beta r + ik_c r], \tag{17}$$

where k_c is the outflow velocity and β is the inverse characteristic width of the condensate [38]. The integrals $\chi_{ij} = \int \phi_j\phi_i^*\, d\mathbf{r}$ for $i \neq j$ can be evaluated in elliptical coordinates in terms of the Bessel functions [39]

$$\chi_{ij} = \beta^2 l_{ij}\left[\frac{1}{\beta}J_0(k_c l_{ij})K_1(\beta l_{ij}) + \frac{1}{k_c}J_1(k_c l_{ij})K_0(\beta l_{ij})\right], \tag{18}$$

where $l_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. We assumed that the CCs are well separated: $l_{ij}\beta \gg 1$, so that for $i \neq j$ we have $\chi_{ij} \ll 1$ as follows from Eq. (18). The integrals in the off-diagonal terms in $\boldsymbol{Q}$ are of the same order in smallness as χ_{ij}. Since at the threshold $|a_i(t)|^2$ are small, the off-diagonal terms in $\boldsymbol{Q}$ are quadratic in small quantities and can be neglected to the linear order in Eq. (16), so that $\boldsymbol{Q}$ is the diagonal matrix with elements $|a_i|^2 a_i q$, where $q = 2\pi \int |\phi|^4 r\, dr$. Using Eq. (17) one can show that $d_{ij} \ll \chi_{ij}$ for $i \neq j$ and can be neglected to the linear order, so that $\boldsymbol{D} = d\boldsymbol{I}$, where $d = 2\pi \int \phi\nabla^2\phi r\, dr$. The matrix $\boldsymbol{\Xi}$ has 1 on the diagonal and χ_{ij} as $ij-$th element. Taking the inverse and keeping only up to the linear order terms gives a matrix with 1 on the diagonal and $-\chi_{ij}$ as the $ij-$th element. Finally, to the linear order terms the elements of matrix $\boldsymbol{\Gamma}$ are $\Gamma_{ii} = f_i p$ and $\Gamma_{ij} = (f_i p_{ij} + f_j p_{ij}^*)$ for $i \neq j$, where $p = 2\pi \int P(r)|\phi|^2 r\, dr$, and $p_{ij} = \int P_j\phi_j\phi_i^*\, d\mathbf{r}$. To the linear order terms in small quantities $|a_i(t)|^2$, p_{ij} and χ_{ij} for $i \neq j$ Eqs. (16) become

$$\dot{a_i} = \mathrm{i}da_i + a_i(f_i(t)p - \gamma_c) + f_i(t)\sum_{i\neq j} a_j p_{ij} - \mathrm{i}q(\tilde{U} - i\tilde{\sigma})|a_i|^2 a_i, \tag{19}$$

where we also used the observation that terms $p_{ij}^* - p\chi_{ij}$ are quadratic in small quantities.

Equation (19) coincides with Eq. (9) if one adds noise, let $\Psi_i = a_i \exp(\mathrm{i}dt)$, $U = q\tilde{U}$, $\sigma = q\tilde{\sigma}$, $\gamma_{\mathrm{inj}} = f_i(t)$, and $K_{ij} = p_{ij}$.

Supplementary Information

Quadratic constrained optimisation (QCO) problems

QCO and QUBO problems are known to be strongly NP-hard in general [18] (with the decision problem to be NP-complete), meaning that an efficient way of solving the QCO or QUBO problem can be used to solve all problems in the complexity class NP, which includes a vast number of important problems such as partitioning, the travelling salesman problem, graph isomorphisms, factoring, nonlinear optimisation beyond quadratic, etc. The NP-hardness also implies that starting with medium-sized matrices QCO or QUBO problems are difficult to solve on a classical computer as the number of required operations grows exponentially fast with N. The actual time required to find the solution also depends on the matrix structure. For instance, for positive definite matrices, the QCO problem remains NP-hard due to the non-convex constraints but can be effectively approximated using an SDP relaxation [40] with the performance guarantee $\pi/4$ [18]. Sparsity also plays an important role: for sufficiently sparse matrices fast methods exist [41]. As for many other hard optimisation problems, there are three types of algorithms for solving QCO problems on a classical computer: exact methods that find the optimal solution to the machine precision, approximate algorithms that generate the solution within a performance guarantee and heuristic algorithms where suitability for solving a particular problem comes from some empirical testing [42]. Exact methods can be used to solve small to medium matrix instances, as they typically involve branch-and-bound algorithms and the exponential worst-case runtime. The heuristic algorithms such as simulated annealing can quickly deliver a decent, but suboptimal (and possibly infeasible) solution [43]. Finally, QCO is known to be in the APX-hard class of problems [44], so there is no polynomial-time approximation algorithm that gives the value of the objective function that is arbitrarily close to the optimal solution (unless P = NP). The problem becomes even more challenging when the task is to find not an approximation to the global minimum of the objective function, but the minimisers as needed for instance in image reconstruction. The values of the objective

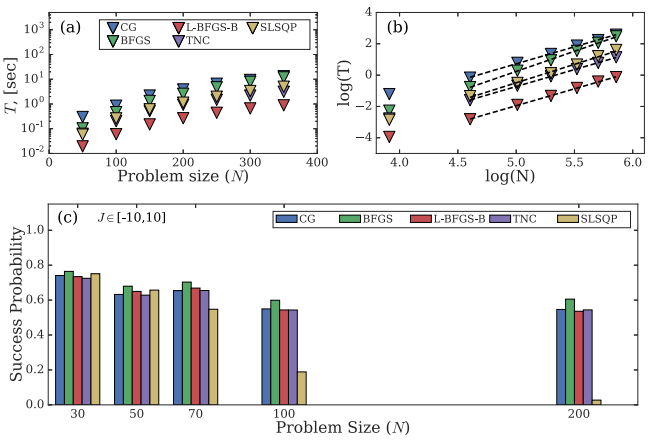

FIG. 6: The performance of the various local optimisers when solving QCO problem of sizes up to $N = 200$. The run-time dependence on the matrix size N is shown in (a) and in a log scale in (b). The success probability of 99% is shown in (c) where each algorithm starts from the same 100 random initial states. The probabilities are averaged over 25 dense coupling matrices with randomly generated elements in $[-10, 10]$.

functions can be very close, but for the entirely different sets of minimizers.

The interest in solving QCO problems is justified in its own right. These problems appear in many real-life applications in a context of a phase retrieval – recovering a general signal from the magnitude of its Fourier transform – which is perceived as a notoriously hard optimisation problem [40]. It arises in X-ray crystallography [45], diffraction imaging [46], astronomical imaging [47], optics [48], microscopy [49], biomedical applications [50].

Performance of local optimisation algorithms for QCO problem

At each iteration of the Monte-Carlo and the Basin-Hopping methods we use the L-BFGS-B algorithm, since it has shown better results for the QCO problems in terms of both performance and the quality of solution compared to other available algorithms in scipy.optimize.minimize library such as the sequential least squares programming (SLSQP), nonlinear conjugate gradient algorithm (CG), truncated Newton (TNC) algorithm, and BFGS. In comparison with the BFGS algorithm, the L-BFGS-B (limited memory BFGS) algorithm exploits an estimation of the inverse Hessian matrix. Each algorithm was supplied with the analytical Jacobian. The performance of the algorithms is shown in Fig. 6. The L-BFGS-B algorithms is the fastest in comparison with all the other algorithms (see Fig. 6(a,b)) while the success probabilities are comparable (Fig. 6(c)).

All numerical calculations were performed on a MacBook Pro (15-inch, 2016), 2.7 GHz Intel Core i7, 16 GB 2133 MHz LPDDR3.

---

[1] Brewster, D. Planetary machines. *The Edinburgh Encyclopedia* **16,** 624 (1830).

[2] Thomson, J. On an integrating machine having a new kinematic principle. *Van Nostrand's Eclectic Engineering Magazine* **15(94),** 301(1876).

[3] Cairns, W. J., Crank, J., & Lloyd, E. C. Some improvements in the construction of a small scale differential analyser and a review of recent applications. *Armament Res. Dept., Theoretical Res. Memo* **27,** 44 (1944).

[4] Feynman, R. R. Simulating physics with computers. *Int. J. Theor. Phys.* **21,** 467 (1982).

[5] Georgescu, I. M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86,** 153 (2014).

[6] Yung, M. H., Nagaj, D., Whitfield, J. D., & Aspuru-Guzik, A. Simulation of classical thermal states on a quantum computer: A transfer-matrix approach. *Physical Review A* **82(6),** 060302 (2010).

[7] Sinha, S. & Russer, P. Quantum computing algorithm for electromagnetic field simulation. *Quantum Information Processing* **9(3),** 385-404 (2010).

[8] Johnson, M. W. *et al.* Quantum annealing with manufactured spins. *Nature* **473(7346),** 194-198 (2011).

[9] Utsunomiya, S., Takata, K. & Yamamoto, Y. Mapping of Ising models onto injection-locked laser systems. *Opt. Express* **19**, 18091(2011).

[10] Marandi, A., Wang, Z., Takata, K., Byer, R.L. & Yamamoto, Y. Network of time-multiplexed optical parametric oscillators as a coherent Ising machine. *Nat. Phot.* **8**, 937-942 (2014).

[11] Inagaki, T. *et al.* Large-scale Ising spin network based on degenerate optical parametric oscillators. *Nature Photonics* **10**,415?419 (2016).

[12] McMahon, P.L. *et al.* A fully programmable 100-spin coherent Ising machine with all-to-all connections. *Science*, **354** 614-617 (2016).

[13] Takeda, Y. *et al.* Boltzmann sampling for an XY model using a non-degenerate optical parametric oscillator network. *Quantum Science and Technology* **3(1),** 014004 (2017).

[14] Nixon, M., Ronen, E., Friesem, A. A. & Davidson, N. Observing geometric frustration with thousands of coupled lasers. *Phys. Rev. Lett.* **110**, 184102 (2013).

[15] Berloff, N. G. *et al.* Realizing the classical XY Hamiltonian in polariton simulators. *Nat. Mat.* **16(11)**, 1120 (2017).

[16] Dung, D. *et al.* Variable potentials for thermalized light and coupled condensates. *Nat. Phot.* **11(9)**, 565 (2017).

[17] Kalinin, K. P. & Berloff, N. G. Blockchain platform with proof-of-work based on analog Hamiltonian optimisers. *arXiv:1802.10091* (2018).

[18] Zhang, S. & Huang, Y. Complex quadratic optimization and semidefinite programming. *SIAM J. Optim.* **16**, 871 (2006).

[19] Leliu, T. *et al.* Combinatorial optimization using dynamical phase transitions in driven-dissipative systems. *Phys. Rev. E* **95,** 022118 (2017).

[20] Gardiner, C. & Zoller, P. Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics (Vol. 56, Springer Science & Business Media, Berlin, 2004).

[21] Aleiner, I. L., Altshuler, B. L., & Rubo, Y. G. Radiative coupling and weak lasing of exciton-polariton condensates. *Physical Review B* **85(12),** 121301 (2012).

[22] Kuramoto, Y. Chemical oscillations, waves, and turbulence (Dover, New York, 2012).

[23] Pikovsky, A., Rosenblum, M., & Kurths, J. Synchronization (Cambridge University Press, Cambridge, England, 2001)

[24] Kourtchatov, S. Y., Likhanskii, V. V., Napartovich, A. P., Arecchi, F. T., & Lapucci, A. Theory of phase locking of globally coupled laser arrays. *Physical Review A* **52(5),** 4089 (1995).

[25] Kozyreff, G. *et al.* Global coupling with time delay in an array of semiconductor lasers. *Phys. Rev. Letts* **85,** 3809 (2000).

[26] Oliva, R. A., & Strogatz, S. H. Dynamics of a large array of globally coupled lasers with distributed frequencies. *International journal of Bifurcation and Chaos* **11(09),** 2359-2374 (2001).

[27] Keeling, J. & Berloff, N. G. Spontaneous rotating vortex lattices in a pumped decaying condensate. *Physical review letters* **100(25),** 250401 (2008).

[28] Leleu, T., Yamamoto, Y., Utsunomiya, S., & Aihara, K. Combinatorial optimization using dynamical phase transitions in driven-dissipative systems. *Physical Review E* **95(2),** 022118 (2017).

[29] The numerical parameters and the initial conditions for GD algorithm Eqs. (6,8) are as follows: $\epsilon = \epsilon_0 \max_i \sum_j |J_{ij}|$, $\rho_{th} = \rho_{0,th} \max_i \sum_j |J_{ij}|$, $U = 0$, $\sigma = 1$, $\gamma(t = 0) = -\max_i \sum_j |J_{ij}|$, $\rho_i(t = 0) = 0$, $\theta_i(t = 0) = 0$, $\epsilon_0 = 0.05$, $\rho_{0,th} = 0.1$. For GD-mod algorithm Eqs. (6,9, 10) we choose $\hat{\epsilon} = 1$ in addition to the parameters listed above.

[30] K.P.Kalinin & N.G.Berloff. Gain-dissipative simulators for solving the NP-hard problems. in preparation (2018).

[31] Suchomel, H. *et al.* An electrically pumped polaritonic lattice simulator. arXiv:1803.08306 (2018)

[32] Berloff, N. G. & Svistunov, B. V. Scenario of strongly nonequilibrated Bose-Einstein condensation. *Physical Review A* **66(1),** 013603 (2002).

[33] Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **16(5),** 1190-1208 (1995).

[34] Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)* **23(4),** 550-560 (1997).

[35] Wales, D. J., & Doye, J. P. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A* **101(28),** 5111-5116 (1997).

[36] Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. Optimization by simulated annealing. *Science* **220(4598),** 671-680. (1983).

[37] Sigurdsson, H., Liew, T. C., & Shelykh, I. A. Parity solitons in nonresonantly driven-dissipative condensate channels. *Physical Review B* **96(20),** 205406 (2017).

[38] Lagoudakis, P. G. & Berloff, N. G. A polariton graph simulator. *New J. Phys.* **19**, 125008 (2017).

[39] Kalinin, K. P., Lagoudakis, P. G., & Berloff, N. G. Matter wave coupling of spatially separated and unequally pumped polariton condensates. *Physical Review B* **97(9),** 094512 (2018).

[40] Candes, E. J., Eldar, Y. C., Strohmer, T., & Voroninski, V. Phase retrieval via matrix completion. *SIAM review* **57(2),** 225-251 (2015).

[41] Shechtman, Y., Beck, A., & Eldar, Y. C. GESPAR: Efficient phase retrieval of sparse signals. *IEEE transactions on signal processing* **62(4),** 928-938 (2014).

[42] Dunning, I., Gupta, S. & Silberholz, J. What Works Best When? A Systematic Evaluation of

Heuristics for Max-Cut and QUBO. To appear in INFORMS Journal on Computing, (2018).

[43] Kochenberger, G. *et al.* The unconstrained binary quadratic programming problem: a survey. *J Comb Optim* **28**, 5881 (2014).

[44] Papadimitriou, C. H. & Yannakakis, M. Optimization, approximation, and complexity classes. *J. Comput. Syst. Sci.* **43(3)**, 425440 (1991).

[45] Harrison, R. W. Phase problem in crystallography. *JOSA a* **10(5),** 1046-1055 (1993).

[46] Bunk, O. *et al.* Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels. *Acta Crystallographica Section A: Foundations of Crystallography* **63(4),** 306-314 (2007).

[47] Fienup, C. & Dainty, J. Phase retrieval and image reconstruction for astronomy. *Image Recovery: Theory and Application* **231,** 275 (1987).

[48] Walther, A. The question of phase retrieval in optics. *Optica Acta: International Journal of Optics* **10(1),** 41-49 (1963).

[49] Miao, J., Ishikawa, T., Shen, Q., & Earnest, T. Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes. *Annu. Rev. Phys. Chem.* **59,** 387-410 (2008).

[50] Dierolf, M. *et al.* Ptychographic X-ray computed tomography at the nanoscale. *Nature* **467(7314),** 436 (2010).