Combinatorial optimization using the Lagrange primal-dual dynamics of parametric oscillator networks

Sri Krishna Vadlamani[*]

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Tianyao Patrick Xiao

Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

Eli Yablonovitch

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, California 94720, USA

Coupled-bistable-oscillator machines have recently generated significant interest due to their observed ability to rapidly produce high-quality solutions to nondeterministic-polynomial-time-complete optimization problems. While the dynamics of such systems are often derived in the literature, it has hitherto been unclear why exactly the system dynamics perform optimization so well. This paper answers this question by presenting a complete equivalence between coupled-oscillator machines and the primal-dual method of Lagrange multipliers. This equivalence explains how coupled-oscillator solvers implement the correct optimization constraints and find high-quality solutions. The equivalence also provides precise mathematical meaning to each system component and enables the principled design of systems that implement more-sophisticated optimization algorithms. We simulate the system dynamics and demonstrate (1) that its performance is competitive with performance of the best-known digital algorithms, (2) that the circuit is robust with regard to large component variations, hinting that the traditional shortcomings of analog computation may be less important for these applications, and (3) that the circuit consumes extremely low amounts of power (on the order of milliwatts) and energy (approximately 100 nJ) per optimization even for problems of 2000 variables.

I. INTRODUCTION

Combinatorial optimization encompasses several important problems, such as the traveling-salesman problem, the knapsack problem, integer programming, protein folding, and many more that are of great interest to the operations-research, computer-science, and other scientific communities. Unfortunately, many of these problems have been shown to be nondeterministic polynomial-time (NP) complete [1]; consequently, it has so far proved extremely difficult to devise fast polynomial-time algorithms to solve them. Conventional digital algorithms are based on either provable approximations that provide a lower bound on solution quality [2] or heuristics (and metaheuristics) to search for higher-quality solutions [3].

There has been significant recent interest in the idea that physical dynamics can possibly obtain high-quality approximate solutions to these problems by exploiting built-in physical maximum and minimization principles such as least action, least time, free-energy minimization, and minimum-entropy generation. One way of doing this is to first establish a correspondence between the mathematical variables and the objective function, on one hand, and the physical variables and the properties of an appropriate physical system, on the other; the next step would then be to try to embed an existing or novel optimization algorithm in the system's dynamical equations. Ideally, such a system will be both physically realizable and capable of finding solutions of quality comparable to or superior to solutions found by state-of-the-art digital algorithms.

In the literature, physics-based optimization machines have often been designed to tackle the magnetic Ising problem, relying on the fact that any NP-complete problem

---

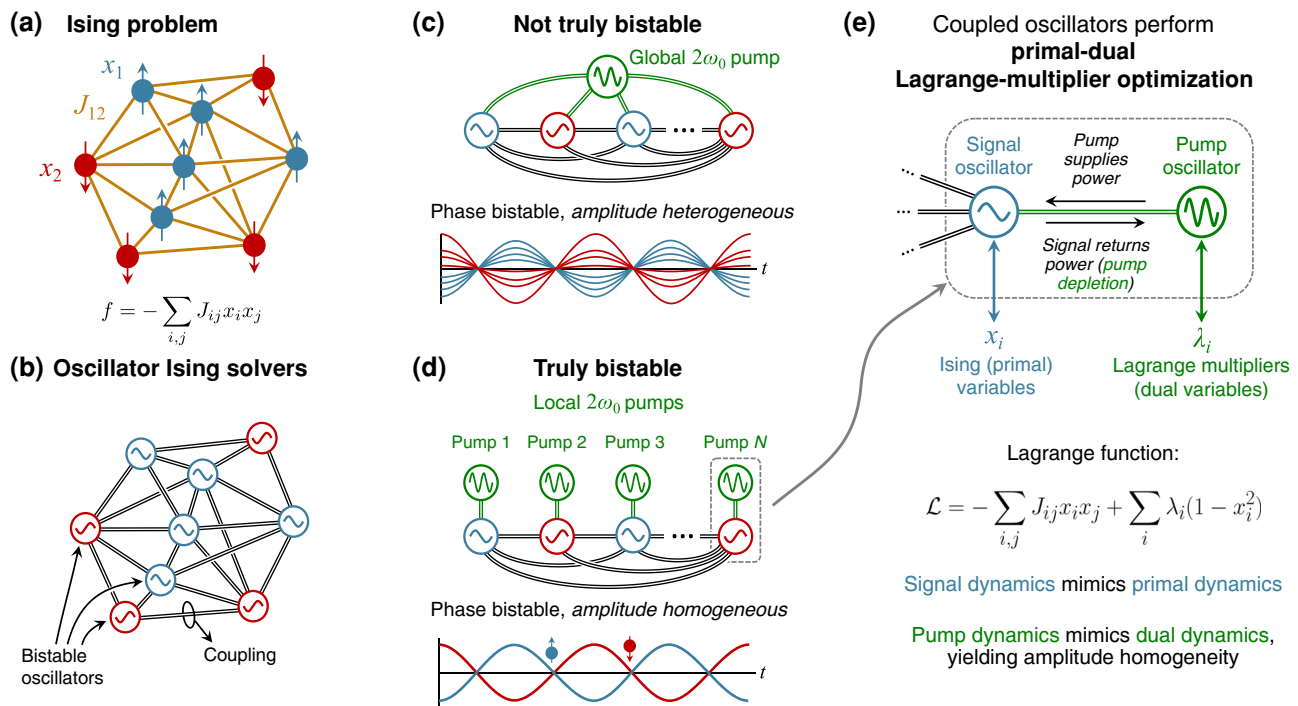[*]Corresponding author: srikv@mit.edu

FIG. 1.   (a) The Ising problem asks for the configuration of binary spins that minimizes the quadratic Ising interaction energy of the spins. (b) Oscillator Ising solvers represent spins by bistable oscillators that are coupled to one another to implement the Ising coupling J_{ij}. (c) Earlier Ising solvers such as the original coherent Ising machine [5] used a single pump to excite all the spin signal oscillators, leading to phase bistability but amplitude heterogeneity—this is not the true bistability that the Ising problem requires. (d) In our Ising solver, each signal oscillator is excited by its own separate pump oscillator—this yields true bistability, i.e., both phase bistability and amplitude homogeneity. (e) In such a system, we show in this paper that the signal oscillators naturally behave as the Ising variables, while the pump oscillators act as the corresponding Lagrange multipliers. The pumps continuously adjust themselves, through pump depletion, to enforce the Ising optimization constraints, i.e., $x_i^2 = 1$ for all i. The evolution of both types of oscillator is determined by the system Lagrange function.

can be reduced to the Ising problem in polynomial time [4]. The Ising problem [Fig. 1(a)] consists of a set of N interacting spins, each of which has two possible orientations, and the challenge is to find the configuration of spins **x** that minimizes the total interaction energy f, given by

$$f(\mathbf{x}) = -\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij} x_i x_j, \qquad (1)$$

where $x_i = \pm 1$ is the binary orientation of the ith spin and J_{ij} is the Ising interaction strength between the ith spin and the jth spin. Ising solvers are typically benchmarked on instances of the equivalent max-cut problem.

Published physical Ising solvers include quantum [6] and classical [7–9] machines that use the adiabatic principle, Hopfield/Boltzmann-machine hardware solvers that minimize the Ising energy through simulated annealing by performing iterative matrix operations to update the binary variables [10–17], memcomputing [18,19], chaotic dynamical systems [20–23], and a number of approaches based on coupled bistable dynamical elements, such as stochastic magnetic bits [24,25] or coupled oscillators.

Several coupled-oscillator systems have been proposed, analyzed [26–28], and physically implemented: laser parametric oscillators [5,29,30], injection-locked *LC* oscillators [31,32], CMOS ring oscillators [33,34], phase-transition oscillators [35], coupled multicore-fiber lasers [36], and coupled polaritonic cavities [37]. While the dynamics of these coupled-bistable-oscillator systems are often derived in the literature, it has so far not been clear what exactly those dynamics mean in terms of optimization—why do coupled-bistable-oscillator solvers obtain the high-quality results that they do? This paper resolves this issue by demonstrating an exact mathematical connection between coupled-bistable-oscillator solvers and the method of Lagrange multipliers.

In the coupled-bistable-oscillator approaches, each oscillator is forced into bistability [Fig. 1(b)] by a physical nonlinearity, and chooses one of the two stable states on the basis of the strengths of its interactions with other oscillators. Nonlinear parametric amplification can produce phase-bistable oscillations at frequency ω_0 via the coherent interaction between a $2\omega_0$ second-harmonic pump oscillator and a ω_0 signal oscillator. However, phase bistability by itself is insufficient to represent Ising spins;

one also needs amplitude homogeneity to achieve true bistability [20].

In this paper, we show that the dynamics of general coupled-parametric-oscillator systems map exactly onto the differential equations of the "primal-dual" method of Lagrange multipliers. The ω_0 signal oscillator dynamics in each spin maps exactly onto the dynamics of the Ising variables x_i, also called the "primal variables," while the $2\omega_0$ pump oscillator dynamics in each spin maps exactly onto the dynamics of the corresponding Lagrange multipliers λ_i, also called the "'"dual variables" [Fig. 1(e)]. In other words, the signal oscillators implement the primal step, while the pump oscillators implement the dual step in the primal-dual method; these terms and the method are explained in detail in later sections.

It is known in optimization theory that the purpose of the dual dynamics in the Lagrange-multiplier method is to enforce constraints, which boils down to enforcing amplitude homogeneity $x_i^2 = 1$ in the case of the Ising problem where the spins are binary by definition. Since we show that the $2\omega_0$ pump dynamics (also known as "pump depletion") map exactly onto the dual dynamics, this means that the pump oscillators themselves lead to amplitude homogeneity, and thus to true bistability, in their associated signal oscillators. We note that it is essential that each signal oscillator have its own corresponding pump oscillator [Fig. 1(d)] for the equivalence to Lagrange multipliers to hold, since each Ising variable x_i has its own corresponding Lagrange multiplier λ_i in the Lagrange-multiplier method. If a global pump is used, as in Fig. 1(c), the system cannot implement N different independent binary constraints, thus leading to the problem of amplitude heterogeneity [5].

Vadlamani *et al.* [38] showed only the primal part (that is, that the signal dynamics and primal dynamics are equivalent) of the Lagrange-multiplier equivalence and proposed that auxiliary hardware would be needed to implement the dual portion. The current paper shows that no additional hardware is needed, and that the intrinsic pump depletion is sufficient to implement the dual dynamics. Moreover, the complete equivalence presented in this paper (1) explains how coupled-oscillator solvers implement optimization constraints and why they find high-quality solutions, (2) provides precise mathematical meaning to each physical component in the circuit, and (3) enables the design of circuits that map onto more-sophisticated optimization algorithms such as the augmented Lagrange primal-dual algorithm. Throughout most of this paper, we use electrical *LC*-oscillator systems as an expository model for general coupled-oscillator systems and to generate numerical results; it should be noted however, that all the results, especially the mathematical equivalence to Lagrange multipliers, are general and are not limited merely to the electrical case. Numerical simulations of the equations of motion reveal that the performance of the oscillator Lagrange-multiplier

approach on benchmark problems is competitive with that of the best-known digital algorithms and is robust with regard to large circuit component variations. Moreover, our simulations (with thermal noise taken into account) reveal that the circuit consumes extremely low amounts of power (on the order of milliwatts) and energy (approximately 100 nJ) per optimization even for problems of size 2000.

This paper expands on our previous theoretical work [39] by providing full derivations of the dynamics and noise models; detailed benchmarking on a diverse set of combinatorial optimization problems; sensitivity analyses to component variations, noise, and dissipative losses; and derivation of technology requirements and scaling trends from these results. The paper is organized as follows: Section II presents the equations of motion for a network of electrical parametric *LC* oscillators, which acts as a proxy for all the coupled-oscillator approaches, and shows that pump depletion constrains the oscillators to fixed-amplitude binary states. Section III briefly reviews the primal-dual Lagrange-multiplier method for constrained optimization. Section IV presents the complete equivalence between the oscillator dynamics and the Lagrange-multiplier method, and discusses an extension to the augmented Lagrange method. In Sec. V, we numerically simulate the circuit dynamics and compare the results against two other algorithms on several large max-cut problems in the Gset and BiqMac problem sets [40,41]. Section VI concludes the paper.

II. ISING-ENERGY MINIMIZATION WITH COUPLED PARAMETRIC *LC* OSCILLATORS

A general degenerate parametric oscillator consists of a signal oscillator at frequency ω_0 and a pump oscillator at frequency $2\omega_0$. The general equations of motion for the slowly varying amplitudes of the two oscillators are given by Yariv and Louisell [42]:

$$\frac{dA_s}{dt} = P_s - \frac{A_s}{\tau_s} + \kappa_s A_s A_p, \tag{2}$$

$$\frac{dA_p}{dt} = P_p - \frac{A_p}{\tau_p} - \kappa_p A_s^2, \tag{3}$$

where A_s and A_p are the signal and pump oscillator amplitudes, P_s and P_p represent power sources inside these oscillators, τ_s and τ_p are the loss rates of each of the oscillators, $\kappa_s A_s A_p$ is the amplitude boost supplied to the signal oscillator by the pump, and $\kappa_p A_s^2$ is the amplitude loss caused in the pump by the signal.

For ease of analysis, throughout the rest of this paper we use a circuit model of electrical parametric oscillators, whose dynamics is shown to be described by the equations above. However, it should be kept in mind that all

the results discussed in this paper transfer readily to coupled parametric oscillators of any type, implemented in any other platform.

To construct the electrical Ising solver, we map an interacting ensemble of N Ising spins to a network of N resistively coupled parametric LC oscillators. This system was discussed in Refs. [38,43,44]. Parametric amplification, enabled by a capacitive nonlinearity in the LC oscillator, ensures that each signal oscillator's steady state is bistable in phase, indicating that these systems can be used to implement binary Ising spins. However, to correctly implement ± 1 spins, one also needs amplitude stability in addition to phase bistability. We show how both these conditions are achieved in this section.

A. Mapping Ising spins to parametric oscillators

A linear LC oscillator supports sinusoidal oscillations $A \cos(\omega_0 t + \phi)$ of arbitrary amplitude A and phase ϕ [Fig. 2(a), left], where $\omega_0 = 1/\sqrt{LC}$ is the natural frequency of the LC cavity. Connecting this oscillator (henceforth called the "signal oscillator") to a pump oscillator at frequency $2\omega_0$ by means of a second-order nonlinear capacitor induces the phenomenon of parametric amplification and introduces constraints on the signal oscillator's phase and amplitude. The parametric oscillator circuit, from Ref. [45], is shown in Fig. 2(b), where the nonlinear capacitor couples the signal and pump oscillators. The nonlinear capacitor's characteristic is as follows:

$$Q = C_0 V_c + C_N V_c^2, \tag{4}$$

where C_0 is the linear capacitance and C_N is the second-order nonlinear capacitance. Such a nonlinear capacitance can be implemented by common semiconductor devices such as p-n-junction or Schottky diodes.

The signal oscillator derives voltage-amplitude gain from the pump oscillator if the signal oscillation's phase is one of two particular values [separated by π radians, represented in blue in Fig. 2(a), middle] dictated by the pump; the signal oscillations that are in quadrature to these special values experience loss. If this phase-sensitive gain exceeds the signal oscillator's resistive losses, only the chosen quadrature survives and the oscillator becomes phase bistable, as shown in the middle in Fig. 2(a).

Phase bistability is not sufficient to implement binary spins—the oscillations must also be amplitude stable [i.e., all oscillators must reach the same saturation amplitude, see Fig. 2(a), right]. When parametric gain is first introduced, the signal oscillator amplitude increases exponentially, drawing more and more power from the pump to sustain its growth. This continues until the pump voltage V_p in Fig. 2(b) gets depleted. The signal and pump then exchange power back and forth until both amplitudes settle around a steady-state value. This pump-depletion
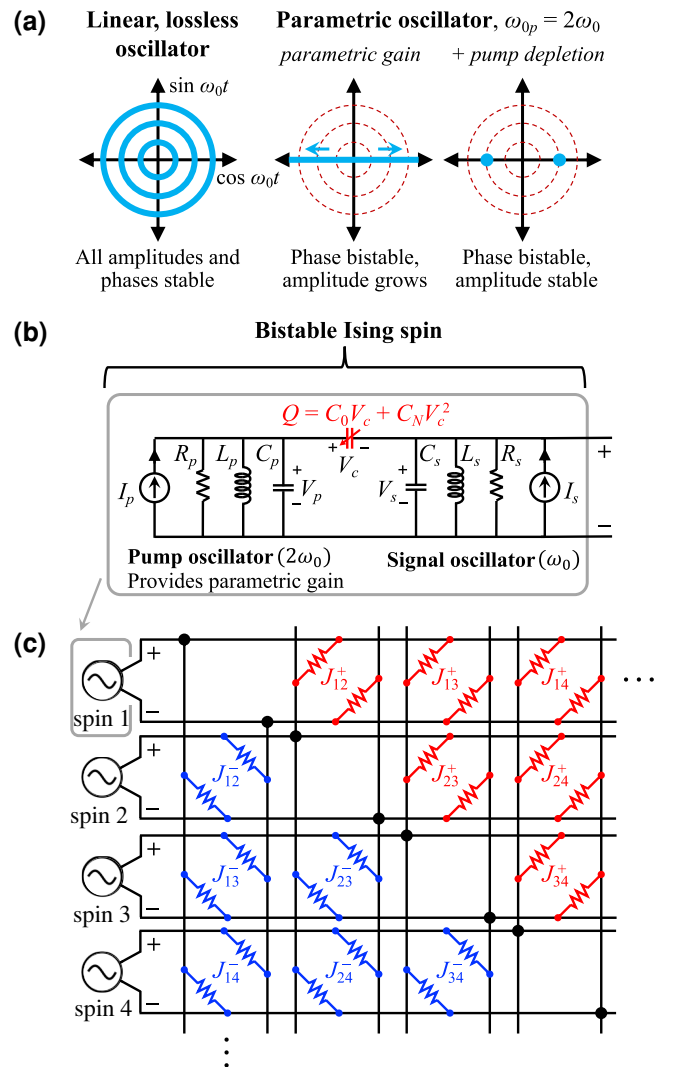

FIG. 2. (a) A linear, lossless signal oscillator (left) supports oscillations of arbitrary phase and amplitude. A parametrically amplified signal oscillator (middle) supports oscillations only at two phases separated by π radians. Signal amplitude is further constrained to a single value by depletion of the pump power (right). (b) Parametric LC oscillators, used as bistable Ising spins, consist of a ω_0 signal oscillator and a $2\omega_0$ pump oscillator connected by a quadratic nonlinear capacitor. (c) All-to-all Ising coupling between spins is implemented via a resistive crossbar. The sign of J_{ij} determines whether the red resistors or the blue resistors are used.

mechanism makes the oscillator truly bistable [Fig. 2(a), right].

The equations of motion for the parametric oscillator circuit in Fig. 2(b) are derived with use of Kirchoff's circuit equations in Appendix A. A key step in the derivation is the use of the well-known slowly-varying-amplitude approximation, which assumes that the amplitude envelopes (A_s and A_p) of the oscillating voltages (V_s and V_p) vary slowly compared with the

frequency of the harmonic oscillations themselves. Under this
approximation, the amplitude envelopes evolve as follows:

$$\frac{dA_s}{dt} = \frac{1}{2}\left[\frac{I_s}{C_0 + C_s} - \frac{A_s}{R_s\left(C_0 + C_s\right)} + \frac{C_N \omega_0 A_s A_p}{C_0 + C_s}\right], \tag{5}$$

$$\frac{dA_p}{dt} = \frac{1}{2}\left[\frac{I_p}{C_0 + C_p} - \frac{A_p}{R_p\left(C_0 + C_p\right)} - \frac{C_N \omega_0 A_s^2}{C_0 + C_p}\right]. \tag{6}$$

While the equations above were derived for the specific
nonlinear-*LC*-oscillator circuit shown in Fig. 2(b), it is
clear that they are identical to the general equations (2)
and (3). Again, in Eqs. (5) and (6), the first term is a
power source, the second term corresponds to internal dis-
sipation, and the third term represents the exchange of
power between the pump and the signal. The signal A_s
has a parametric gain term that is proportional to the pump
amplitude, while the pump A_p has a loss term correspond-
ing to the transfer of energy to the signal oscillator. This
term is responsible for pump depletion, which limits the
parametric gain and the signal amplitude. In the following,
we assume that the signal oscillator does not have its own
power source—instead, I_s corresponds to noise power with
a time-averaged current of zero.

B. Dynamics of dissipatively coupled parametric oscillators

Figure 2(c) shows a scheme, also used by Wang and
Roychowdhury [31], to resistively couple the bistable
parametric *LC* oscillators to implement the spin-spin inter-
actions J_{ij} in the Ising problem. For simplicity, we consider
the case where the interaction weights are binary, $J_{ij} = \pm 1$, but the scheme can straightforwardly be extended to
any intermediate-valued weights (see Appendix B 2).

If two spins i and j are ferromagnetically coupled ($J_{ij} = 1$), a pair of straight-linking resistors [red resistors in
Fig. 2(c)] is used to connect the corresponding oscillators.
This works because, if the oscillators are not in the same
phase, the resultant voltage differences across the connect-
ing resistors will cause current flows that flip the phase
of one of the oscillators. Similarly, antiferromagnetic cou-
pling ($J_{ij} = -1$) is implemented by a pair of cross-linking
connecting resistors [blue resistors in Fig. 2(c)].

The equations of motion for the full network of N
coupled identical parametric *LC* oscillators [Fig. 2(c)]
are derived from Kirchoff's circuit laws in Appendix B
with use of the slow-varying-amplitude approximation. To
derive exact mathematical correspondences in the later
sections, we also make the assumption that the pump inter-
nal resistance R_p is large enough that it can be ignored
relative to the loss of pump energy to the signal oscilla-
tor; the effect of finite R_p is investigated in Sec. V and

Appendix F 2. Now, since introducing the J_{ij} coupling
modifies only the "signal" part of the circuit, which oper-
ates at frequency ω_0, the pump equation remains the same
as Eq. (6), but with R_p dropped and with the appropriate
subscripts i included:

$$\frac{dA_{pi}}{dt} = \frac{1}{2}\left[\frac{I_{pi}}{C_0 + C_p} - \frac{C_N \omega_0 A_{si}^2}{C_0 + C_p}\right]. \tag{7}$$

The more-involved evolution equation for the ith signal
oscillator is as follows:

$$\frac{dA_{si}}{dt} = -\left[\frac{N_i A_{si}}{4RC_{0s}} - \frac{1}{4RC_{0s}}\sum_{j:j\neq i} J_{ij} A_{sj}\right]$$
$$- \frac{A_{si}}{2R_s C_{0s}} + \frac{C_N \omega_0 A_{pi} A_{si}}{2C_{0s}}, \tag{8}$$

where $C_{0s} = C_0 + C_s$, R_s is the signal internal resistance,
R is the coupling resistance, and $N_i := \sum_{j:j\neq i}|J_{ij}|$ is the
number of nonzero connections to the ith oscillator.

The form of Eq. (8) can be generalized to any coupled-
parametric-oscillator network. The term in brackets in Eq.
(8) captures the net loss that the signal amplitude A_{si} expe-
riences due to its connections to the other oscillators, the
middle term is the internal oscillator loss, and the last term
is the parametric gain that is supplied from the pump, as
in Eq. (5). While it is the case that the slowly-varying-
amplitude approximation is widely accepted and Eqs. (8)
and (7) have clear physical interpretations, we neverthe-
less simulated the N-spin circuit in Fig. 2(c) in LTspice,
and we compare its performance with that of the slowly-
varying-amplitude approximation for $N = 5$ and $N = 15$
in Appendix B 5.

C. Rewriting the equations of motion differently

We now re-express the system's dynamics to better
elucidate its algorithmic functionality. First, the pump
dynamics equation (7) can be rewritten as

$$\frac{dA_{pi}}{dt} = \frac{C_N \omega_0}{2\left(C_0 + C_p\right)}\left(A_{\text{sat}}^2 - A_{si}^2\right), \tag{9}$$

where we have introduced $A_{\text{sat}}^2 := I_p/C_N \omega_0$.

Next, by incorporating the pump amplitude into a new
variable,

$$\Lambda_i := \frac{C_N \omega_0 A_{pi}}{2} - \frac{1}{2R_s} - \frac{N_i}{4R}, \tag{10}$$

we observe that the dynamics for A_{si} and A_{pi} can be re-
expressed as follows:

$$\frac{dA_{si}}{dt} = -\frac{1}{2C_{0s}}\frac{\partial \mathcal{L}}{\partial A_{si}}, \tag{11}$$

$$\frac{d\Lambda_i}{dt} = \frac{C_N^2 \omega_0^2}{4\left(C_0 + C_p\right)} \left(A_{\text{sat}}^2 - A_{si}^2\right) = \frac{C_N^2 \omega_0^2}{4\left(C_0 + C_p\right)} \frac{\partial \mathcal{L}}{\partial \Lambda_i},$$

(12)

where $\mathcal{L}$ is defined as

$$\mathcal{L}\left(\mathbf{A}_s, \Lambda\right) = -\frac{1}{4R} \sum_{i,j} J_{ij} A_{si} A_{sj} + \sum_i \Lambda_i \left(A_{\text{sat}}^2 - A_{si}^2\right),$$

(13)

with $\mathbf{A}_s$ and Λ being vectors whose components are A_{si} and Λ_i, respectively. We call $\mathcal{L}$ the "Lagrange function" of the problem. The above equations show that the signal and pump amplitudes respectively perform simultaneous gradient descent and ascent on the same function $\mathcal{L}$.

Notably, the first term in Eq. (13) has the form of the Ising interaction energy in Eq. (1), except that the amplitudes are not strictly binary. According to Eq. (11), the amplitudes of the oscillators evolve to minimize this Ising-like function. However, this does not fully describe the dynamics, due to the presence of the second term in the Lagrange function. We will show that these equations of motion are actually an exact implementation of the primal-dual method of Lagrange multipliers. In the next section, we provide a brief overview of the primal-dual method of Lagrange multipliers, and in Sec. IV, we make the isomorphism between the circuit and the primal-dual Lagrange-multiplier method more explicit. Before we move on, we mention as an aside that the circuit with a single global pump in Fig. 1(c) has a similar corresponding Lagrange function $\mathcal{L}'$ but this function has only one "pump" variable Λ, unlike Eq. (13), which has N "pump" variables Λ_i, $i \in \{1, \ldots, N\}$.

III. OVERVIEW OF LAGRANGE MULTIPLIERS

Here we provide a brief overview of the method of Lagrange multipliers, a very-well-known procedure for solving constrained optimization problems. This discussion is primarily drawn from the textbooks by Bertsekas [46] and Boyd and Vandenberghe [47].

Let $f(\mathbf{x})$ be an objective function of N variables, and let the point $\mathbf{x}^*$ locally minimize $f(\mathbf{x})$ among all $\mathbf{x}$ that simultaneously satisfy the constraints $g_1(\mathbf{x}) = 0, \ldots, g_p(\mathbf{x}) = 0$. Then, the theory of Lagrange multipliers tells us that one can find numbers $\lambda_1^*, \ldots, \lambda_p^*$, referred to as "Lagrange multipliers," such that

$$\nabla f(\mathbf{x}^*) = -\sum_{i=1}^p \lambda_i^* \nabla g_i(\mathbf{x}^*).$$

(14)

Letting λ be the vector whose components are λ_i, and defining the Lagrange function $L(\mathbf{x}, \lambda)$ as

$$L(\mathbf{x}, \lambda) := f(\mathbf{x}) + \sum_{i=1}^p \lambda_i g_i(\mathbf{x}),$$

(15)

we see that the constrained local minimum $\mathbf{x}^*$ and its associated multipliers λ^* satisfy

$$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda^*) = 0, \quad \nabla_\lambda L(\mathbf{x}^*, \lambda^*) = 0.$$

(16)

If a candidate point $\left(\mathbf{x}', \lambda'\right)$ satisfies these conditions, then $\mathbf{x}'$ is a stationary point of $f(\mathbf{x})$ subject to the constraints. Equation (16) converts the search for constrained stationary points of $f(\mathbf{x})$ into one for unconstrained stationary points of $L(\mathbf{x}, \lambda)$.

Certain "well-structured" problems (e.g., convex problems) satisfy "strong duality":

$$\min_{\mathbf{x}} \left(\max_\lambda L(\mathbf{x}, \lambda) \right) = \max_\lambda \left(\min_{\mathbf{x}} L(\mathbf{x}, \lambda) \right).$$

(17)

The point $(\mathbf{x}^*, \lambda^*)$ at which the equality holds is the global constrained optimum of the problem (see Appendix C for a saddle-point interpretation). The *primal-dual method*, also known as the "method of multipliers" [46], finds this optimum by iteratively solving the nested min-max optimization problem on the right-hand side of Eq. (17). Starting from an initial point $(\mathbf{x}^{(0)}, \lambda^{(0)})$, the algorithm first holds λ fixed and minimizes L using several gradient-descent steps in $\mathbf{x}$. Then $\mathbf{x}$ is held fixed and one step of gradient ascent on L is performed in the λ directions. This alternating procedure of fast minimization and slow maximization is repeated until convergence is achieved. In the limit of zero step size, the iterative algorithm is represented by a pair of differential equations:

$$\frac{d\mathbf{x}}{dt} = -\kappa \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda),$$

(18)

$$\frac{d\lambda}{dt} = \kappa' \nabla_\lambda L(\mathbf{x}, \lambda)$$

(19)

for appropriate step sizes κ and κ'. Equation (18) performs gradient descent on the Lagrange function to optimize $\mathbf{x}$, while Eq. (19) performs gradient ascent on L to optimize λ. This procedure is also called the "primal-dual algorithm" [48], where the descent in $\mathbf{x}$ is the *primal step* and the ascent in λ is the *dual step*. Strong duality guarantees the convergence of the algorithm to the constrained global optimum of f. A fast maximization over λ in conjunction with a slow minimization over $\mathbf{x}$ is also a valid algorithm since it represents the left-hand side of Eq. (17).

Most difficult problems are highly nonconvex and satisfy only weak duality:

$$\min_{\mathbf{x}} \left(\max_{\lambda} L(\mathbf{x}, \lambda) \right) \geq \max_{\lambda} \left(\min_{\mathbf{x}} L(\mathbf{x}, \lambda) \right). \qquad (20)$$

In this case, performing a fast maximization over λ in conjunction with a slow minimization over $\mathbf{x}$ is more accurate since the left-hand side of Eq. (20) is the required constrained minimum. Alternatively, one could use the more powerful augmented Lagrangian method of multipliers [46] when strong duality is not satisfied. The augmented Lagrange function, $L_\alpha(\mathbf{x}, \lambda)$, is defined as

$$L_\alpha(\mathbf{x}, \lambda) := L(\mathbf{x}, \lambda) + \frac{\alpha}{2} \left(\sum_{i=1}^{p} (g_i(\mathbf{x}))^2 \right) \qquad (21)$$

for a positive parameter α. Bertsekas [46] shows that if $L_\alpha(\mathbf{x}, \lambda)$ is used in place of $L(\mathbf{x}, \lambda)$ in Eqs. (18) and (19) and the system is initialized close to a local optimum $(\mathbf{x}^*, \lambda^*)$ of a weakly dual problem, the equations will converge to $(\mathbf{x}^*, \lambda^*)$.

IV. SIGNAL DYNAMICS PERFORMS PRIMAL STEP, PUMP DYNAMICS PERFORMS DUAL STEP

We now apply the method of Lagrange multipliers to the Ising optimization problem, whose merit function $f(\mathbf{x})$ is given in Eq. (1), with the constraint that each of the N spins is binary: $x_i = +1$ or $x_i = -1$. This binary constraint can be written as $g_i(\mathbf{x}) = 1 - x_i^2 = 0$ for all i from 1 to N. The Lagrange function for the Ising problem is then given by

$$L(\mathbf{x}, \lambda) = - \sum_{i=1}^{N} \sum_{j=1}^{N} J_{ij} x_i x_j + \sum_{i=1}^{N} \lambda_i \left(1 - x_i^2 \right), \qquad (22)$$

where λ_i is the Lagrange multiplier associated with the constraint on the ith spin. Substituting Eq. (22) into Eqs. (18) and (19), we derive the update equations for the primal-dual method of multipliers:

$$\frac{dx_i}{dt} = -\kappa \frac{\partial L}{\partial x_i} = -2\kappa \left(-\sum_{j=1}^{N} J_{ij} x_j - \lambda_i x_i \right), \qquad (23)$$

$$\frac{d\lambda_i}{dt} = \kappa' \frac{\partial L}{\partial \lambda_i} = \kappa' \left(1 - x_i^2 \right). \qquad (24)$$

Notably, the Ising problem's Lagrange function in Eq. (22) is isomorphic to the oscillator-network Lagrange function $\mathcal{L}$ [Eq. (13)] we defined in Sec. II. Moreover, the equations of motion of the method of multipliers, Eqs. (23) and (24), are in perfect correspondence with the oscillator network's equations of motion, Eqs. (11) and (12). More precisely,

TABLE I. Mapping of variables in the method of Lagrange multipliers to the coupled-oscillator network.

Problem variable	Physical variable
Spin variable x_i	$(1/A_{\text{sat}}) \times A_{si}$
Lagrange multiplier λ_i	$4R \times \Lambda_i$
Coupling matrix J_{ij}	J_{ij}
Lagrange function L	$\left(4R/A_{\text{sat}}^2 \right) \times \mathcal{L}$
Step size κ	$1/8RC_{0s}$
Step size κ'	$C_N^2 \omega_0^2 R A_{\text{sat}}^2 / C_{0p}$

the signal equation (23) exactly corresponds to the primal equation (11), while the pump equation (24) exactly corresponds to the dual equation (12). One needs only to make the identifications given in Table I to complete the correspondence. Besides A_{si}, Λ_i, and $\mathcal{L}$, all other physical parameters in Table I are fixed constants. Therefore, the coupled-oscillator circuit implements the two differential equations that describe the primal-dual method of multipliers.

The signal voltages A_{si} of the oscillators play the role of the Ising variables x_i, while the Λ_i variables play the role of the Lagrange multipliers. The Λ_i variables correspond physically to the gain supplied to each oscillator from the pump. In Eq. (10) for Λ_i, the term $\frac{1}{2} C_N \omega_0 A_{pi}$ is a negative conductance that corresponds to parametric gain. Since the pump voltage A_{pi} is the only time-varying component of the gain conductance, the time evolution of the Lagrange multipliers λ_i is fully contained in the dynamics of the pump oscillator voltages. Therefore, in summary, the signal oscillator voltages A_{si} play the role of the Ising variables x_i, while the pump oscillator voltages A_{pi} play the role of the corresponding Lagrange multipliers λ_i. This fact also highlights the importance of having a separate pump oscillator for every signal oscillator—the circuit with a single global pump in Fig. 1(c) implements only one Lagrange multiplier and fails to mimic the correct Ising Lagrange function, Eq. (22), which has N Lagrange multipliers.

Pump depletion performs the role of Lagrange-multiplier feedback to constrain the signal voltages. When the system reaches a steady state, all of the signal voltages satisfy the binarization constraint such that $x_i = \pm 1$. This can also be seen in Eq. (9): in the steady state ($dA_{pi}/dt = 0$), the amplitude of every oscillator is the same and equals A_{sat}. Therefore, pump depletion, which is equivalent to the dual step in the Lagrange method, ensures amplitude homogeneity of all the signal voltages in the steady state. This insight supersedes our claim in an earlier publication [38] that a separate feedback circuit would be necessary to implement the λ feedback. The primal-dual Lagrange algorithm is entirely self-contained in the dynamics of parametric oscillators.

A comment about pump loss is in order. In the mathematical primal-dual method, the Lagrange multipliers λ_i

change only in response to the values of the primal variables x_i [i.e., $d\lambda_i/dt \propto (1 - x_i^2)$]; the dual evolution does not involve the multipliers λ_i themselves. Intuitively, the Lagrange multipliers evolve in a manner that applies a constraint exclusively on the primal variables. In the physical system, on the other hand, since λ_i is encoded into the pump amplitudes A_{pi}, a finite pump loss introduces a loss term that depends on A_{pi} into the equation for dA_{pi}/dt, thus causing a deviation from the primal-dual method. It was for this reason that we set $R_p = \infty$ in Sec. II B just before Eq. (7). Table II shows that Q of 5000 seems to be sufficient for faithful performance, while lower values of Q lead to performance degradation.

A. Implementing the augmented Lagrangian method

Since the Ising problem does not satisfy strong duality, the augmented Lagrange function $L_\alpha(\mathbf{x}, \lambda)$ provides a theoretically higher-quality solution. For the Ising problem, this is given by

$$L_\alpha(\mathbf{x}, \lambda) = L(\mathbf{x}, \lambda) + \frac{\alpha}{2} \sum_{i=1}^{N} \left(1 - x_i^2\right)^2, \quad (25)$$

where $L(\mathbf{x}, \lambda)$ is from Eq. (22).

The augmented Lagrange equations of motion are as follows:

$$\frac{dx_i}{dt} = -2\kappa \left(-\sum_{j=1}^{N} J_{ij} x_j - \lambda_i x_i - \alpha x_i + \alpha x_i^3 \right), \quad (26)$$

$$\frac{d\lambda_i}{dt} = \kappa' \left(1 - x_i^2\right). \quad (27)$$

The equations are essentially the same as before except for an additional cubic nonlinear term that appears in Eq. (26). This nonlinear term, in addition to offering the theoretical optimization advantages discussed in Ref. [46], also ensures that the signal voltages remain closer to the saturation amplitude than in the plain Lagrange method. To implement this term physically, the signal resistor R_s in Fig. 2(b) is replaced with a nonlinear resistor with characteristic $I = G_0 V + G_N V^3$. In an electrical circuit, a simple practical implementation is a pair of parallel p-n-junction diodes that conduct in opposite directions.

Since the additional resistor is in the signal part of the circuit, the pump equations remain unchanged. The equation of motion for the signal circuit, derived in Appendix D, is as follows:

$$\dot{A}_{si} = \left[-\frac{N_i}{4RC_{0s}} A_{si} + \frac{1}{4RC_{0s}} \sum_{j:j \neq i} J_{ij} A_{sj} \right] + \frac{C_N \omega_0 A_{pi}}{2C_{0s}} A_{si}$$

$$- \frac{G_0}{2C_{0s}} A_{si} - \frac{3G_N}{8C_{0s}} A_{si}^3. \quad (28)$$

This equation can be cast in the form of Eq. (26) by rewriting it as

$$\dot{A}_{si} = -\frac{1}{4RC_{0s}} \left[-\sum_{j=1}^{N} J_{ij} A_{sj} - 4R \right.$$

$$\times \left(\Lambda_i - \frac{G_0}{2} - \frac{3G_N A_{\text{sat}}^2}{8} \right) A_{si}$$

$$\left. - \frac{3G_N R A_{\text{sat}}^2}{2} A_{si} + \frac{3G_N R}{2} A_{si}^3 \right] \quad (29)$$

and making the identification

$$x_i := \frac{A_{si}}{A_{\text{sat}}}, \quad (30)$$

$$\lambda_i := 4R \left(\Lambda_i - \frac{G_0}{2} - \frac{3G_N A_{\text{sat}}^2}{8} \right), \quad (31)$$

$$\alpha := \frac{3G_N R A_{\text{sat}}^2}{2}. \quad (32)$$

As before, the Lagrange multiplier corresponds to the gains supplied by the pump oscillators, but with an additional fixed offset. Since the nonlinear resistor is in the signal part of the circuit, the pump continues to evolve according to Eq. (12).

V. NUMERICAL RESULTS

In this section, we present the results of the numerical simulation of Eqs. (28) and (12) for quadratic binary optimization problems (which are readily converted to Ising instances through the simple procedures in Appendix E) of sizes 50, 100, 250, and 500 from the BiqMac collection, and max-cut problems of sizes 800 and 2000 from the Gset collection.

The natural frequencies of the signal and pump oscillators were set to $\omega_0 = 1$ GHz and $2\omega_0 = 2$ GHz, respectively. The nonlinear part of the capacitor that connects the signal and pump oscillators, C_N, was chosen so that the modulation on the capacitance is 10% of the signal capacitance C_s at an applied voltage of 1 V. The linear part of the connecting capacitance, C_0, was assumed to be zero because any nonzero C_0 can be absorbed into C_s and the pump capacitance C_p (made clear in the derivations in Appendix B). The voltage saturation amplitude A_{sat} of the signal oscillations was set to 10 mV. Binary weights $J = \pm 1$ were implemented according to the scheme in Fig. 2(c) by our connecting pairs of signal oscillators with resistors of a common value R in the appropriate configuration. Values of J other than ± 1 were implemented with the use of a geometric series of resistances centered at R and the binary expansion of J (see Appendix B 2). To ensure that the isomorphism with Lagrange multipliers holds, the

TABLE II.   Performance of our approach on selected Gset 800-variable max-cut problems compared with the performance of other algorithms. The oscillator Lagrange-multiplier approaches vastly outperform the Goemans-Williamson method but do not quite match up to the method of Leleu *et al.* Performance improves as the pump quality factor increases. For the last five columns, the best solution quality and the median solution quality are reported for ten independent runs, while the results for Leleu *et al.* [20] are for 20 runs. UB, upper bound.

| Problem | Goemans-Williamson method | Metric | Leleu *et al.* [20] | Oscillators, plain Lagrange method | Oscillators, augmented Lagrange method | Oscillators, augmented Lagrange method with pump loss and thermal noise | | |
						$Q_p = 100$	$Q_p = 500$	$Q_p = 5000$
1	11 272	Best	11 624	11 580	11 613	9946	11 460	11 613
	UB 12 838	Median	11 624	11 552	11 558	9931	11 416	11 570
2	11 277	Best	11 620	11 575	11 596	10 001	11 480	11 597
	UB 12 844	Median	11 620	11 554	11 572	9936	11 416	11 577
6	1813	Best	2178	2143	2173	448	2019	2168
	UB 3387	Median	2178	2124	2144	432	1967	2134
7	1652	Best	2006	1975	1973	283	1819	1984
	UB 3224	Median	2006	1950	1955	265	1794	1966

pump is henceforth assumed to have no internal dissipative loss, unless noted otherwise. The effect of pump resistance and noise on the performance is discussed later in this section and in Table II. Further details on how parameters were chosen and their effect on the solver's performance are provided in Appendix F.

A. Dynamics of the solver

The slowly-varying-amplitude equations (28) and (12) were simulated with use of MATLAB's built-in ode45 ordinary-differential-equation (ODE) solver for a total time of 50 µs with all the signal capacitor voltages starting at the noise level $\sqrt{kT/C_s} \approx 5$ µV and the initial pump voltages being set such that there is gain right from $t = 0$.

Appendix F provides further details of the initial conditions and the simulation setup. The signal and pump oscillator voltage amplitudes for the first 800-vertex problem in Gset are shown in Fig. 3. The oscillators corresponding to spins 1, 2, and 7 are plotted to depict the diversity of behaviors observed in the system: spin 2 starts out near the noise level but immediately settles down to a steady state of −10 mV (logical −1), spin 1 flips from logical +1 to logical −1 after an initial period of evolution, and spin 7 undergoes rapid repeated flipping between −1 and +1 and has relatively large fluctuations in its oscillation amplitude.

The time evolution of the pump voltages is shown in the bottom panel in Fig. 3. The pump voltage indicates how much parametric gain is being supplied to the corresponding signal oscillator to maintain a steady-state amplitude of ±10 mV. As explained previously, the pump voltage dynamics directly tracks the time evolution of the Lagrange multipliers. Spin 7, which has large fluctuations in the signal voltage and thus frequent deviations from the binary constraint, has correspondingly large fluctuations in its pump voltage.

An important feature of the time evolution of the signal amplitudes in the top panel in Fig. 3 that needs to be addressed is the rapid oscillations that they exhibit
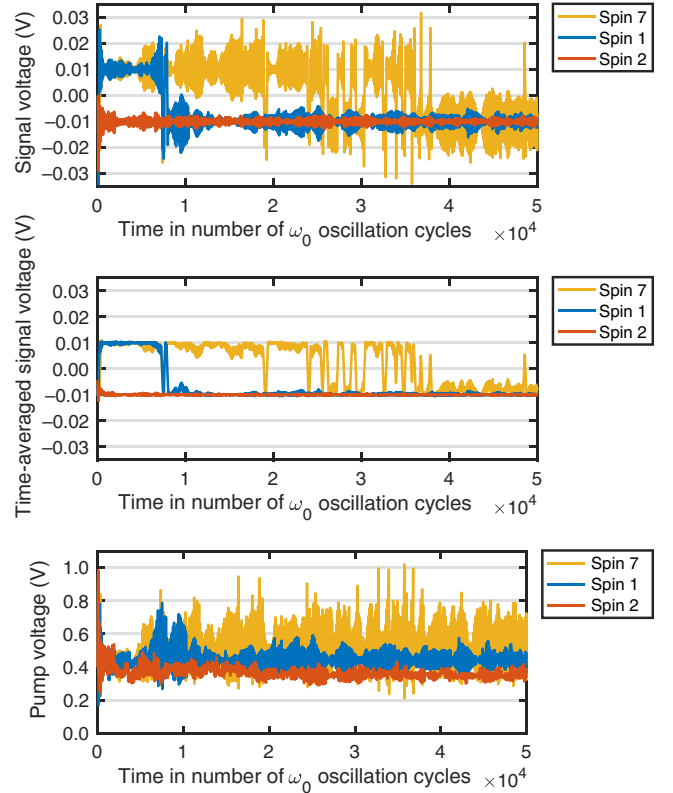

FIG. 3.   Temporal dynamics of the signal (top), window time average of the temporal dynamics that shows clamping to the saturation amplitude ($A_{sat} = \pm 10$ mV) (middle), and pump (bottom) oscillator voltage amplitudes for spins 1, 2, and 7 in an 800-variable Gset problem. The signal voltages start from noise before eventually settling to ±10 mV.

about the saturation amplitudes of ± 0.01 V. How is this behavior consistent with our claim earlier that our system possesses "amplitude homogeneity"? Close inspection of the top panel in Fig. 3 reveals that all the spins actually do obey amplitude homogeneity in the sense that the means of the voltage oscillations V_{si} are always centered at the saturation amplitudes ± 0.01 V for all spins i. This is seen clearly in the middle panel in Fig. 3, which depicts a running time average of the depiction in the top panel in Fig. 3. The key point is that there is no amplitude heterogeneity across spins, which was the problem in earlier methods where different spins were settling to different saturation amplitudes. On the basis of our investigations in Appendix F 1 b, the reason for the rapidly oscillating overshoots and undershoots about the saturation amplitude is the low value of C_p, the pump capacitance. We know that the pump dynamics exert a driving force on the signal voltages toward homogeneity (i.e., toward the saturation amplitude A_{sat}), thus imposing the constraints. Low C_p leads to fast pump dynamics, and consequently, aggressive movement toward the saturation amplitude followed by immediate overshoots/undershoots because of insufficient pump damping.

At each point in time, the collection of signal voltages A_{si} can be converted to a binary solution vector by taking the sign of each element. This allows the computation of an instantaneous max-cut value, shown in Fig. 4 for two problems: Gset problem 1, with 800 variables, and Gset problem 22, with 2000 variables. Most of the progress toward the best-known solution is made at early times, with a slowdown in improvements as time progresses. The best instantaneous objective-function value within the 50-μs simulation window is declared as the solution of the run.

B. Quality of solution

To understand how performance scales with size, we used BiqMac benchmark problems of size 50, 100, 250, and 500, and Gset benchmark problems of size 800 and 2000. Our problem set consisted of ten problems of each size for a total of 60 problems. The solver was run 10 times with random independent initial conditions on each problem and the best and median solutions obtained over the ten runs were recorded for each problem. The results for Gset problems 1 and 2 (800 spins, 0,1 weights) and problems 6 and 7 (800 spins, $-1, 0, 1$ weights) are presented in Table II. A more-comprehensive list is provided in Appendix F 4.

We use the performance of the Goemans-Williamson algorithm as a baseline for comparison, as well as to provide a theoretical upper bound on the quality of the max-cut solution. The best-known solutions to these specific max-cut problem instances are from Leleu *et al.* [20]. For our coupled-oscillator approach, we include the quality of
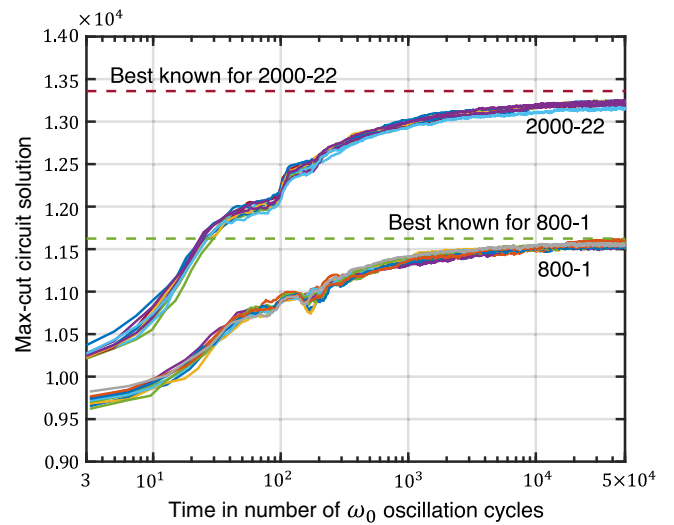

FIG. 4. Temporal evolution of the max-cut objective function. The different colors correspond to ten independent runs of the circuit starting from random noise for each of two problems: Gset problem 1, with 800 vertices (800-1), and problem 22, with 2000 vertices (2000-22).

the solution found without and with the nonlinear resistor, i.e., for the plain Lagrange multipliers and the augmented Lagrange methods, respectively. Finally, we include the results of the coupled-oscillator network under less-ideal conditions: the pump circuit is made lossy (parameterized by the quality factor of the pump oscillator, $Q_p = R_p \sqrt{C_p/L_p}$), and Johnson thermal noise is incorporated into both the signal circuit and the pump circuit. The noise model and the associated stochastic-differential-equation (SDE) approach are described in Appendix B 3.

We note several key findings from Table II. First, the coupled-parametric-oscillator network far outperforms the basic Goemans-Williamson algorithm. Second, the physical system that implements the augmented Lagrange method generally performs better than the plain Lagrange method, although the difference between the two methods is not always significant.

Introducing loss in the pump circuit leads to a reduction in performance. This is not surprising because the addition of pump loss breaks the exact correspondence with Lagrange multipliers as pointed out in Sec. IV. The performance improves as the pump quality factor increases, with the results for $Q_p > 5000$, even with thermal noise included in both the signal circuit and the pump circuit, being similar to the results in the lossless, noiseless case. Piezoelectric acoustic wave filters that are used in cellphone receivers have quality factors of this magnitude [49,50] and could be used in physical implementations of an electrical Ising solver.

Finally, the algorithm in Leleu *et al.* [20], for which a physical implementation was recently proposed [21], finds higher-quality solutions compared with the
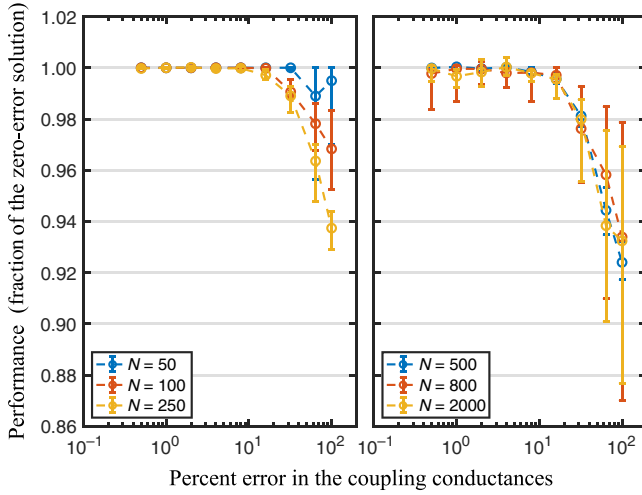
FIG. 5.   Effect of error in the coupling conductance values on solution quality. For each size N, four problems were picked and three random faulty circuits were run per problem; the results were normalized by the performance of error-free circuits. Medians and 25th-percentile and 75th-percentile values are reported.

Lagrange-multiplier solver. This is possibly due to nongradient chaotic dynamics that does not get stuck at fixed points or limit cycles. Lagrange multipliers, on the other hand, follow gradient-based dynamics in the form of alternating descent and ascent. The remainder of the oscillator results in this section (and in the appendixes) are for the augmented Lagrange method, assuming a lossless pump and no noise.
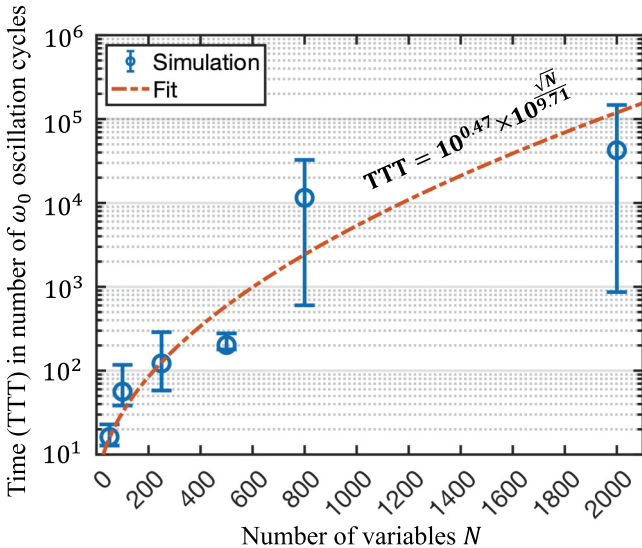

FIG. 6.   Time to target as a function of problem size. Ten problems were solved per problem size, and the median, 25th percentile, and 75th percentile for the time to target are shown. The first four points, for sizes 50, 100, 250, and 500, are for the BiqMac problems, while the last two points, for sizes 800 and 2000, are for Gset. The number of cycles scales as $10^{\sqrt{N}/9.71}$.

C. Robustness with regard to coupling-resistance imperfections

Sensitivity to component imperfections is one of the long-standing criticisms of analog computers. The present application has some built-in tolerance to these imperfections because the problem demands binary answers, even though the processing is done on analog signals in continuous time. In our network of coupled parametric oscillators, a main source of component imperfections is the (up to) N^2 resistors connecting the oscillators together. A connection weight J_{ij} is proportional to the conductance of the connecting resistors, given by $G_{ij} = |J_{ij}|/R$. Errors in these conductance values can cause the wrong problem to be solved by the hardware, in turn leading to nonoptimal solutions to the original problem.

We find that, in general, the Lagrange-multiplier oscillator network is insensitive to errors in the coupling resistors, as shown in Fig. 5. We assumed that the conductance G_{ij} had a Gaussian distribution with a mean given by $|J_{ij}|/R$ and a standard deviation that is a certain percentage of the mean. For each problem size, four problems were tackled; for each problem, results were obtained for three circuits sampled with randomized conductance errors for each error level. The results, plotted in Fig. 5, indicate that conductance errors as large as 10% are tolerated without significant reduction in the quality of the solution even for 2000-spin problems. This level of precision is well within the capability of modern programmable-resistive-memory devices [51]. The resilience of the system performance to pump resistance variation is discussed in Appendix F 2.

D. Time to target

Next, we extract the dependence of the time to target (TTT) on the problem size. A run of the solver on a given problem is considered *successful* if the instantaneous objective-function value breaches 97% of the best-known value for that problem at some point during the 50-μs duration of the run. We define the TTT for a successful run as the first time the 97% mark is crossed. For an unsuccessful run, the TTT is the full 50 μs. The TTT for the problem is then equal to the sum of the TTT of all the runs divided by the number of successful runs. This metric measures the average time spent between two successes. Figure 6 shows how the TTT depends on the number of variables N in our problem set. Although these problems are drawn from two different benchmark sets, the TTT, in number of ω_0 oscillation cycles, scales as $10^{0.47} \times 10^{\sqrt{N}/9.71} = 10^{0.47} \times 1.27^{\sqrt{N}}$, which is $\mathcal{O}(2^{\sqrt{N}})$, corroborating previous work on solvers of this type that noted similar scaling [15,52].

E. Power consumption

Finally, one can estimate the power consumption of our circuits by adding up the power dissipated by all the
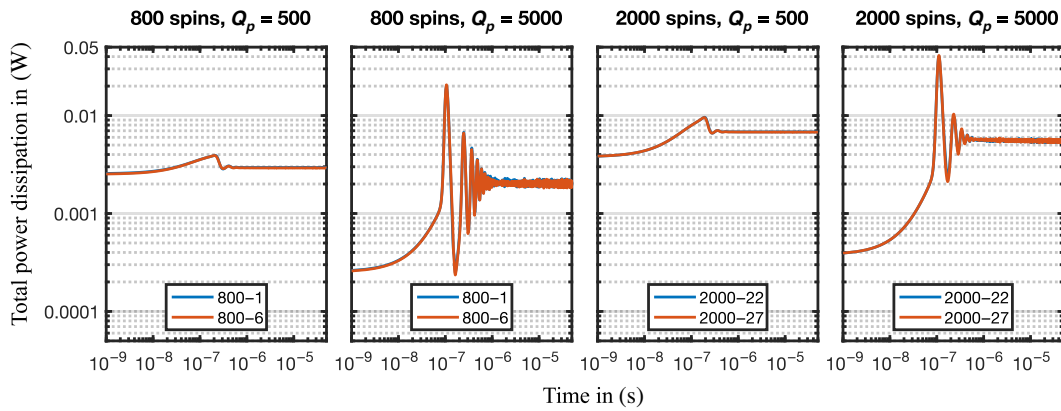
FIG. 7. Total power dissipation in the circuit as a function of time for different problems, obtained from stochastic-differential-equation simulations of the circuit with the thermal noise included. The two panels on the left depict power dissipation for 800-spin Gset problem 1 (800-1) and problem 6 (800-6) for two different pump Q values, whereas the two panels on the right depict the corresponding results for 2000-spin Gset problem 22 (2000-22) and problem 27 (2000-27). The blue and orange curves are very close and overlap in all the panels.

resistors in the system—the coupling resistors between spins, and the signal and pump resistors inside each spin. Figure 7 depicts the evolution of the total power dissipation of the circuit for two 800-spin and 2000-spin problems from Gset, for two values of the pump oscillator Q per problem. These plots were obtained from stochastic-differential-equation simulations in MATLAB of the circuit equations with thermal noise included (see Appendix B 3 for the details of the noise model). The key observations from Fig. 7 are as follows: (1) The total power dissipation is at the level of *milliwatts*, which is several orders of magnitude more efficient than standard CPUs and graphics processing units, which consume hundreds of watts [53]. This fact, combined with the run time of 50 μs, yields an energy consumption of tens of nanojoules per optimization, making this approach to combinatorial optimization highly attractive for future research and experimentation. A back-of-the-envelope calculation supporting these power-dissipation numbers is provided in Appendix F 3. (2) The power dissipation fluctuates significantly early on in the evolution when the system moves rapidly toward higher-quality objective-function values (see Fig. 4 for comparison), followed by a settling down that corresponds to the later, slow optimization phase (Fig. 4).

VI. CONCLUSION

In this paper, we showed that coupled-parametric-oscillator Ising solvers exactly implement the primal-dual Lagrange-multiplier optimization algorithm. The signal oscillator voltages represent the binary problem variables x_i, while the pump oscillator voltages represent the corresponding Lagrange multipliers λ_i. Moreover, the equations of motion of the signal oscillators match the primal (descent) equations, while the evolution equations of

the pump oscillators match the dual (ascent) equations of the Lagrange-multiplier method. This complete equivalence (1) explains why coupled-oscillator solvers find high-quality solutions (2) and provides precise mathematical meaning to each physical component in the circuit. Moreover, using the above-mentioned insights, we showed that one can readily implement a more-sophisticated algorithm—the augmented Lagrange-multiplier method.

Next, we showed that numerical simulations of the slowly-varying-amplitude circuit equations on 800-spin and 2000-spin problems yielded results that were significantly better than the results obtained with the Goemans-Williamson algorithm, and that were competitive with the best-known Ising solvers [20]. The time to target was found to scale as $\mathcal{O}(2^{\sqrt{N}})$, where N is the problem size—a result that is consistent with other work in the literature [15,52]. Next, our simulations (with thermal noise taken into account) reveal that the circuit consumes extremely low amounts of power (on the order of milliwatts) and energy per optimization (approximately 100 nJ) even for problems of size 2000, making the system highly attractive for further research and experimentation. Finally, we also showed that the quality of the solutions obtained by the oscillator solvers was robust with regard to significant errors in the circuit components used to program the J_{ij}—an intriguing result that points to the possibility that the traditional shortcomings of general-purpose analog computation might not be a stumbling block for this application. We note that all our theoretical and numerical results are true for coupled-parametric-oscillator systems of all types in general—the electrical circuit in this paper was chosen only as a model system to illustrate the equivalence and properties of such solvers more concretely.

These encouraging results suggest a promising research direction where such circuit solvers are designed to

implement more-sophisticated optimization algorithms for a multitude of important problems in operations research and machine learning, and are used as analog coprocessors or accelerators alongside standard digital chips. We hope this work will instigate further bidirectional research into the design of novel physical systems that naturally perform optimization (physical optimizers) of various flavors and the construction of novel optimization algorithms from physical dynamics.

The BiqMac and Gset problems that were used to evaluate the solution techniques in this paper are publicly available at the resources listed in Refs. [40,41]. The MATLAB and LTspice codes that were developed to generate the results will be shared by the authors upon request.

ACKNOWLEDGMENTS

APPENDIX A: SINGLE PARAMETRIC *LC* OSCILLATOR—EQUATIONS OF MOTION

Before we start the derivation, we note that the current source in the signal circuit, I_s, is just a noise source in our system. Therefore, it can be dropped while we are considering the evolution of the system. We retain it in the current derivation simply to obtain general expressions but will drop it as soon as the discussion specializes to our situation.

The circuit equations for Fig. 8 are as follows:

$$C_s \dot{V}_s = I_3 - I_4, \quad C_p \dot{V}_p = I_2 - I_3, \tag{A1}$$

$$V_s = L_s \left(\dot{I}_4 - \dot{I}_5 \right), \quad V_p = L_p \left(\dot{I}_1 - \dot{I}_2 \right), \tag{A2}$$

$$V_s = R_s \left(I_s + I_5 \right), \quad V_p = R_p \left(I_p - I_1 \right), \tag{A3}$$

$$I_3 = C_0 \left(\dot{V}_p - \dot{V}_s \right) + 2C_N \left(V_p - V_s \right) \left(\dot{V}_p - \dot{V}_s \right). \tag{A4}$$

I_3 can be eliminated by substituting Eq. (A4) into Eq. (A1). Then, Eqs. (A1) and (A3) can be used to express I_1, I_2, I_4, and I_5 in terms of voltages and the current sources.
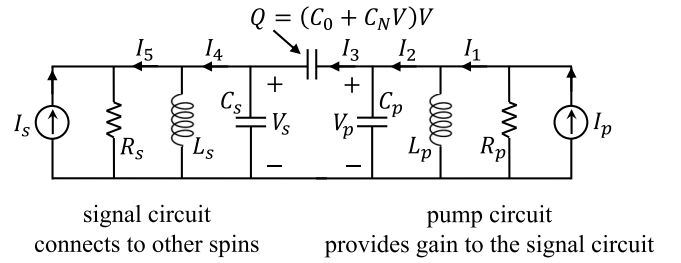

FIG. 8. The signal (left) and pump (right) *LC* oscillators are connected by a nonlinear capacitor.

Finally, plugging all these expressions into Eq. (A2) yields

$$V_s = L_s \left(C_0 \left(\ddot{V}_p - \ddot{V}_s \right) + 2C_N \left(\dot{V}_p - \dot{V}_s \right)^2 \right. $$
$$\left. + 2C_N \left(V_p - V_s \right) \left(\ddot{V}_p - \ddot{V}_s \right) - C_s \ddot{V}_s - \frac{\dot{V}_s}{R_s} + \dot{I}_s \right), \tag{A5}$$

$$V_p = L_p \left(-C_0 \left(\ddot{V}_p - \ddot{V}_s \right) - 2C_N \left(\dot{V}_p - \dot{V}_s \right)^2 \right. $$
$$\left. - 2C_N \left(V_p - V_s \right) \left(\ddot{V}_p - \ddot{V}_s \right) - C_p \ddot{V}_p - \frac{\dot{V}_p}{R_p} + \dot{I}_p \right). \tag{A6}$$

In Eq. (A5), we retain only terms that oscillate at ω_0 or contribute to oscillations at ω_0. Similarly, in Eq. (A6), we retain only terms that oscillate at $2\omega_0$ or contribute to oscillations at $2\omega_0$. These equations simplify to

$$\dot{I}_s = \frac{V_s}{L_s} + \left(C_0 + C_s \right) \ddot{V}_s + \frac{\dot{V}_s}{R_s} + 4C_N \dot{V}_p \dot{V}_s $$
$$+ 2C_N V_p \ddot{V}_s + 2C_N V_s \ddot{V}_p, \tag{A7}$$

$$\dot{I}_p = \frac{V_p}{L_p} + \left(C_0 + C_p \right) \ddot{V}_p + \frac{\dot{V}_p}{R_p} + 2C_N \dot{V}_s^2 + 2C_N V_s \ddot{V}_s. \tag{A8}$$

At this point, we make the redefinitions $C_s := C_0 + C_s$ and $C_p := C_0 + C_p$ for notational convenience. Next, we perform the slowly-varying-amplitude approximation by expressing all the currents and voltages involved as follows:

$$I_p = \frac{I_{pc}}{2} e^{2i\omega_0 t} + \text{c.c.}, \quad \dot{I}_p = 2i\omega_0 \frac{I_{pc}}{2} e^{2i\omega_0 t} + \text{c.c.}, \tag{A9}$$

$$I_s = \frac{I_{sc} - iI_{ss}}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.}, \quad \dot{I}_s = i\omega_0 \frac{I_{sc} - iI_{ss}}{2} e^{i(\omega_0 t + \phi_s)} $$
$$+ \text{c.c.}, \tag{A10}$$

$$V_s = \frac{A_s - iB_s}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.}, \quad \dot{V}_s = i\omega_0 \frac{A_s - iB_s}{2} $$
$$\times e^{i(\omega_0 t + \phi_s)} + \frac{\dot{A}_s - i\dot{B}_s}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.}, \tag{A11}$$

$$\ddot{V}_s = 2i\omega_0 \frac{\dot{A}_s - i\dot{B}_s}{2} e^{i(\omega_0 t + \phi_s)} - \omega_0^2 \frac{A_s - iB_s}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.},$$
$$\text{(A12)}$$

$$V_p = \frac{A_p}{2} e^{i(2\omega_0 t)} + \text{c.c.}, \quad \dot{V}_p = 2i\omega_0 \frac{A_p}{2} e^{i(2\omega_0 t)}$$
$$+ \frac{\dot{A}_p}{2} e^{i(2\omega_0 t)} + \text{c.c.}, \quad \text{(A13)}$$

$$\ddot{V}_p = 4i\omega_0 \frac{\dot{A}_p}{2} e^{i(2\omega_0 t)} - 4\omega_0^2 \frac{A_p}{2} e^{i(2\omega_0 t)} + \text{c.c.}, \quad \text{(A14)}$$

where A_s is the cosine component of V_s and B_s is its sine component. Plugging these expressions into Eq (A7), we get

$$i\omega_0 (I_{sc} - iI_{ss}) = \frac{A_s - iB_s}{L_s} + C_s \left(2i\omega_0 \left(\dot{A}_s - i\dot{B}_s\right)\right.$$
$$\left. - \omega_0^2 (A_s - iB_s)\right)$$
$$+ \frac{i\omega_0 (A_s - iB_s) + \dot{A}_s - i\dot{B}_s}{R_s}$$
$$+ C_N \left[2 \left(2i\omega_0 A_p + \dot{A}_p\right)(-i\omega_0 (A_s + iB_s)\right.$$
$$+ \dot{A}_s + i\dot{B}_s)$$
$$+ A_p \left(-2i\omega_0 \left(\dot{A}_s + i\dot{B}_s\right) - \omega_0^2 (A_s + iB_s)\right)$$
$$\left. + (A_s + iB_s) \left(4i\omega_0 \dot{A}_p - 4\omega_0^2 A_p\right)\right] e^{i(-2\phi_s)}.$$
$$\text{(A15)}$$

Equating the imaginary parts on both sides, recognizing that $1/\omega_0^2 = L_1 C_s$, rearranging terms, and setting $\phi_s = 3\pi/4$, we have

$$\dot{A}_s = \frac{I_{sc}}{2C_s} - \frac{A_s}{2R_s C_s} + \frac{C_N \omega_0 A_p}{2C_s} A_s + \frac{\dot{B}_s}{2R_s \omega_0 C_s} - \frac{C_N}{2\omega_0 C_s}$$
$$\times \left[2\dot{A}_s \dot{A}_p - 2\omega_0 \dot{A}_p B_s - 2\omega_0 \dot{B}_s A_p\right]. \quad \text{(A16)}$$

The first term on the right-hand side of Eq. (A16) is the injection from the current source, the second term is the internal resistive loss, the third term is the gain provided by the pump to the signal cosine component. The fourth term can be ignored because its magnitude is $\mathcal{O}$ (loss coefficient/ω_0), which is small by the slowly-varying-amplitude approximation. The first term in the brackets is small compared with $(C_N \omega_0 A_p/2C_s) A_s$ (again by the slowly-varying-amplitude approximation) and can be dropped. Finally, the second and third terms in the brackets can be dropped because they are of size $\mathcal{O}$ (gain coefficient/ω_0). The cosine amplitude dynamics is then

$$\dot{A}_s = \frac{I_{sc}}{2C_s} - \frac{A_s}{2R_s C_s} + \frac{C_N \omega_0 A_p}{2C_s} A_s. \quad \text{(A17)}$$
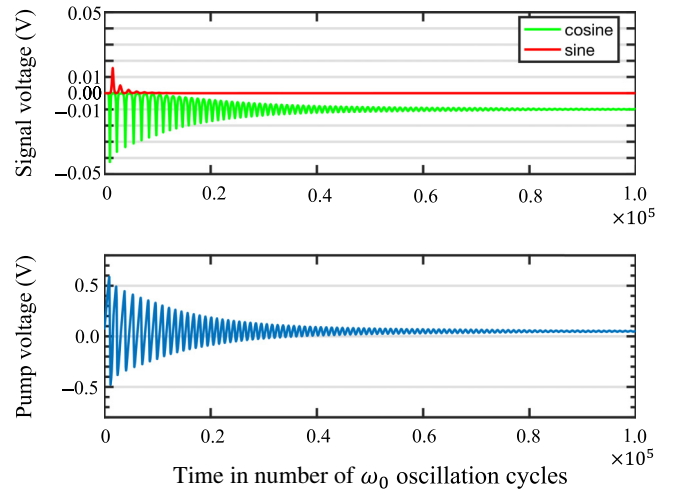

FIG. 9.   Simulation of the slowly-varying-amplitude equations for a single spin system—a signal oscillator connected to a pump oscillator. The sine component decays to zero early in the evolution.

Equating the real parts on both sides of Eq. (A15), we get for the amplitude of the sine component

$$\dot{B}_s = \frac{I_{ss}}{2C_s} - \frac{B_s}{2R_s C_s} - \frac{C_N \omega_0 A_p}{2C_s} B_s - \frac{\dot{A}_s}{2R_s \omega_0 C_s} + \frac{C_N}{2\omega_0 C_s}$$
$$\times \left[2\dot{B}_s \dot{A}_p + 2\omega_0 \dot{A}_p A_s + 2\omega_0 \dot{A}_s A_p\right]. \quad \text{(A18)}$$

The third term on the right-hand side is a parametric *loss* term and not a gain term, due to which the sine component never grows to the same order of magnitude as the cosine component—this will be verified shortly.

The equivalent of Eqs. (A16) and (A18) for the pump circuit is

$$\dot{A}_p = \frac{I_{pc}}{2C_p} - \frac{A_p}{2R_p C_p} + \frac{C_N}{4C_p \omega_0} \left(\dot{A}_s^2 - \dot{B}_s^2\right)$$
$$+ \frac{C_N}{4C_p \omega_0} \left(2\omega_0^2 \left(B_s^2 - A_s^2\right) + 4\omega_0 \left(A_s \dot{B}_s + \dot{A}_s B_s\right)\right).$$
$$\text{(A19)}$$

Simulating Eqs. (A16), (A18), and (A19) using MATLAB's ode15i implicit ODE solver leads to the plots in Fig. 9, confirming that the sine component B_s decays to zero very early and can be ignored. It is henceforth dropped in all our equations. Performing the slowly-varying-amplitude approximation on Eq. (A19) and dropping terms that contain B_s, we obtain the following final pump amplitude-evolution equation:

$$\dot{A}_p = \frac{I_{pc}}{2C_p} - \frac{A_p}{2R_p C_p} - \frac{C_N \omega_0 A_s^2}{2C_p}. \quad \text{(A20)}$$
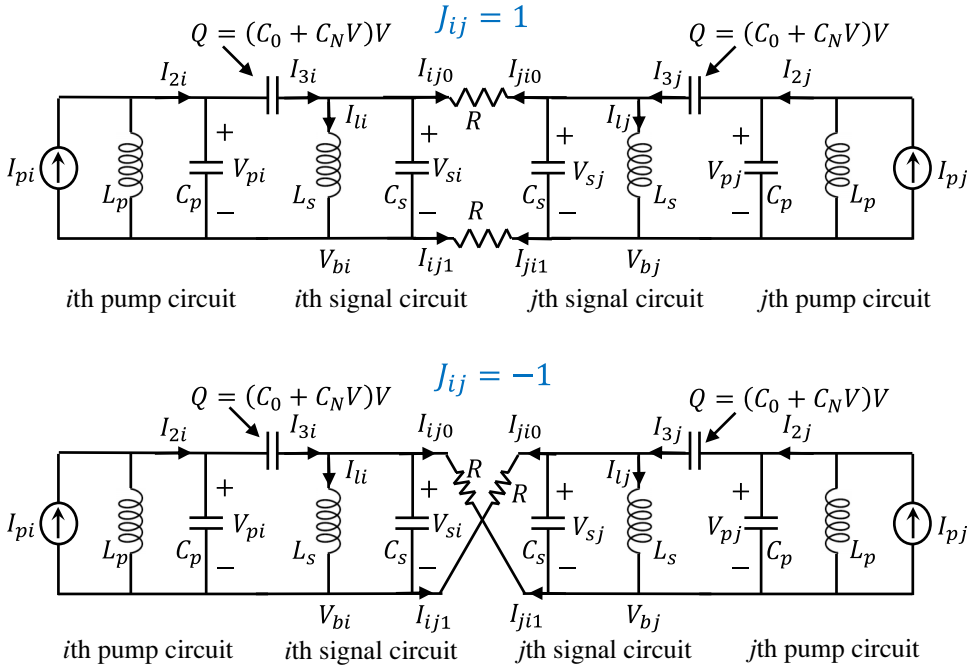
FIG. 10. The various currents and voltages when one focuses on the coupling between the ith parametric oscillator and the jth parametric oscillator. The top and bottom panels show the $J_{ij} = 1$ and $J_{ij} = -1$ cases, respectively.

APPENDIX B: COUPLED PARAMETRIC LC OSCILLATORS—EQUATIONS OF MOTION

In this appendix, we derive the equations of motion for a network of N coupled parametric oscillators with all-to-all coupling with J_{ij} taking values ± 1. The general case of sparser/non-± 1 coupling is dealt with later in this appendix.

Let the parametric oscillators be labeled from $i = 1$ to $i = N$. The notation we use is indicated in Fig. 10, where we focus on the coupling between the ith parametric oscillator and the jth parametric oscillator. One of the terminals of the capacitor in the oscillator labeled $i = 1$ (not shown in Fig. 10) is arbitrarily chosen as its "bottom" terminal, and its other terminal is labeled its "top terminal." For each oscillator that is connected to $i = 1$ through a $+1$ connection, the terminal in that oscillator that is directly connected to the bottom terminal of $i = 1$ is labeled its "bottom terminal." Similarly, for each oscillator that is connected to $i = 1$ through a -1 connection, the terminal in that oscillator that is directly connected to the bottom terminal of $i = 1$ is labeled its "top terminal." We continue this process until all terminals in the circuit are labeled. If two terminals are connected by a $+$ connection and one of them is the bottom terminal of its host oscillator, the other terminal is labeled the "bottom terminal" of its own host oscillator. If two terminals are connected by a $-$ connection and one of them is the bottom terminal of its host oscillator, the other terminal is labeled the "top terminal" of its own host oscillator. Through this process, we

can identify the bottom terminals of all the oscillators. The "bottom" labeling is shown in Fig. 10 for $J_{ij} = 1$ and $J_{ij} = -1$.

Let the potential at the "bottom" terminal of oscillator i be V_{bi}. The current that flows out from the bottom terminal of the ith oscillator into the resistor that connects it to the jth oscillator is i_{ij1}. Similarly, the current that flows out from the top terminal of the ith oscillator into the resistor that connects it to the jth oscillator is i_{ij0}. In the ith oscillator, the voltage difference between the top terminal and the bottom terminal of the capacitor is denoted by V_i, the current passing through the inductor from the top terminal to the bottom terminal is I_{li}, and the current passing through the capacitor from the top terminal to the bottom terminal is I_{ci}. All this notation is again indicated in Fig. 10.

If we represent the set $\{1, 2, \ldots, N\}$ by $[N]$, the circuit equations are as follows:

$$C_s \dot{V}_{si} = I_{3i} - I_{li} - \sum_j I_{ij0}, \quad C_p \dot{V}_{pi} = I_{2i} - I_{3i}$$

$$\text{for all } i \in [N], \tag{B1}$$

$$V_{si} = L_s \dot{I}_{li}, \quad V_{pi} = L_p \left(\dot{I}_{pi} - \dot{I}_{2i} \right) \text{ for all } i \in [N], \tag{B2}$$

$$I_{3i} = C_0 \left(\dot{V}_{pi} - \dot{V}_{si} \right) + 2C_N \left(V_{pi} - V_{si} \right) \left(\dot{V}_{pi} - \dot{V}_{si} \right)$$

$$\text{for all } i \in [N], \tag{B3}$$

$$\sum_j I_{ij0} + \sum_j I_{ij1} = 0 \text{ for all } i \in [N], \tag{B4}$$

$$V_{b1} = 0, \tag{B5}$$

$$I_{ijk} = k\left[\left(\frac{V_{bi} - V_{bj}}{R}\right)\left(\frac{1 + J_{ij}}{2}\right) + \left(\frac{V_{bi} - V_{bj} - V_{sj}}{R}\right)\left(\frac{1 - J_{ij}}{2}\right)\right]$$

$$+ (1 - k)\left[\left(\frac{V_{bi} + V_{si} - V_{bj} - V_{sj}}{R}\right)\left(\frac{1 + J_{ij}}{2}\right) + \left(\frac{V_{bi} + V_{si} - V_{bj}}{R}\right)\left(\frac{1 - J_{ij}}{2}\right)\right]$$

for all $i \in \{1, 2, \ldots, N\}$, for all $j \in \{1, 2, \ldots, N\}$, $j \neq i$, and for all $k \in \{0, 1\}$. $\tag{B6}$

Equations (B1)–(B4) and (B6) are the current law, voltage law, and device characteristics at different places in the circuit. Equation (B5) fixes the voltage reference by setting the potential of the bottom terminal of the first oscillator to zero.

Equation (B6) yields

$$\sum_j I_{ij0} = \sum_{j:j\neq i}\left[\frac{V_{bi} - V_{bj} + V_{si}}{R} - \frac{V_{sj}}{R}\left(\frac{1 + J_{ij}}{2}\right)\right]. \tag{B7}$$

Next, plugging Eq. (B6) into Eq. (B4), we get

$$\sum_{j:j\neq i} 2\left(\frac{V_{bi} - V_{bj}}{R}\right) = -\frac{(N-1)V_{si}}{R} + \sum_{j:j\neq i}\frac{V_{sj}}{R}. \tag{B8}$$

Finally, substituting Eq. (B8) into Eq. (B7), we get

$$2R\left(\sum_j I_{ij0}\right) = (N-1)V_{si} - \sum_{j:j\neq i} J_{ij}V_{sj}. \tag{B9}$$

We solve Eqs. (B1)–(B3) in the same way as before, retaining only the ω_0 and $2\omega_0$ terms in the signal and pump equations respectively, to obtain

$$-\frac{N-1}{2R}\dot{V}_{si} + \frac{1}{2R}\sum_{j:j\neq i} J_{ij}\dot{V}_{sj} = \frac{V_{si}}{L_s} + (C_0 + C_s)\ddot{V}_{si}$$

$$+ 4C_N\dot{V}_{pi}\dot{V}_{si} + 2C_N V_{pi}\ddot{V}_{si} + 2C_N V_{si}\ddot{V}_{pi}, \tag{B10}$$

$$\dot{I}_{pi} = \frac{V_{pi}}{L_p} + (C_0 + C_p)\ddot{V}_{pi} + 2C_N\dot{V}_{si}^2 + 2C_N V_{si}\ddot{V}_{si}. \tag{B11}$$

1. Slowly-varying-amplitude approximation

Making the substitution $C_{0s} := C_0 + C_s$ and $C_{0p} := C_0 + C_p$ and plugging into Eq. (B10) the slowly varying amplitudes from above, we get the following signal equations:

$$\dot{A}_{si} - \frac{C_N\omega_0 A_{pi}}{2C_{0s}}A_{si} + \frac{C_N}{2\omega_0 C_{0s}}\left[2\dot{A}_{si}\dot{A}_{pi} - 2\omega_0\dot{A}_{pi}B_{si} - 2\omega_0\dot{B}_{si}A_{pi}\right]$$

$$= \frac{1}{2C_{0s}}\left[-\frac{N-1}{2R}\left(A_{si} - \frac{\dot{B}_{si}}{\omega_0}\right) + \frac{1}{2R}\sum_{j:j\neq i} J_{ij}\left(A_{sj} - \frac{\dot{B}_{sj}}{\omega_0}\right)\right], \tag{B12}$$

$$\dot{B}_{si} + \frac{C_N\omega_0 A_{pi}}{2C_{0s}}B_{si} - \frac{C_N}{2\omega_0 C_{0s}}\left[2\dot{B}_{si}\dot{A}_{pi} + 2\omega_0\dot{A}_{pi}A_{si} + 2\omega_0\dot{A}_{si}A_{pi}\right]$$

$$= \frac{1}{2C_{0s}}\left[-\frac{N-1}{2R}\left(B_{si} + \frac{\dot{A}_{si}}{\omega_0}\right) + \frac{1}{2R}\sum_{j:j\neq i} J_{ij}\left(B_{sj} + \frac{\dot{A}_{sj}}{\omega_0}\right)\right]. \tag{B13}$$

As before, we drop Eq. (B13) and all terms that contain B_{si} in Eq. (B12), and we apply the slowly-varying-amplitude approximation to the remaining terms in Eq. (B12) to obtain

$$\dot{A}_{si} = \left[-\frac{N-1}{4RC_{0s}}A_{si} + \frac{1}{4RC_{0s}}\sum_{j:j\neq i} J_{ij}A_{sj}\right] + \frac{C_N\omega_0 A_{pi}}{2C_{0s}}A_{si}. \tag{B14}$$

Equation (B11) reduces to

$$\dot{A}_{pi} = \frac{I_{pci}}{2C_{0p}} - \frac{C_N\omega_0 A_{si}^2}{2C_{0p}}.$$

(B15)

2. Extension to cases where we do not have $J_{ij} = \pm 1$ or all-to-all connections

So far, we have considered only J matrices in which all the entries were chosen from $\{-1, 1\}$. In this section, we describe the modifications required to generalize the coupled-*LC*-oscillator circuit when the J values take on arbitrary real values expressed in binary form $\ldots b_2 b_1 b_0 \cdot b_{-1} b_{-2} \ldots$. We let the sign of J_{ij}, positive or negative, be represented by s_{ij}. That is, $J_{ij} = s_{ij} |J_{ij}|$. The circuit equations (B1) to (B5) carry over, while Eq. (B6) is modified to

$$I_{ijk} = k \left[\left(\frac{V_{bi} - V_{bj}}{R_{ij}} \right) \left(\frac{1 + s_{ij}}{2} \right) + \left(\frac{V_{bi} - V_{bj} - V_{sj}}{R_{ij}} \right) \left(\frac{1 - s_{ij}}{2} \right) \right]$$
$$+ (1 - k) \left[\left(\frac{V_{bi} + V_{si} - V_{bj} - V_{sj}}{R_{ij}} \right) \left(\frac{1 + s_{ij}}{2} \right) + \left(\frac{V_{bi} + V_{si} - V_{bj}}{R_{ij}} \right) \left(\frac{1 - s_{ij}}{2} \right) \right],$$

for all $i \in \{1, 2, \ldots, N\}$ for all $j \in \{1, 2, \ldots, N\}$, $j \neq i$, and for all $k \in \{0, 1\}$

(B16)

By our following the same procedure as before, Eq. (B8) is changed to

$$\sum_{j:j \neq i} 2 \left(\frac{V_{bi} - V_{bj}}{R_{ij}} \right) = -\sum_{j:j \neq i} \frac{V_{si}}{R_{ij}} + \sum_{j:j \neq i} \frac{V_{sj}}{R_{ij}}.$$

(B17)

To implement an arbitrary J_{ij}, we use a "common" coupling resistor R, and binary multiples of it, $R_m = 2^m R$. That is, R_{-1} is $R/2$ and R_2 is $4R$. If $|J_{ij}|$ is written in binary form up to 3-bit precision as

$$|J_{ij}| = b_1 2^1 + b_0 2^0 + b_{-1} 2^{-1},$$

(B18)

J_{ij} is implemented by our setting

$$\frac{1}{R_{ij}} = \frac{b_1}{R_{-1}} + \frac{b_0}{R_0} + \frac{b_{-1}}{R_1} = \frac{1}{R} |J_{ij}|.$$

(B19)

To see that this setting indeed does the job, we plug this expression for R_{ij} into Eq. (B17) and complete the calculation to see that Eq. (B14) is changed to

$$\dot{A}_{si} = \left[-\frac{\sum_{j:j \neq i} |J_{ij}|}{4RC_{0s}} A_{si} + \frac{1}{4RC_{0s}} \sum_{j:j \neq i} J_{ij} A_{sj} \right]$$
$$+ \frac{C_N\omega_0 A_{pi}}{2C_{0s}} A_{si}.$$

(B20)

The pump equation (B15) remains unchanged.

3. Including thermal noise in the coupling resistors

Thermal noise is incorporated into the circuit by our adding noise voltage sources $V_{ij0}^{(n)}$ and $V_{ij1}^{(n)}$ in series with the coupling resistors R_{ij0} and R_{ij1}, respectively. Further,

the thermal noise generated by the internal resistors of the signal and pump *LC* oscillators is modeled by our adding noise current sources $I_{si}^{(n)}$ and $I_{pi}^{(n)}$ in parallel to the two tanks, respectively. Then the counterparts of Eqs. (B10) and (B11) are as follows:

$$-\frac{\sum_{j:j \neq i} |J_{ij}|}{2R} \dot{V}_{si} + \frac{1}{2R} \sum_{j:j \neq i} J_{ij} \dot{V}_{sj} - \frac{1}{2R} \sum_{j:j \neq i} |J_{ij}| \dot{V}_{ij1}^{(n)}$$
$$+ \frac{1}{2R} \sum_{j:j \neq i} |J_{ij}| \dot{V}_{ij0}^{(n)} + \dot{I}_{si}^{(n)}$$
$$= \frac{V_{si}}{L_s} + \frac{\dot{V}_{si}}{R_s} + C_{0s} \ddot{V}_{si} + 4C_N \dot{V}_{pi} \dot{V}_{si}$$
$$+ 2C_N V_{pi} \ddot{V}_{si} + 2C_N V_{si} \ddot{V}_{pi},$$

(B21)

$$\dot{I}_{pi}^{(n)} + \dot{I}_{pi} = \frac{V_{pi}}{L_p} + \frac{\dot{V}_{pi}}{R_p} + C_{0p} \ddot{V}_{pi} + 2C_N \dot{V}_{si}^2 + 2C_N V_{si} \ddot{V}_{si}.$$

(B22)

The slowly-varying-amplitude approximation restricts the above equations to small frequency windows around ω_0 and $2\omega_0$, respectively, which means that only band-pass-filtered versions of the white-noise terms $I_{pi}^{(n)}$, $I_{si}^{(n)}$, $V_{ij0}^{(n)}$, and $V_{ij1}^{(n)}$ are retained in the slowly-varying-amplitude equations. If the impulse responses of band-pass filters centered at about ω_0 and $2\omega_0$ are $h_s(t)$ and $h_p(t)$, respectively, the convolution operator is represented by Λ, and the slowly varying cosine and sine noise amplitudes are represented by $A^{(n)}$ and $B^{(n)}$ with the appropriate subscripts (additional I in the superscript to represent currents), we have the

following expressions for the filtered noise terms:

$$h_s(t) * I_{si}^{(n)}(t) = \frac{A_{si}^{(n,I)} - iB_{si}^{(n,I)}}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.},$$

$$h_s(t) * v_{ij0}^{(n)} = \frac{A_{ij0}^{(n)} - iB_{ij0}^{(n)}}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.},$$

$$h_s(t) * v_{ij1}^{(n)} = \frac{A_{ij1}^{(n)} - iB_{ij1}^{(n)}}{2} e^{i(\omega_0 t + \phi_s)} + \text{c.c.},$$

$$h_p(t) * I_{pi}^{(n)} = \frac{A_{pi}^{(n,I)} - iB_{pi}^{(n,I)}}{2} e^{i(2\omega_0 t)} + \text{c.c.}$$

Using standard formulae and assuming that $A^{(n)}$ and $B^{(n)}$ are identically distributed but independent random processes, the two-time correlation functions of the slowly varying noise amplitudes are as follows:

$$\langle A_{si}^{(n,I)}(0) A_{si}^{(n,I)}(\tau) \rangle = \langle B_{si}^{(n,I)}(0) B_{si}^{(n,I)}(\tau) \rangle$$
$$= \frac{4kT}{R_s} \frac{h_s(\tau) * h_s(-\tau)}{\cos(\omega_0 \tau)}, \quad \text{(B23)}$$

$$\langle A_{ij0}^{(n)}(0) A_{ij0}^{(n)}(\tau) \rangle = \langle B_{ij0}^{(n)}(0) B_{ij0}^{(n)}(\tau) \rangle$$
$$= 4kTR_{ij} \frac{h_s(\tau) * h_s(-\tau)}{\cos(\omega_0 \tau)}, \quad \text{(B24)}$$

$$\langle A_{ij1}^{(n)}(0) A_{ij1}^{(n)}(\tau) \rangle = \langle B_{ij1}^{(n)}(0) B_{ij1}^{(n)}(\tau) \rangle$$
$$= 4kTR_{ij} \frac{h_s(\tau) * h_s(-\tau)}{\cos(\omega_0 \tau)}, \quad \text{(B25)}$$

$$\langle A_{pi}^{(n,I)}(0) A_{pi}^{(n,I)}(\tau) \rangle = \langle B_{pi}^{(n,I)}(0) B_{pi}^{(n,I)}(\tau) \rangle$$
$$= \frac{4kT}{R_p} \frac{h_p(\tau) * h_p(-\tau)}{\cos(2\omega_0 \tau)}. \quad \text{(B26)}$$

The slowly varying versions of Eqs. (B21) and (B22) are then

$$\dot{A}_{si} = -\frac{\sum_{j:j\neq i} |J_{ij}|}{4RC_{0s}} A_{si} + \frac{1}{4RC_{0s}} \sum_{j:j\neq i} J_{ij} A_{sj}$$

$$- \frac{1}{2R_s C_{0s}} A_{si} + \frac{C_N \omega_0 A_{pi}}{2C_{0s}} A_{si} + \frac{A_{si}^{(n,I)}}{2C_{0s}}$$

$$- \frac{1}{4RC_{0s}} \sum_{j:j\neq i} |J_{ij}| A_{ij1}^{(n)} + \frac{1}{4RC_{0s}} \sum_{j:j\neq i} |J_{ij}| A_{ij0}^{(n)}, \quad \text{(B27)}$$

$$\dot{A}_{pi} = \frac{I_{pci} + A_{pi}^{(n,I)}}{2C_{0p}} - \frac{1}{2R_p C_{0p}} A_{pi} - \frac{C_N \omega_0 A_{si}^2}{2C_{0p}}. \quad \text{(B28)}$$

If the signal and pump band-pass filters are assumed to be perfectly rectangular with unit real frequency response, the impulse responses $h_s(\tau)$ and $h_p(\tau)$ satisfy

$$h_s(\tau) * h_s(-\tau) = h_s(\tau) = \frac{2\sin(\Delta\omega\tau/2)}{\pi\tau} \cos(\omega_0\tau), \quad \text{(B29)}$$

$$h_p(\tau) * h_p(-\tau) = h_p(\tau) = \frac{2\sin(\Delta\omega\tau/2)}{\pi\tau} \cos(2\omega_0\tau). \quad \text{(B30)}$$

Next, we introduce noise processes $n_{si}(t)$ and $n_{pi}(t)$ to capture the noise terms in Eqs. (B27) and (B28):

$$n_{si} = 2RA_{si}^{(n,I)} - \sum_{j:j\neq i} |J_{ij}| A_{ij1}^{(n)} + \sum_{j:j\neq i} |J_{ij}| A_{ij0}^{(n)}, \quad \text{(B31)}$$

$$n_{pi} = A_{pi}^{(n,I)}. \quad \text{(B32)}$$

Since the pump noise term (B32) is straightforward to implement in the MATLAB sde solver, we shift our attention to the signal noise. In words, Eq. (B31) tells us that the processes n_{si}, of which there are N, are linear combinations of $\mathcal{O}(N^2)$ independent Gaussian noise processes. We conclude from standard random-process theory that the n_{si} are Gaussian random processes too. This means they should be expressible as a linear combination of N independent Gaussian noise processes w_i instead of $\mathcal{O}(N^2)$ of them. This is a desirable representation because Eq. (B27) will then take the matrix-vector form

$$\dot{\mathbf{x}} = P\mathbf{x} + \frac{1}{4RC_{0s}} Q\mathbf{w} \quad \text{(B33)}$$

for some matrices P and Q. P is readily extracted from Eq. (B27), whereas Q is such that

$$n_{si} = \sum_j Q_{ij} w_j. \quad \text{(B34)}$$

Equation (B33) is a Langevin stochastic differential equation and can readily be simulated with the use of MATLAB's sde function.

We show next how to compute the matrix Q that leads to correlations that are consistent with Eq. (B31). The two-point correlation of $n_{si}(t)$ with itself (its autocorrelation) is as follows:

$$\langle n_{si}(0)n_{si}(\tau)\rangle = 4R^2 \langle A_{si}^{(n,I)}(0)A_{si}^{(n,I)}(\tau)\rangle$$
$$+ 2\sum_{j\,:\,j\neq i} |J_{ij}|^2 \langle A_{ij0}^{(n)}(0)A_{ij0}^{(n)}(\tau)\rangle$$
$$= 8kTR\left(\frac{2R}{R_s} + \sum_{j\,:\,j\neq i}|J_{ij}|\right)\frac{2\sin(\Delta\omega\tau/2)}{\pi\tau}. \tag{B35}$$

The two-point correlation between $n_{si}(t)$ and $n_{sk}(t)$ for $i \neq k$ is as follows:

$$\langle n_{si}(0)n_{sk}(\tau)\rangle = -8kTR|J_{ik}|J_{ik}\frac{2\sin(\Delta\omega\tau/2)}{\pi\tau}. \tag{B36}$$

Assuming that the processes w_i have autocorrelation $2\sin(\Delta\omega\tau/2)/\pi\tau$ and using Eq. (B34), we get

$$\langle n_{si}(0)n_{si}(\tau)\rangle = \left(\sum_j Q_{ij}^2\right)\frac{2\sin(\Delta\omega\tau/2)}{\pi\tau}, \tag{B37}$$

$$\langle n_{si}(0)n_{sk}(\tau)\rangle = \left(\sum_j Q_{ij}Q_{kj}\right)\frac{2\sin(\Delta\omega\tau/2)}{\pi\tau}. \tag{B38}$$

Equating the right-hand sides of Eqs. (B35) and (B37) and those of Eqs. (B36) and (B38), we see that Q is obtained by our performing the Cholesky decomposition of a matrix M constructed as follows:

$$M_{ik} = -8kTR|J_{ik}|J_{ik}\ \text{if}\ i \neq k, \tag{B39}$$

$$M_{ik} = 8kTR\left(\frac{2R}{R_s} + \sum_{j\,:\,j\neq i}|J_{ij}|\right)\ \text{if}\ i = k. \tag{B40}$$

This completes the discussion of the signal noise.

At this point, we comment on our code implementation. Firstly, our circuit had a cubic nonlinear saturating internal conductance $I = G_{\text{lin}}V + G_{\text{nonlin}}V^3$ in the signal circuit, so the $2R/R_s$ term in Eq. (B40) was replaced with $2R\left(G_{\text{lin}} + 9G_{\text{nonlin}}A_{\text{sat}}^2\right)$. Secondly, we faced difficulties with generating band-limited white noise with sinc autocorrelation, which is what the w_i need to be. For this reason, we simply used pure white noise (Dirac-δ autocorrelation) for the w_i. This assumption translates to forcing the slowly varying amplitudes to have Dirac-δ autocorrelation instead of the sinc autocorrelation that was derived in Eqs. (B23)–(B26).

TABLE III. LTspice-testing circuit parameter values.

Parameter	Value ($N = 5$)	Value ($N = 15$)
Signal capacitance C_s	$\frac{1}{2\pi}$ nF	$\frac{1}{2\pi}$ nF
Signal inductance L_s	$\frac{1}{2\pi}$ nH	$\frac{1}{2\pi}$ nH
Pump capacitance C_p	$\frac{0.5}{4\pi}$ nF	$\frac{0.5}{4\pi}$ nF
Pump inductance L_p	$\frac{2}{4\pi}$ nH	$\frac{2}{4\pi}$ nH
Linear capacitance C_0	0	0
Nonlinear capacitance C_N	$\frac{0.1}{2\pi}$ nF/V	$\frac{0.1}{2\pi}$ nF/V
Saturation voltage A_{sat}	0.1 V	0.07 V
Coupling resistance R	125 Ω	250 Ω
Pump internal resistance R_p	∞	∞
Signal internal conductance G_s	$1/R\ \Omega^{-1}$	$1/R\ \Omega^{-1}$
Nonlinear conductance G_N	$\frac{1}{RA_{\text{sat}}^2}\ \Omega^{-1}\,\text{V}^{-2}$	$\frac{1}{RA_{\text{sat}}^2}\ \Omega^{-1}\,\text{V}^{-2}$
Total time T	5 μs	5 μs

4. Extension to the case of nonzero local magnetic fields h_i

In a more-general form of the Ising problem, each spin also experiences a local magnetic field that adds to the total energy. The Hamiltonian is then $H = -\sum_i h_i x_i - \sum_{ij} J_{ij}x_i x_j$. This expression can be interpreted as an $N + 1$-spin Ising Hamiltonian, where $h_i/2$ are the $J_{i,N+1}$ connection coefficients of the first N spins to a newly introduced $N + 1$th spin that is fixed to orientation $+1$. This viewpoint enables us to minimize this new Hamiltonian by simply adding an ac voltage source with phase corresponding to $+1$ to the original circuit and connecting it to the other oscillators through $h_i/2$ resistors in a manner exactly analogous to that for the J resistors. The final signal equation of motion is

$$\dot{A}_{si} = \left[-\frac{|h_i|/2 + \sum_j |J_{ij}|}{4RC_{0s}}A_{si} + \frac{1}{4RC_{0s}}\frac{h_i}{2}A_{\text{sat}} \right.$$
$$\left. + \frac{1}{4RC_{0s}}\sum_{j\,:\,j\neq i}J_{ij}A_{sj} \right] + \frac{C_N\omega_0 A_{pi}}{2C_{0s}}A_{si}. \tag{B41}$$

5. LTspice versus slowly-varying-amplitude approximation

To ascertain the reliability of the slowly-varying-amplitude approximation vis-à-vis the performance of the real circuit, we simulated the coupled-N-spin-parametric-oscillator circuit shown in Fig. 2(c) using both the slowly-varying-amplitude approximation and the circuit simulator LTspice for three fully connected problems for each of two problem sizes, $N = 5$ and $N = 15$. Table III lists the circuit parameters that were used to generate the results reported in this section.

Figures 11(a)–11(c) and Figs. 11(d)–11(f) depict the temporal evolution of the signal and pump oscillator
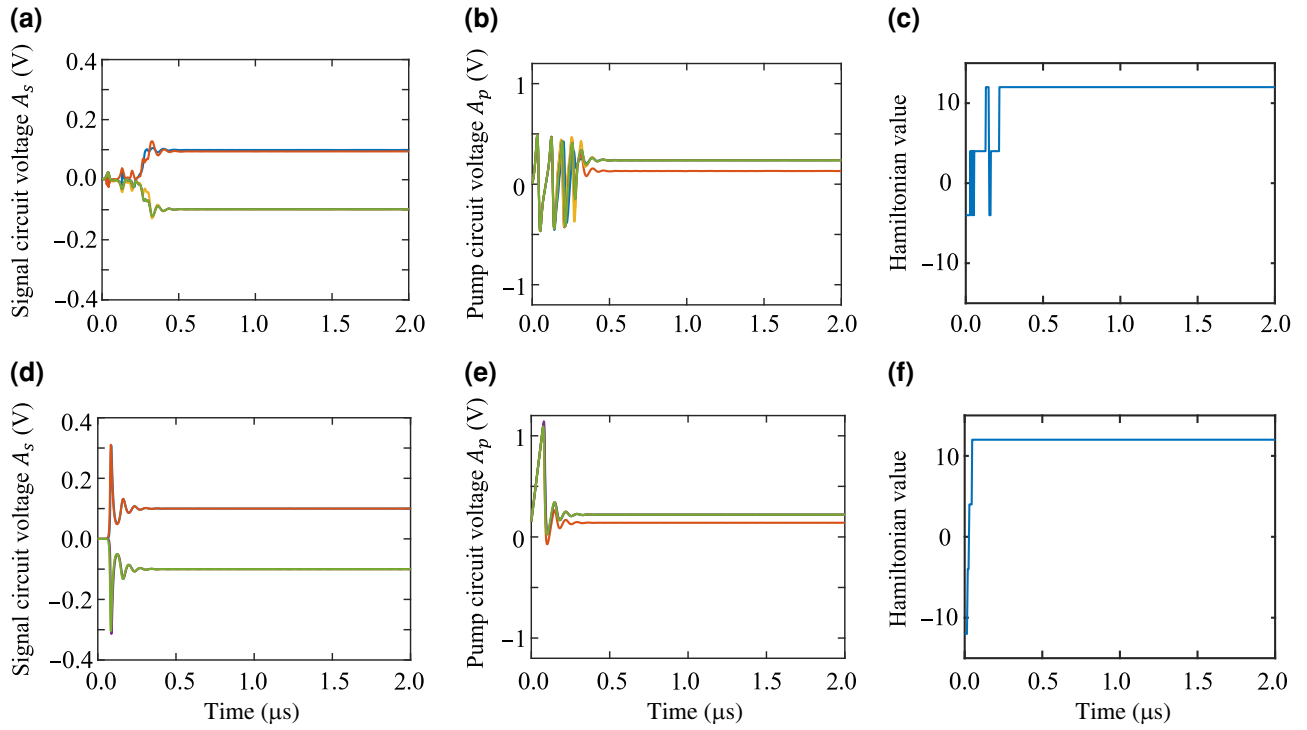
FIG. 11.   (a)–(c) LTspice results and (d)–(f) slowly-varying-amplitude-approximation results for an $N = 5$ problem . Temporal evolution of the (a),(d) amplitudes of the signal oscillators of all five spins, (b),(e) amplitudes of the pump oscillators of all five spins, and (c),(f) instantaneous Hamiltonian value. The signal and pump oscillators settle to the same values in both cases, and both solvers successfully find the global optimum ($H = 12$).

voltage amplitudes of a five-spin circuit obtained from the LTspice and slowly-varying-amplitude-approximation simulations, respectively. Figures 11(a) and 11(d) and Figs. 11(b) and 11(e) demonstrate that the signal and pump amplitudes both settle to the same values in both

simulations. Moreover, we see in Figs. 11(c) and 11(f) that both simulators successfully find the optimal solution—$H = 12$ in this case. A significant difference between the two simulations that is apparent from the plots is the behavior at early times—we believe this is due to

TABLE IV.   Comparison of the results of the LTspice and slowly-varying-amplitude-approximation simulations. Each problem was run 30 times (for 5 μs) with independent initial conditions, and the best Hamiltonian value and the Hamiltonian value at the final time step were collected for each run. The statistics of these runs are then compared against the true optimum.

				How solution of a run is chosen					
				Best value in run			Last value in run		
Problem	Size N	True optimum	Simulator	25th percentile	Median	75th percentile	25th percentile	Median	75th percentile
1	5	12	LTspice	12	12	12	12	12	12
			Slowly-varying-amplitude approximation	12	12	12	12	12	12
2	5	12	LTspice	4	12	12	4	12	12
			Slowly-varying-amplitude approximation	12	12	12	12	12	12
3	5	8	LTspice	8	8	8	8	8	8
			Slowly-varying-amplitude approximation	8	8	8	8	8	8
4	15	74	LTspice	74	74	74	74	74	74
			Slowly-varying-amplitude approximation	74	74	74	74	74	74
5	15	82	LTspice	82	82	82	74	82	82
			Slowly-varying-amplitude approximation	74	78	82	74	78	82
6	15	66	LTspice	66	66	66	66	66	66
			Slowly-varying-amplitude approximation	50	58	66	50	58	66

errors arising from the large time steps that we noticed were being chosen internally by LTspice during the simulation. Attempts to force LTspice to use smaller time steps using standard syntax did not succeed.

Table IV presents the results of the simulations on three problems each of size $N = 5$ and $N = 15$.

Each problem was run 30 times (for 5 μs each) with independent noisy starting conditions for the signal circuits, and two types of solution metric were recorded: (1) the best Ising solution during each 5-μs run and (2) the Ising solution at the last time step. For each problem, the median and 25% and 75% quantiles of these 30 runs are reported for both of the aforementioned solution metrics. It is clear from the reported value that both solvers succeed in finding the true optimum for all the problems, but with different success rates.

We believe the reason for this discrepancy is twofold: (1) the large time steps that LTspice takes in the simulations, hence losing important dynamical information, and (2) and the use of the nonstandard parameters listed in Table III instead of the optimal parameters listed in Table II. We observed experimentally that the optimal parameters in Table II yielded a near-100% success rate for the slowly-varying-amplitude solver for problems of this size, but led to tremendous numerical instabilities in LTspice. We used the less-optimal parameters in Table III in this section, sacrificing the performance of the slowly-varying-amplitude solver along the way, to obtain meaningful LTspice results and illustrate the consistency between the methods.

APPENDIX C: DUALITY AND THE SADDLE-POINT NATURE OF (x*, λ*)

Let us say we are searching for *constrained global minima* instead of *constrained local minima*. The problem we are trying to solve is as follows:

$$\text{minimize} \quad f(\mathbf{x})$$
$$\text{subject to} \quad g_i(\mathbf{x}) = 0, \ i = 1, \dots, p.$$

Standard optimization textbooks show that this problem can be rewritten as

$$\min_{\mathbf{x}: \ g_i(x)=0 \ \forall i} f(\mathbf{x}) = \min_{\mathbf{x}} \left(\max_{\lambda} L(\mathbf{x}, \lambda) \right), \qquad \text{(C1)}$$

where $L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_i \lambda_i g_i(\mathbf{x})$ is the Lagrange function. We have converted a constrained optimization problem into an *unconstrained* nested min-max optimization problem. The well-known min-max inequality that is true for arbitrary functions tells us that

$$\min_{\mathbf{x}} \left(\max_{\lambda} L(\mathbf{x}, \lambda) \right) \geq \max_{\lambda} \left(\min_{\mathbf{x}} L(\mathbf{x}, \lambda) \right). \qquad \text{(C2)}$$

This relation holds for any optimization problem and is also called "weak duality." For some special optimization problems—which include many common convex optimization problems—we actually have equality:

$$\min_{\mathbf{x}} \left(\max_{\lambda} L(\mathbf{x}, \lambda) \right) = \max_{\lambda} \left(\min_{\mathbf{x}} L(\mathbf{x}, \lambda) \right). \qquad \text{(C3)}$$

The above relation says that the constrained global minimum $\mathbf{x}^*$ of $f(\mathbf{x})$ and its associated multiplier λ^* form a saddle point of $L(\mathbf{x}, \lambda)$. To see why they form a saddle point of $L(\mathbf{x}, \lambda)$, note that $(\mathbf{x}, \arg\max_\lambda L(\mathbf{x}, \lambda))$ on the left-hand side represents a "1D" curved slice of the full space that passes through $(\mathbf{x}^*, \lambda^*)$. Moreover, $L(\mathbf{x}, \lambda)$ is minimized over this slice at $(\mathbf{x}^*, \lambda^*)$. Therefore, as we move away from $(\mathbf{x}^*, \lambda^*)$ along the tangent to this slice, L increases. Similarly, the right-hand side says that, over the "1D" curved slice represented by $(\arg\min_\mathbf{x} L(\mathbf{x}, \lambda), \lambda)$, $L(\mathbf{x}, \lambda)$ is maximized at $(\mathbf{x}^*, \lambda^*)$. Therefore, as we move away from $(\mathbf{x}^*, \lambda^*)$ along the tangent to this slice, L decreases.

APPENDIX D: AUGMENTED LAGRANGE CIRCUIT USING NONLINEAR RESISTORS—EQUATIONS OF MOTION

We insert a nonlinear resistor with the characteristic $I = G_0 V + G_N V^3$ in parallel with all the signal-circuit capacitors in the system to implement the cubic nonlinearity required by the augmented Lagrange equations of motion. The circuit equations from before, Eqs. (B1)–(B6), remain the same except for the first equation in Eq. (B1), which changes to

$$C_s \dot{V}_{si} = I_{3i} - I_{li} - \sum_j I_{ij0} - G_0 V_{si} - G_N V_{si}^3. \qquad \text{(D1)}$$

Solving all the equations as before, we find the counterparts of Eqs. (B10) and (B11) are as follows:

$$-\frac{N-1}{2R}\dot{V}_{si} + \frac{1}{2R}\sum_{j:j\neq i} J_{ij}\dot{V}_{sj} = \frac{V_{si}}{L_s} + C_{0s}\ddot{V}_{si} + 4C_N\dot{V}_{pi}\dot{V}_{si}$$
$$+ 2C_N V_{pi}\ddot{V}_{si} + 2C_N V_{si}\ddot{V}_{pi} + G_0\dot{V}_{si} + 3G_N V_{si}^2\dot{V}_{si}, \qquad \text{(D2)}$$

$$\dot{I}_{pi} = \frac{V_{pi}}{L_p} + C_{0p}\ddot{V}_{pi} + 2C_N\dot{V}_{si}^2 + 2C_N V_{si}\ddot{V}_{si}. \qquad \text{(D3)}$$

Using the slowly-varying-amplitude approximation and ignoring the sine components, we obtain

$$\dot{A}_{si} = \left[-\frac{N-1}{4RC_{0s}}A_{si} + \frac{1}{4RC_{0s}}\sum_{j\,:\,j\neq i} J_{ij}A_{sj} \right]$$
$$+ \frac{C_N\omega_0 A_{pi}}{2C_{0s}}A_{si} - \frac{G_0}{2C_{0s}}A_{si} - \frac{3G_N}{8C_{0s}}A_{si}^3, \quad \text{(D4)}$$

$$\dot{A}_{pi} = \frac{I_{pci}}{2C_{0p}} - \frac{C_N\omega_0 A_{si}^2}{2C_{0p}}. \quad \text{(D5)}$$

APPENDIX E: TRANSLATING QUADRATIC BINARY AND MAX-CUT INSTANCES INTO ISING INSTANCES

1. Quadratic binary to Ising

The BiqMac collection specifies quadratic binary $(0,1)$ *minimization* problems by listing the coefficients Q_{ij} of the terms $x_i x_j$ in the quadratic objective function. The coefficients form a symmetric matrix Q. The problem is stated precisely and recast as an Ising *maximization* problem below:

$$\mathbf{x}^* = \arg\min_{x_i\in\{0,1\}\,\forall i} \sum_{ij} Q_{ij} x_i x_j = \arg\min_{x_i\in\{0,1\}\,\forall i} \mathbf{x}^T Q \mathbf{x}$$

$$= \arg\max_{x_i=\pm1\,\forall i} \left[-\frac{1}{4}\mathbf{1}^T Q\mathbf{1} - \frac{1}{4}\sum_i Q_{ii} \right.$$

$$\left. -\frac{1}{4}2\mathbf{1}^T Q\mathbf{x} - \frac{1}{4}\mathbf{x}^T \widetilde{Q}\mathbf{x} \right]$$

$$= \arg\max_{x_i=\pm1\,\forall i} K + \frac{1}{4}\mathbf{h}^T\mathbf{x} + \frac{1}{4}\mathbf{x}^T J\mathbf{x},$$

where $\widetilde{Q}$ is the same as matrix Q but with the principal diagonal zeroed out, the effective Ising matrix $J_{ij} := -\widetilde{Q}_{ij}$, the effective Zeeman vector $\mathbf{h} := -2Q\mathbf{1}$, and the constant $K := -\frac{1}{4}\mathbf{1}^T Q\mathbf{1} - \frac{1}{4}\sum_i Q_{ii}$.

2. Max-cut to Ising

The Gset collection specifies max-cut problems by listing the edges ij and their weights w_{ij}. The max-cut optimization problem is stated and recast as an Ising problem below:

$$\mathbf{x}^* = \arg\max_{x_i=\pm1\,\forall i} \frac{1}{8}\sum_{ij} w_{ij}\left(x_i - x_j\right)^2$$

$$= \arg\max_{x_i=\pm1\,\forall i} \frac{1}{4}\sum_{ij} w_{ij} + \frac{1}{4}\sum_{ij}\left(-w_{ij}\right)x_i x_j$$

$$= \arg\max_{x_i=\pm1\,\forall i} K + \frac{1}{4}\sum_{ij} J_{ij} x_i x_j,$$

where we introduced the effective Ising matrix $J_{ij} := -w_{ij}$ and the constant $K := \frac{1}{4}\sum_{ij} w_{ij}$.

APPENDIX F: NUMERICAL RESULTS AND PARAMETER CHOICES

The slowly-varying-amplitude circuit equations were run on the Beasley binary quadratic problems in the Biq-Mac problem set [41] and on problems 1–10 (size 800) and problems 22–31 (size 2000) of the Gset max-cut problem set [40]. The Beasley binary quadratic problems involve minimizing a quadratic objective function where the feasible set is $0/1$ vectors and the function coefficients are positive and negative integers. Gset max-cut problems 1–5 and 22–26 have only $0,1$ edge weights, while problems 6–10 and 27–31 have $-1,0,1$ weights.

1. Parameter choices

Table V lists the circuit parameter definitions and values that were used in the simulations. The prefixes "signal" and "pump" are used to refer to components of the ω_0 signal and $2\omega_0$ pump circuits, respectively.

The linear capacitance and inductance values were chosen to set the natural frequency of the signal and pump oscillators to $\omega_0 = 1$ GHz and $2\omega_0 = 2$ GHz, respectively. The nonlinear capacitance C_N is chosen so that the modulation on the capacitance is 10% of C_s at an applied voltage of 1 V. The linear part C_0 of the nonlinear capacitance is assumed to be zero because any nonzero C_0 can be absorbed into C_s and C_p (made clear in the derivations in Appendix B). The signal internal resistance R_s and the linear part G_0 of the nonlinear conductor are merged into G_s. The voltage saturation amplitude A_{sat} of the signal oscillations is set to 10 mV.

Binary weights $J = \pm1$ are implemented according to the scheme in Fig. 2(c) by our connecting pairs of signal oscillators with resistors of a common value R in the appropriate configuration. Values of J other than ±1 are

TABLE V. Circuit parameter definitions and values.

Parameter	Value
Signal capacitance C_s	$(1/2\pi)$ nF
Signal inductance L_s	$(1/2\pi)$ nH
Pump capacitance C_p	$(0.01/4\pi)$ nF
Pump inductance L_p	$(100/4\pi)$ nH
Linear connecting cap C_0	0
Nonlinear connecting cap C_N	$(0.1/2\pi)$ nF/V
Signal saturation voltage A_{sat}	0.01 V
Common coupling resistance R	$(500\Gamma/47.94)$ Ω
Pump internal resistance R_p	∞
Signal internal conductance G_s	$1/R$ Ω^{-1}
Cubic nonlinear conductance G_N	$(1/RA_{\text{sat}}^2)$ $\Omega^{-1}\text{V}^{-2}$

implemented by our using a geometric series of resistances centered at R and the binary expansion of J (see Appendix B 2). We use a different value of R for each problem, set heuristically using the quantity Γ, which we call the "average coordination number" of the problem:

$$\Gamma = \frac{1}{N} \left(\sum_i \sum_{j:j \neq i} |J_{ij}| + \frac{1}{2} \sum_i |h_i| \right). \qquad \text{(F1)}$$

In the special case of 0/1 connections, Γ is the average number of nonzero connections to each spin. For the first Gset problem of size 800, we empirically found that our setting $R = 500\ \Omega$ satisfied the slowly-varying-amplitude approximation and led to good performance. This problem has an average coordination number Γ of 47.94. For other problems of average coordination number Γ', we set $R = (500\Gamma'/\Gamma)\ \Omega$.

To ensure that the isomorphism with Lagrange multipliers holds, the pump was assumed to have no internal dissipative loss, unless noted otherwise. The cubic coefficient G_N of the nonlinear conductor used in the augmented Lagrangian method is chosen so that the linear and cubic conductances are equal at the saturation voltage A_{sat}.

a. ODE-solver and SDE-solver settings

All simulations were run for a total (circuit) time of 50 µs. The noiseless calculations were done with the MATLAB ode45 solver while the noisy cases were run with the sde solver. The ode45 solver adaptively picks time steps, while a step size of 1 ns was chosen for the sde-solver calculations.

The rms noise voltage across the signal capacitor in equilibrium is given by $V_{\text{noise}} = \sqrt{kT/C_{0s}} \approx 5$ µV for the C_{0s} chosen in the main text. All the signal-circuit capacitor voltages at $t = 0$ start out at this noise level in all our computations with the initial condition for the ode45-solver computations following a continuous uniform distribution between $-V_{\text{noise}}$ and V_{noise} and the sde solver initial condition being chosen uniformly randomly from the discrete set $\{-V_{\text{noise}}, V_{\text{noise}}\}$.

On the other hand, there was no randomness in our choice of the initial condition for the pump voltages. The initial pump voltage is the same for all the spins and is chosen such that the system experiences net gain right from $t = 0$. From Eq. (B20), the losses of the various oscillation modes of the circuit are proportional to the eigenvalues of the matrix X, whose elements are $X_{ij} = \delta_{ij} \left(\sum_{k:k \neq i} |J_{ik}| \right) - (1 - \delta_{ij}) J_{ij}$, where δ_{ij} is the Kronecker δ. In the presence of a nonlinear saturating conductor, the losses increase further. We choose the initial pump voltage to create a gain that is a factor of 1.1 times larger than the 50th least loss in the system. If

the ith eigenvalue of a matrix M is denoted by $\lambda_i(M)$, and X is as defined earlier in this paragraph, our initial pump voltage for all oscillators in all computations is as follows:

$$A_{pi}(0) = \frac{1}{C_N \omega_0} \times 1.1 \times \left(\frac{\lambda_{50}(X)}{2R} + G_0 \right).$$

The nonlinear contribution to initial loss is ignored because all the signal amplitudes are initially at the noise level.

b. Pump capacitance

We recall from the discussion of Lagrange multipliers in the main text that a good heuristic method to find constrained minima of optimization problems that satisfy only weak duality is to perform a fast gradient ascent in λ and a slow gradient descent in $\mathbf{x}$. Equation (D5) tells us that the speed of gradient ascent in the A_{pi} (which are proportional to the Lagrange multipliers) directions is inversely proportional to C_{0p}. Therefore, reducing C_{0p} should increase the speed of pump voltage evolution, bringing the dynamics closer to the prescribed heuristic. This is demonstrated in Fig. 12, which shows that reducing the pump capacitance C_{0p} does indeed improve the solution quality. To produce the results shown in Fig. 12, the algorithm was run 10 times on the first Gset 800-spin problem for each value of C_{0p} on the x axis. The plot depicts the median and the 25th and 75th percentiles of the ten runs for each C_{0p} as a fraction of the best-known solution for this problem.
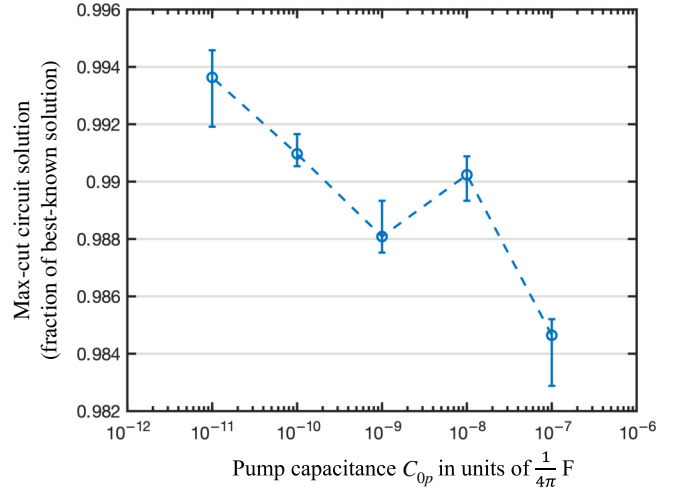

FIG. 12. Median, 25th-percentile, and 75th-percentile performance as a fraction of the best-known solution (henceforth called "normalized performance quartiles") of ten runs of the algorithm on the first Gset problem of size 800 for different values of pump capacitance.
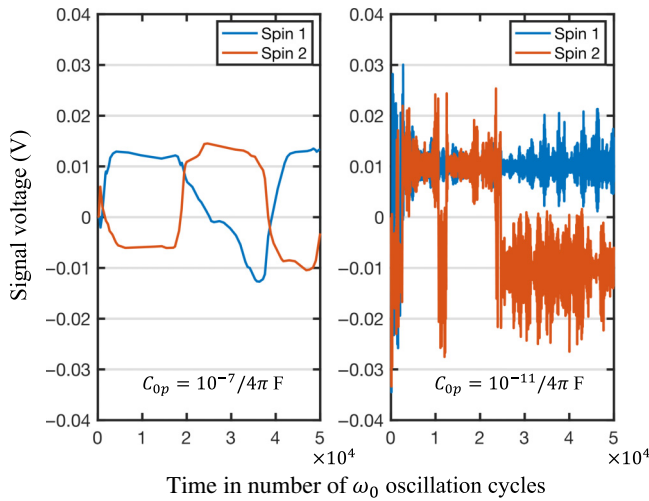
FIG. 13. Slow variation of signal voltage at high C_{0p} (left) and fast variation at low C_{0p} (right).

Reducing the pump capacitance C_{0p} increases the speed with which the pump equation responds to deviations of the signal voltage from the saturation amplitude, and this in turn increases the speed of voltage variations in the signal circuit itself, as shown in Fig. 13.

We used $C_{0p} = 10^{-11}/4\pi$ in our simulations due to its better performance. One possible danger of using too small a C_{0p} is that the fast variations it generates in the slowly varying amplitude could lead to a violation of the slowly-varying-amplitude approximation itself. Figure 14 shows an enlargement for the $C_{0p} = 10^{-11}/4\pi$ case and shows that the variation is on the order of hundreds of cycles, well within the validity regime of the slowly-varying-amplitude approximation.
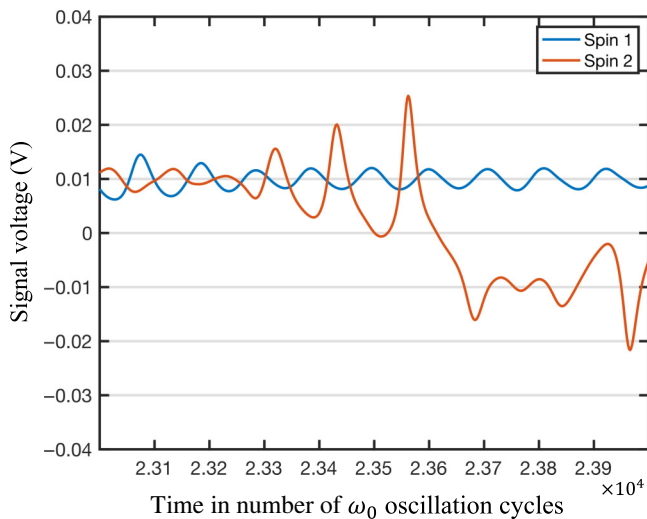

FIG. 14. Enlargement of the right panel in Fig. 13 showing that the slowly-varying-amplitude approximation is still valid.
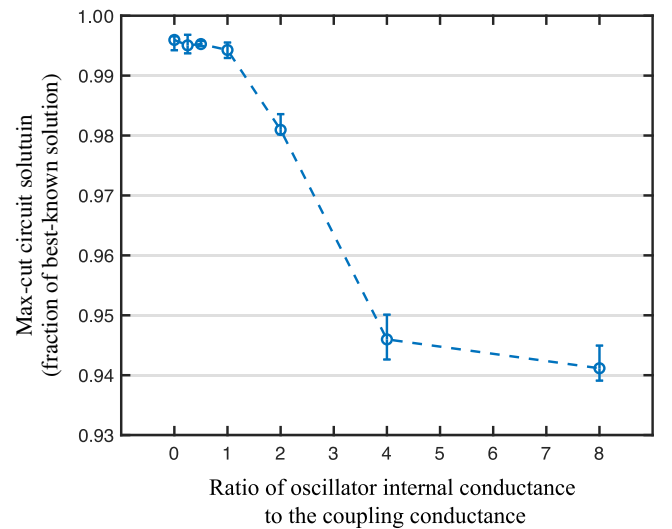

FIG. 15. Normalized performance quartiles of ten algorithm runs on Gset problem 6 (size 800) versus signal oscillator internal conductance G_0.

c. Effect of varying the strength of the internal nonlinear saturating conductor in the signal oscillators

The internal signal saturating conductor that implements the augmented Lagrange method is expressed as follows:

$$I = G_0 V + G_N V^3, \tag{F2}$$

G_N is pegged to G_0/A_{sat}^2. This ensures that, once the signal amplitude reaches A_{sat}, the nonlinearity kicks in and limits the voltage. Scaling G_0 up increases the "steepness" of the nonlinear barrier faced by the signal voltage. Numerical simulations with $G_0 = 1/R$, where R is the
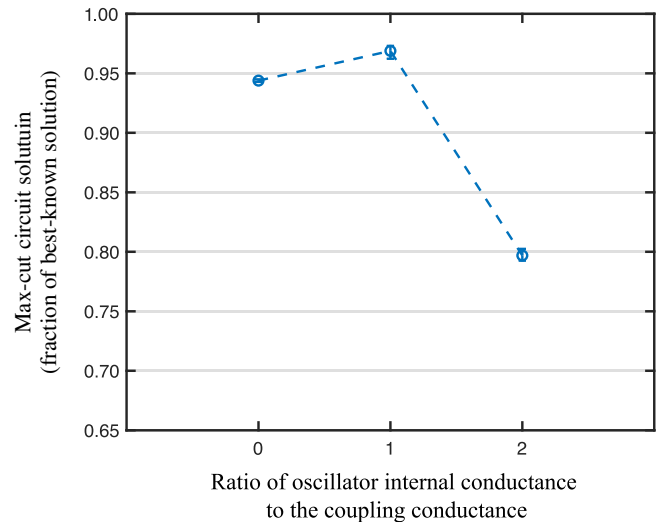

FIG. 16. Normalized performance quartiles of ten algorithm runs on Gset problem 27 (size 2000) versus signal oscillator internal conductance G_0.

common coupling resistance, yielded performance that matched or bettered the no-nonlinearity performance for the 800-spin and 2000-spin Gset problems—this demonstrates that the augmented Lagrange method is indeed better than the plain version. This is shown in Figs. 15 and 16 and also in Tables VI and VII.

In Figs. 15 and 16, the x axis shows the ratio $G_0/(1/R)$, while the plots themselves show the median, 25th-percentile, and 75th-percentile performance over ten runs at each x-axis point as a fraction of the best-known solution.

d. Effect of varying the nonlinear capacitance C_N

The product of C_N and A_{pi} is the parametric gain of the ith signal oscillator. Therefore, it is intuitive that varying C_N should not have much effect because the pump voltage can compensate for the change (Fig. 17).

2. Effect of variations in the pump Q between different spins

From the derivations provided earlier, the evolution of the pump voltage of the ith spin, under the assumption that the pump circuit is perfectly resistance free (infinite Q), is given by

$$\frac{dA_{pi}}{dt} = \frac{1}{2}\left[\frac{I_p}{C_0 + C_p} - \frac{C_N \omega_0 A_{si}^2}{C_0 + C_p}\right]. \tag{F3}$$

At any fixed point (steady state) of this equation, the right-hand side is zero, implying that the ith signal amplitude needs to satisfy $A_{si}^2 = I_p/C_N \omega_0 = A_{\text{sat}}^2$. In this case, we achieve perfect amplitude homogeneity across spins. In the presence of finite internal parallel pump resistance R_{pi},

however, the equation changes to

$$\frac{dA_{pi}}{dt} = \frac{1}{2}\left[\frac{I_p}{C_0 + C_p} - \frac{A_{pi}}{R_{pi}(C_0 + C_p)} - \frac{C_N \omega_0 A_{si}^2}{C_0 + C_p}\right]. \tag{F4}$$

At the steady state of this equation, the modified $A_{\text{sat},i}$ for each spin i is $A_{si}^2 = [I_p - (A_{pi}/R_{pi})]/C_N \omega_0 = A_{\text{sat},i}^2$. This formula reduces to the A_{sat} formula in the previous paragraph if $R_{pi} = \infty$ and gradually deviates from it as R_{pi} is reduced. Moreover, if the pump internal resistances R_{pi} are different for different spins i, both amplitude homogeneity and the equivalence of the pump evolution to Lagrange Ising constraint imposition are disturbed. Even if the R_{pi} values were equal for all i (but still finite), the pump evolution equation would still differ from the ideal Lagrange-multiplier dual equation $d\lambda_i/dt = \kappa'(x_{\text{sat}}^2 - x_i^2)$.

The last three columns in Table II show the impact of a finite pump Q factor (but pump Q was the same for all the spins i) on the performance. As expected, theperformance was worse for lossy pumps with low Q values and better for high Q values, which approximate lossless pumps.

In this section, we reported simulations of Lagrange Ising solver circuits that had random variations in the pump Q between spins for two 800-spin problems from the Gset problem set (Gset problems 1 and 2). Two values of the mean pump Q_{mean} were chosen, 500 and 5000. Each pump resistor in the circuit was chosen to be a Gaussian random variable with mean Q_{mean} and standard deviation dQ expressed as a percentage of Q_{mean}. The five percentages that were picked were 0%, 5%, 10%, 20%, and 30%. For
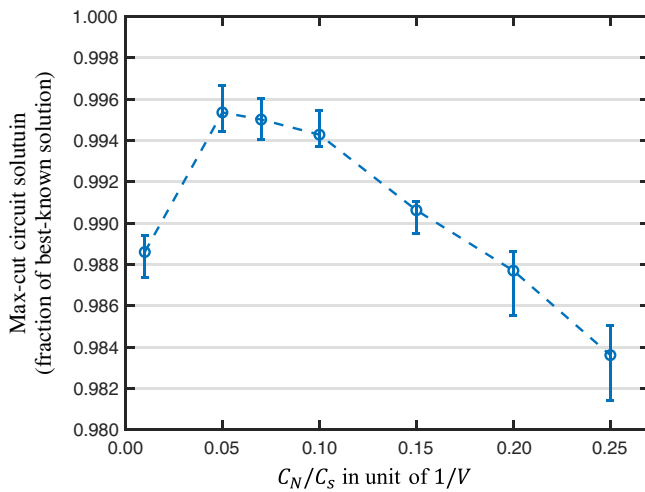

FIG. 17. Normalized performance quartiles of ten algorithm runs on Gset problem 1 (size 800) versus nonlinear capacitance C_N.
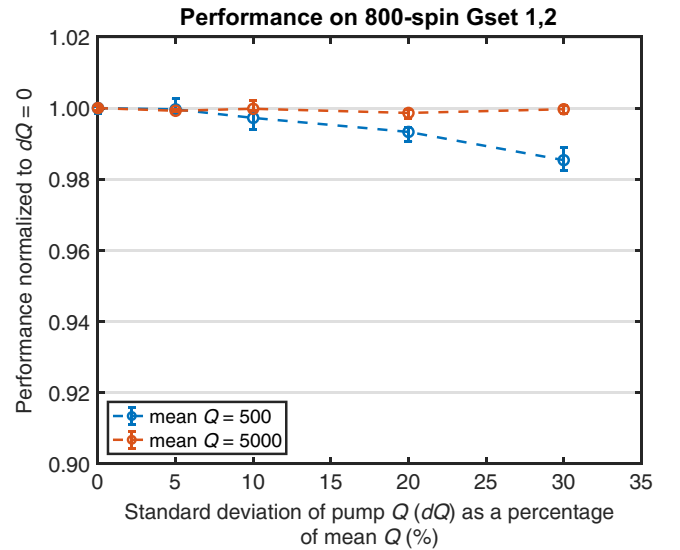

FIG. 18. Effect of pump Q variations between spins on the performance on 800-spin Gset problems 1 and 2.

each Gset problem, and for each mean Q_{mean} and standard deviation dQ, three random circuits were generated. In other words, six runs were performed in all for each pair (Q_{mean}, dQ). Thermal noise was taken into account, and an SDE solver was used to obtain results after a total time of 50 μs. The resultant performance was normalized to the performance of circuits with the same Q for all pumps (i.e., $dQ = 0$). The medians and quartiles of the normalized results jointly across both problems are plotted in Fig. 18.

It is clear that pump Q variations do not seem to matter for high mean Q ($Q_{mean} = 5000$ in Fig. 18) but begin to lead to performance degradations for lower mean Q ($Q_{mean} = 500$ in Fig. 18).

3. Back-of-the-envelope check of the power-dissipation numbers

In this subsection, we provide a quick back-of-the-envelope calculation of the circuit power dissipation for 800-spin Gset problem 1. The total power dissipation P_{tot} is the sum of the signal P_s, pump P_p, and coupling P_c dissipations. P_s is the sum of the dissipations in the linear and nonlinear parts of the signal internal resistances across all the spins and is given by $P_s \approx 800 \times 2 \times A_{sat}^2/R_s =$

$1600 \times (100 \times 10^{-6})/500 = 0.32$ mW. Since each spin in Gset problem 1 is connected to approximately 48 other spins on average, the total worst-case dissipation in the coupling resistors P_c is approximately $48 \times 800/2 \times (100 \times 10^{-6})/500 \times 2 \approx 8$ mW. Since this worst-case estimate was obtained by our assuming that all the spin connections in the network were frustrated (there exists no edge for which the orientations of the spins connected to it satisfy the edge value), the actual coupling dissipation has to be significantly lower than this. Finally, using $Q_{pump} = 500$ (the corresponding pump resistance R_p is approximately 50 kΩ for the values used in this paper) and $V_{pump} \approx 0.5$ V (from the bottom panel in Fig. 3), we find the pump dissipation P_p is approximately $800 \times 0.25/(5 \times 10^4) \approx 4$ mW. The total dissipation is therefore upper bounded as $P_{tot} \leq 12.32$ mW.

4. Results for more Gset problems

We present results for the oscillator Lagrange solver for Gset problems 1–10 (size 800) and problems 22–31 (size 2000) in Tables VI and VII. While the Lagrange-multiplier method is outperformed by the approach of Leleu *et al.*, clever amalgamation of the two ideas could lead to better hybrid algorithms in the future.

TABLE VI. Performance on Gset problems 1–10 (size 800) of the Goemans-Williamson (G-W) algorithm, the upper bound (UB) it implies, the approach of Leleu *et al.*, and coupled oscillators without nonlinear resistors (Osc) and with nonlinear resistors (Osc NL). For the last two columns, the best and median values are reported for ten independent runs, while the results for Leleu *et al.* are taken from Ref. [19]. The ratio of the G-W-algorithm performance and the G-W UB is not 88% in rows 6–10 because of the presence of negative weight edges in those problems.

Problem	G-W algorithm	G-W UB	Metric	Leleu *et al.*	Osc	Osc NL
1	11 272	12 838	Best	11 624	11 580	11 613
			Median	11 624	11 552	11 558
2	11 277	12 844	Best	11 620	11 575	11 596
			Median	11 620	11 554	11 572
3	11 289	12 857	Best	11 622	11 588	11 586
			Median	11 622	11 560	11 562
4	11 301	12 871	Best	11 646	11 611	11 641
			Median	11 646	11 586	11 590
5	11 293	12 862	Best	11 631	11 591	11 578
			Median	11 631	11 568	11 562
6	1813	3387	Best	2178	2143	2173
			Median	2178	2124	2144
7	1652	3224	Best	2006	1975	1973
			Median	2006	1950	1955
8	1667	3243	Best	2005	1966	1992
			Median	2005	1948	1961
9	1704	3278	Best	2054	2010	2043
			Median	2054	1991	2006
10	1646	3218	Best	2000	1956	1979
			Median	2000	1940	1955

TABLE VII. Performance on Gset problems 22–31 (2000 vertices) of the approach of Leleu *et al.* and coupled oscillators without nonlinear resistors (Osc) and with nonlinear resistors (Osc NL). In the column headed by "Leleu *et al.*," three center dots are used to indicate values that could not be deduced from Ref. [19].

Problem	Metric	Leleu *et al.*	Osc	Osc NL
22	Best	13 359	13 191	13 255
	Median	$\cdots$	13 176	13 231
23	Best	13 342	13 178	13 277
	Median	13342	13151	13228
24	Best	13 337	13 166	13 259
	Median	13 337	13 150	13 232
25	Best	13 340	13 170	13 263
	Median	13 340	13 154	13 228
26	Best	13 328	13 155	13 252
	Median	$\cdots$	13 142	132 28
27	Best	3341	3171	3275
	Median	3341	3156	3237
28	Best	3298	3132	3230
	Median	3298	3112	3185
29	Best	3405	3221	3328
	Median	3405	3206	3302
30	Best	3413	3252	3332
	Median	$\cdots$	3226	3287
31	Best	3310	3144	3223
	Median	$\cdots$	3125	3203

[1] R. M. Karp, in *Complexity of Computer Computations* (Springer, 1972), p. 85.

[2] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, J. ACM **42,** 1115 (1995).

[3] U. Benlic and J.-K. Hao, Breakout local search for the Max-Cutproblem, Eng. Appl. Artif. Intell. **26,** 1162 (2013).

[4] A. Lucas, Ising formulations of many NP problems, Front. Phys. **2,** 5 (2014).

[5] T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. L. McMahon, T. Umeki, K. Enbutsu, O. Tadanaga, H. Takenouchi, K. Aihara, K.-i. Kawarabayashi, K. Inoue, S. Utsunomiya, and H. Takesue, A coherent Ising machine for 2000-node optimization problems, Science **354,** 603 (2016).

[6] M. W. Johnson, M. H. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, *et al.*, Quantum annealing with manufactured spins, Nature **473,** 194 (2011).

[7] H. Goto, K. Tatsumura, and A. R. Dixon, Combinatorial optimization by simulating adiabatic bifurcations in nonlinear Hamiltonian systems, Sci. Adv. **5,** eaav2372 (2019).

[8] H. Goto, Quantum computation based on quantum adiabatic bifurcations of Kerr-nonlinear parametric oscillators, J. Phys. Soc. Jpn. **88,** 061015 (2019).

[9] H. Goto, K. Endo, M. Suzuki, Y. Sakai, T. Kanao, Y. Hamakawa, R. Hidaka, M. Yamasaki, and K. Tatsumura, High-performance combinatorial optimization based on classical mechanics, Science Advances. **7,** eabe7953 (2021).

[10] P. J. Van Laarhoven and E. H. Aarts, in *Simulated Annealing: Theory and Applications* (Springer, 1987), p. 7.

[11] F. Cai, S. Kumar, T. Van Vaerenbergh, X. Sheng, R. Liu, C. Li, Z. Liu, M. Foltin, S. Yu, Q. Xia, *et al.*, Power-efficient combinatorial optimization using intrinsic noise in memristor Hopfield neural networks, Nat. Electron. **3,** 409 (2020).

[12] M. N. Bojnordi and E. Ipek, in *2016 IEEE International Symposium on High Performance Computer Architecture (HPCA)* (IEEE, 2016), p. 1.

[13] M. Mahmoodi, M. Prezioso, and D. Strukov, Versatile stochastic dot product circuits based on nonvolatile memories for high performance neurocomputing and neurooptimization, Nat. Commun. **10,** 1 (2019).

[14] S. Kumar, J. P. Strachan, and R. S. Williams, Chaotic dynamics in nanoscale NbO_2 Mott memristors for analogue computing, Nature **548,** 318 (2017).

[15] S. Patel, P. Canoza, and S. Salahuddin, Logically synthesized and hardware-accelerated restricted Boltzmann machines for combinatorial optimization and integer factorization, Nat. Electron. **5,** 92 (2022).

[16] C. Roques-Carmes, Y. Shen, C. Zanoci, M. Prabhu, F. Atieh, L. Jing, T. Dubček, C. Mao, M. R. Johnson, V. Ceperić, J. D. Joannopoulos, D. Englund, and M. Soljačić, Heuristic recurrent algorithms for photonic Ising machines, Nat. Commun. **11,** 249 (2020).

[17] D. Pierangeli, G. Marcucci, and C. Conti, Large-scale photonic Ising machine by spatial light modulation, Phys. Rev. Lett. **122,** 213902 (2019).

[18] F. L. Traversa and M. Di Ventra, Polynomial-time solution of prime factorization and NP-complete problems with digital memcomputing machines, Chaos **27,** 023107 (2017).

[19] M. Di Ventra and F. L. Traversa, Perspective: Memcomputing: Leveraging memory and physics to compute efficiently, J. Appl. Phys. **123,** 180901 (2018).

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

[21] S. Reifenstein, T. Leleu, T. McKenna, M. Jankowski, M. Suh, E. Ng, F. Khoyratee, Z. Toroczkai, and Y. Yamamoto, Coherent SAT solvers: A tutorial, Adv. Opt. Photonics **15,** 385 (2023).

[22] B. Molnár, F. Molnár, M. Varga, Z. Toroczkai, and M. Ercsey-Ravasz, A continuous-time MaxSAT solver with high analog performance, Nat. Commun. **9,** 4864 (2018).

[23] M. Ercsey-Ravasz and Z. Toroczkai, Optimization hardness as transient chaos in an analog approach to constraint satisfaction, Nat. Phys. **7,** 966 (2011).

[24] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-bits for invertible logic, Phys. Rev. X **7,** 031014 (2017).

[25] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, Integer factorization using stochastic magnetic tunnel junctions, Nature **573,** 390 (2019).

[26] Y. Haribara, S. Utsunomiya, and Y. Yamamoto, Computational principle and performance evaluation of coherent Ising machine based on degenerate optical parametric oscillator network, Entropy **18,** 151 (2016).

[27] M. C. Strinati, L. Bello, E. G. Dalla Torre, and A. Pe'er, Can nonlinear parametric oscillators solve random Ising models?, Phys. Rev. Lett. **126,** 143901 (2021).

[28] D. Chermoshentsev, A. Malyshev, M. Esencan, E. Tiunov, D. Mendoza, A. Aspuru-Guzik, A. Fedorov, and A. Lvovsky, Polynomial unconstrained binary optimisation inspired by optical simulation, Preprint ArXiv:2106.13167.

[29] 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).

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

[31] T. Wang and J. Roychowdhury, in *Unconventional Computation and Natural Computation*, Vol. 11493, edited by I. McQuillan and S. Seki (Springer International Publishing, Cham, 2019), p. 232.

[32] J. Chou, S. Bramhavar, S. Ghosh, and W. Herzog, Analog coupled oscillator based weighted Ising machine, Sci. Rep. **9,** 1 (2019).

[33] I. Ahmed, P.-W. Chiu, W. Moy, and C. H. Kim, A probabilistic compute fabric based on coupled ring oscillators for solving combinatorial optimization problems, IEEE J. Solid-State Circuits **56,** 2870 (2021).

[34] W. Moy, I. Ahmed, P.-w. Chiu, J. Moy, S. S. Sapatnekar, and C. H. Kim, A 1,968-node coupled ring oscillator circuit for combinatorial optimization problem solving, Nat. Electron. **5,** 310 (317).

[35] S. Dutta, A. Khanna, A. Assoa, H. Paik, D. Schlom, Z. Toroczkai, A. Raychowdhury, and S. Datta, An Ising Hamiltonian solver based on coupled stochastic phase-transition nano-oscillators, Nat. Electron. **4,** 502 (2021).

[36] M. Babaeian, D. T. Nguyen, V. Demir, M. Akbulut, P.-A. Blanche, Y. Kaneda, S. Guha, M. A. Neifeld, and N. Peyghambarian, A single shot coherent Ising machine based on a network of injection-locked multicore fiber lasers, Nat. Commun. **10,** 3516 (2019).

[37] K. P. Kalinin and N. G. Berloff, Global optimization of spin Hamiltonians with gain-dissipative systems, Sci. Rep. **8,** 17791 (2018).

[38] S. K. Vadlamani, T. P. Xiao, and E. Yablonovitch, Physics successfully implements Lagrange multiplier optimization, Proc. Natl. Acad. Sci. **117,** 26639 (2020).

[39] S. K. Vadlamani, T. P. Xiao, and E. Yablonovitch, in *2022 IEEE International Conference on Rebooting Computing (ICRC)* (IEEE, San Francisco, 2022), p. 45.

[40] https://web.stanford.edu/yyye/yyye/gset/.

[41] A. Wiegele, Biq Mac library—A collection of max-cut and quadratic 0-1 programming instances of medium size, Preprint **51,** (2007).

[42] A. Yariv and W. Louisell, 5A2 - theory of the optical parametric oscillator, IEEE J. Quantum. Electron. **2,** 418 (1966).

[43] T. P. Xiao, Ph.D. thesis (University of California, Berkeley, 2019), ISBN: 9781085792172 Publication Title: ProQuest Dissertations and Theses 13886372.

[44] S. K. Vadlamani, Ph.D. thesis (University of California, Berkeley, 2021), ISBN: 9798535557618 Publication Title: ProQuest Dissertations and Theses 28540365.

[45] https://www.nii.ac.jp/qis/first-quantum/e/forStudents/lecture/pdf/noise/chapter11.pdf.

[46] D. Bertsekas, *Nonlinear Programming* (Athena Scientific, 1999).

[47] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).

[48] M. X. Goemans and D. P. Williamson, The primal-dual method for approximation algorithms and its application to network design problems, Approximation Algorithms for NP-Hard Problems (1997), p. 144.

[49] F. Z. Bi and B. P. Barber, Bulk acoustic wave RF technology, IEEE Microw. Mag. **9,** 65 (2008).

[50] C. C. Ruppel, Acoustic wave filter technology—A review, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **64,** 1390 (2017).

[51] T. P. Xiao, C. H. Bennett, B. Feinberg, S. Agarwal, and M. J. Marinella, Analog architectures for neural network acceleration based on non-volatile memory, Appl. Phys. Rev. **7,** 031301 (2020).

[52] 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, Sci. Adv. **5,** eaau0823 (2019).

[53] https://resources.nvidia.com/en-us-tensor-core/nvidia-tensor-core-gpu-datasheet.