Escaping Local Minima in Ising Machines through Analog Floquet Solvers

Nicolas Casilli[1], Seunghwi Kim[2], Sunil Mittal[1], Marvin Onabajo[1], Andrea Alù[2,3], and Cristian Cassella[1,†]

[1]Institute for Nano Systems Innovation (NanoSI), Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02115, USA

[2]Photonics Initiative, Advanced Science Research Center, City University of New York, New York City, New York 10031, USA

[3]Physics Program, Graduate Center, City University of New York, New York City, New York 10016, USA

[†]Corresponding author: c.cassella@northeastern.edu

Supplementary Table S1: Table of Acronyms

Acronym	Definition
QUBO	Quadratic Binary Unconstrained Optimization
NP	Nondeterministic Polynomial
IM	Ising Machine
PO	Parametric Oscillator
AFS	Analog Floquet Solver
CMOS	Complementary Metal-Oxide Semiconductor
D	Dimension, in higher-order spin-solvers
x_p	Resonant cavity pump mode
$\chi^{(2)}$	Type of nonlinearity
ω_p	Pump frequency driving the pump mode
x	Signal mode
$\omega_p/2$	Oscillation frequency of the signal mode parametrically down converted from ω_p
J	Coupling matrix
i	Index of the i^{th} spin
j	Index of the j^{th} spin
J_{ij}	Coupling strength between the i^{th} and j^{th} spins
σ_i, σ_j	i^{th} and j^{th} spin-states
H_{Ising}	Ising Hamiltonian
κ_p	Loss-rate of the pump mode
κ	Loss-rate of the signal mode
r_1	Amplitude or gain saturation coefficient
r_2	Small-signal gain parameter
y	Resonant mode
g	Linear coupling between x and y
Γ	Loss-rate of the resonant mode
ω_x	Natural angular frequency of the signal mode

ω_y	Natural angular frequency of the resonant mode
n	Separation index of Floquet harmonics
M	Separation index of Floquet harmonics
A_n	Amplitude coefficient of the n^{th} Floquet harmonic of the x mode
B_n	Amplitude coefficient of the n^{th} Floquet harmonic of the y mode
ϕ_n	Relative phase shift of the n^{th} Floquet harmonic
μ	Separation frequency between adjacent Floquet harmonic
N	Number of spins in a coupled system
x_i	Signal mode of the i^{th} spin
κ_i	Loss-rate of the signal mode
g_i	Linear coupling between the i^{th} x and y modes
Γ_i	Loss-rate of the i^{th} resonant mode
F_C	Cost Function describing the dynamics of coupled Floquet nodes
η	Interaction, or coupling, term in F_C
$A_{i,n}$	Amplitude coefficient of the n^{th} Floquet harmonic of the i^{th} x mode
$\phi_{i,n}$	Phase of the n^{th} Floquet harmonic of the i^{th} signal mode
μ_i	Separation frequency between adjacent Floquet harmonic of the i^{th} x mode
A_i	Amplitude coefficient of the 1^{st} Floquet harmonic of the i^{th} x mode
η_0	Static component of $\eta(t)$
η_1	Modulated component of $\eta(t)$
Φ_i	Extracted and binarized phase of the i^{th} spin
$\bar{\eta}$	Normalized interaction term of F_c
G	Coupling between x and x_p modes
κ_{ex}	Extra-cavity mode loss-rate
x_{in}	Pump strength exciting the x_p and x modes
X	Amplitude solution form of x
θ	Phase solution form of x
F_{PO}	Cost Function describing the dynamics of coupled POs
η_{PO}	Interaction term in F_{PO}
$B_{i,n}$	Amplitude coefficient of the n^{th} Floquet harmonic of the i^{th} resonant mode (y)
B_i	Amplitude coefficient of the 1^{st} Floquet harmonic of the i^{th} signal mode (y)
$\phi_{x_i} - \phi_{y_i}$	Phase difference between the x and y modes
f_s	Sampling rate used to track the solutions identified by the AFS
$\bar{H}_{min}$	The average of H_{min} over, in this case, 100 runs

I. Analytical Studies of Systems via Coupled Mode Theory

A. Parametric Oscillators and their Coupled Systems

In this section, we address the dynamics of the nonlinear systems presented in the main manuscript. We begin with the model of a single PO, which will be the fundamental backbone for all following analysis in the main manuscript and supplementary material. POs, which consists of modes x_p and x mediated by the $\chi^{(2)}$ nonlinear process, can be described with a system of two equations, Eq. $(S.1a)$ and Eq. $(S.1b)$ [1], [2]:

$$\dot{x_p} = -\left(\frac{\kappa_p}{2} + i\omega_\mathrm{p}\right)x_p - iGx^2 + \sqrt{\kappa_{ex}}x_{in}e^{-i\omega_p t}, \qquad (S.1a)$$

$$\dot{x} = -\left(\frac{\kappa}{2} + i\omega_x\right)x - iG\,x^* x_p. \qquad (S.1b)$$

This system is such that when x_p is driven with a pump signal (ω_p) with strength x_{in} and incurring a loss-rate $\sqrt{\kappa_{ex}}$, x is indirectly excited at $\omega_p/2$ through parametric down-conversion via the intramode coupling G. Under the assumption that $\kappa_p \gg \kappa$ (κ_p being the loss-rate of x_p and κ being the loss-rate of x), the mode x_p can be adiabatically eliminated under a rotating frame of $x \to xe^{-i\frac{\omega_p}{2}t}$ and $x \to xe^{-\frac{i\pi}{4}}$ to produce a single equation, Eq. $(S2)$.

$$\dot{x} = -\left(\frac{\kappa}{2} + r_1|x|^2\right)x + r_2 x^*. \qquad (S.2)$$

Through this adiabatic elimination, the nonlinear parameters linking x and x_p take the form of $r_1 = \frac{2|G|^2}{\kappa_p}$ and $r_2 = \frac{2G\sqrt{\kappa_{ex}}x_{in}}{\kappa_p}$ (being r_1 the gain saturation coefficient and r_2 the small-signal gain parameter) [1]. POs are regarded as highly effective candidates for use as macroscopic spins in Ising Machines due to their inherent phase bistability, which is achieved following the activation of their period-doubling regime following a bifurcation [3], [4], [5], [6], [7], [8], [9]. To understand the conditions via which this bistability appears, we can re-express Eq. $(S.2)$ at steady state after changing the form of x to $Xe^{-i\theta}$ and constraining $\theta = 0, \pi$ [2].

$$0 = -\left(\frac{\kappa}{2} + i\Delta_x + r_1 X^2\right) X + r_2 X. \tag{S.3}$$

When $r_2 > \frac{\kappa}{2}$, we can obtain a solution to Eq. $(S.3)$ as $X_{th} = \sqrt{\frac{r_2 - \frac{\kappa}{2}}{r_1}}$ [2]. For reference, Fig. 1(b) in the main manuscript illustrates the frequency spectrum of the x mode of a PO after undergoing a period doubling bifurcation when driven with $r_2 > \frac{\kappa}{2}$. A collective of POs can then be coupled together with a coupling strength $J_{i,j}$ (such that $J_{i,j} = J_{j,i}$) between the i^{th} and j^{th} POs to create a PO-based IM. In this regard, POs with a phase solution $\theta = 0$ can represent a "spin-up" in the Ising Model while those with a $\theta = \pi$ represent a "spin-down". Such a coupled system of N coupled POs can be modelled using temporal coupled mode theory with a system of equations taking the familiar form, shown in Eq. $(S.4)$:

$$\dot{x_i} = -\left(\frac{\kappa}{2} + r_1 |x_i|^2\right) x_i + r_2 x_i^* - \sum_j^N J_{i,j} x_j. \tag{S.4}$$

Following the same procedure as shown in the main manuscript to determine F_C, it is relatively straightforward to construct $F_{PO}(x_1, x_1^*, x_2 x_2^*, \dots, x_N x_N^*)$, a cost function for the PO-based IM, as shown in Eq. $(S.5)$ [10], [11]:

$$F_{PO} = \sum_i^N \left[\frac{\kappa}{2}|x_i|^2 + \frac{r_1}{2}|x_i|^4 - \frac{r_2}{2}\left({x_i}^2 + {x_i^*}^2\right)\right] + \frac{1}{2}\sum_{i,j}^N x_i^* J_{i,j} x_j. \tag{S.5}$$

F_{PO} obeys $-\frac{\partial F_{PO}}{\partial x_i^*} = \dot{x_i}$ and $-\frac{\partial F_{PO}}{\partial x_i} = \dot{x_i^*}$. To verify this relation, we differentiate F_{PO} with respect to x_i^* and x_i, and we achieve the results seen in Eq. $(S.6a)$ and Eq. $(S.6b)$.

$$-\frac{\partial F_{PO}}{\partial x_i^*} = -\frac{\kappa}{2} x_i - r_1 |x_i|^2 x_i + r_2 x_i^* - igy_i - \sum_j J_{i,j} x_j = \dot{x_i} = \frac{\partial x_i}{\partial t} \qquad (S.6a)$$

$$-\frac{\partial F_{PO}}{\partial x_i} = -\frac{\kappa}{2} x_i^* - r_1 |x_i|^2 x_i^* + r_2 x_i - igy_i^* - \sum_j J_{i,j} x_j^* = \dot{x_i^*} = \frac{\partial x_i^*}{\partial t} \qquad (S.6b)$$

These conditions ensure that $\frac{\partial F_{PO}}{\partial t} = -2\sum_i^N |\dot{x_i}|^2$ (meaning that $\frac{\partial F_{PO}}{\partial t}$ is negative semi-definite throughout the trajectories of x_i), as $\frac{\partial F_{PO}}{\partial t} = \sum_i^N \left[\frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial t} + \frac{\partial F}{\partial x_i^*}\frac{\partial x_i^*}{\partial t}\right] = \sum_i^N [\dot{x_i^*}\,\dot{x_i} + \dot{x_i}\dot{x_i^*}]$. Furthermore, it can be shown that at the steady state or fixed point of the system (where all $\dot{x_i} = 0$), $\frac{\partial F_{PO}}{\partial t} = 0$. Such conditions permit to treat F_{PO} as a Lyapunov function since it is thus bounded from below by 0 [4].

As can be seen in Eq. $(S.5)$, two processes contribute to the minimization of F_{PO}; namely, self-sustaining oscillations of each PO and the collective dynamics between POs [11]. The interaction term describing the collective dynamics of a PO-based IM can be expressed as: $\eta_{PO} = \frac{1}{2}\sum_{i,j}^N x_i^* J_{i,j} x_j$. Note that this term bears significant resemblance to the Ising Hamiltonian, $H_{Ising} = \frac{1}{2}\sum_{i,j}^N \sigma_i J_{i,j} \sigma_j$; in fact, when all the POs' amplitudes are equal, $\eta_{PO} \propto H_{Ising}$, as x_i can be treated as $A_i\sigma_i$, where A_i represents the amplitude of x_i. Furthermore, under such conditions, the minimization of η_{PO} directly corresponds to the minimization of H_{Ising}, as obtained through gradient descent [12]. Yet, while such a reliance on gradient descent optimization offers PO-based IMs unique computational advantages compared to von Neumann computers when solving QUBO problems, PO-based IMs suffer from an intrinsic susceptibility to become trapped in sub-optimal solutions corresponding to local minima without any recourses to escape [4].

B. The Nonlinear Dynamics of Floquet Nodes and their Coupled Systems

In this section, we undertake a more thorough derivation of $\frac{\partial F_C}{\partial t}$ and more generally describe the AFS's computational dynamics by assuming periodic envelopes for both the x and y modes [$x(t) = \sum_{n=-\infty}^{\infty} A_n(t)e^{i(n\mu t+\phi_n(t))}$ and $y(t) = \sum_{n=-\infty}^{\infty} B_n(t)e^{i(n\mu t+\theta_n(t))}$]. Particularly, we show that $\frac{\partial F_C}{\partial t}$ is negative semi-definite despite the activation of the Floquet state.

First, we begin by revisiting the properties of F_C, namely: $-\frac{\partial F_C}{\partial x_i^*} = \dot{x}_i$, $-\frac{\partial F_C}{\partial x_i} = \dot{x}_i^*$, $-\frac{\partial F_C}{\partial y_i^*} = \dot{y}_i$, and $-\frac{\partial F_C}{\partial y_i} = \dot{y}_i^*$. Eqs. $(S.7a)$ - $(S.7d)$, below, show these relations in more explicit forms:

$$-\frac{\partial F_C}{\partial x_i^*} = -\frac{\kappa}{2}x_i - r_1|x_i|^2x_i + r_2x_i^* - igy_i - \sum_j J_{i,j}x_j = \dot{x}_i = \frac{\partial x_i}{\partial t} \qquad (S.7a)$$

$$-\frac{\partial F_C}{\partial x_i} = -\frac{\kappa}{2}x_i^* - r_1|x_i|^2x_i^* + r_2x_i - igy_i^* - \sum_j J_{i,j}x_j^* = \dot{x}_i^* = \frac{\partial x_i^*}{\partial t} \qquad (S.7b)$$

$$-\frac{\partial F_C}{\partial y_i^*} = -\frac{\Gamma}{2}y_i - igx_i = \dot{y}_i = \frac{\partial y_i}{\partial t} \qquad (S.7c)$$

$$-\frac{\partial F_C}{\partial y_i} = -\frac{\Gamma}{2}y_i^* - igx_i^* = \dot{y}_i^* = \frac{\partial y_i^*}{\partial t} \qquad (S.7d)$$

Now it can be observed that the computation of $\frac{\partial F_C}{\partial t}$ requires determining $\sum_i^N \left[\frac{\partial F_C}{\partial x_i}\frac{\partial x_i}{\partial t} + \frac{\partial F_C}{\partial x_i^*}\frac{\partial x_i^*}{\partial t} + \frac{\partial F_C}{\partial y_i}\frac{\partial y_i}{\partial t} + \frac{\partial F_C}{\partial y_i^*}\frac{\partial y_i^*}{\partial t}\right]$, which can be re-expressed as Eq. $(S.8)$.

$$\frac{\partial F_C}{\partial t} = -\sum_i^N [\dot{x}_i^*\dot{x}_i + \dot{x}_i\dot{x}_i^* + \dot{y}_i^*\dot{y}_i + \dot{y}_i\dot{y}_i^*] = -2\sum_i^N [|\dot{x}_i|^2 + |\dot{y}_i|^2] \leq 0 \qquad (S.8)$$

Next, we endeavor to express $|\dot{x_i}|^2 + |\dot{y_i}|^2$ in their periodic forms to understand the implications of the activation of the Floquet states on the minimization of F_C. In this regard, owing to the rotational frame $x \rightarrow xe^{-\frac{i\pi}{4}}$, we can express $x_i = A_i \cos(\mu t + \phi_i)$ and $y_i = -iB_i \sin(\mu t + \phi_i)$, as is done in the main manuscript. Like before, we retain only the first Floquet harmonics since we treat $A_{i,1} \gg \sum_{n=2} A_{i,n}$ and $B_{i,1} \gg \sum_{n=2} B_{i,n}$ (note that the two modes operate in quadrature owing to dispersive coupling mediated by ig, with a constant phase difference of $\frac{\pi}{2}$). To visualize the mode amplitudes and confirm that our assumptions are reasonable, we numerically extracted time and frequency domain decompositions of both x_i and y_i, as well as their phase difference with respect to frequency, as seen in Fig. S1.

Consequently, we can explicitly formulate $|\dot{x_i}|^2$, as seen in Eq. $(S.9)$:

$$|\dot{x_i}|^2 = \dot{x_i}\dot{x_i}^* = \begin{matrix} \left[\dot{A_i}\cos(\mu t + \phi_i) - A_i(\mu)\sin(\mu t + \phi_i)\right] * \\ \left[\dot{A_i}^*\cos(\mu t + \phi_i) - A_i^*(\mu)\sin(\mu t + \phi_i)\right] \end{matrix} \quad (S.9)$$

which simplifies to Eq. $(S.10)$:

$$|\dot{x_i}|^2 = \begin{matrix} \left|\dot{A_i}\right|^2 \cos^2(\mu t + \phi_i) + |A_i|^2(\mu)^2 \sin^2(\mu t + \phi_i) \\ -\frac{1}{2}\left(\dot{A_i}^* A_i + \dot{A_i} A_i^*\right)(\mu)\sin\left(2(\mu t + \phi_i)\right) \end{matrix} \quad (S.10)$$

In a similar fashion, we can express $|\dot{y_i}|^2$ as Eq. $(S.11)$:

$$|\dot{y_i}|^2 = \begin{matrix} \left|\dot{B_i}\right|^2 \sin^2(\mu t + \phi_i) + |B_i|^2(\mu)^2 \cos^2(\mu t + \phi_i) \\ +\frac{1}{2}\left(\dot{B_i}^* B_i + \dot{B_i} B_i^*\right)(\mu)\sin(2(\mu t + \phi_i)) \end{matrix} \quad (S.11)$$

Additionally, we assume that the amplitude terms, A_i and B_i, vary on a time-scale much larger than the oscillation period $\frac{1}{\mu}$. Henceforth, using the slow-envelope approximation, we can set $\dot{A_i} =$

μA_i, $\dot{A}_i^* = \mu A_i^*$, $\dot{B}_i = \mu B_i$, and $\dot{B}_i^* = \mu B_i^*$ to express Eqs. $(S.10)$ and $(S.11)$ as Eqs. $(S.12a)$ and $(S.12b)$, as seen below [13], [14]:

$$|\dot{x}_i|^2 = \left|\dot{A}_i\right|^2\left[1 - \sin\left(2(\mu t + \phi_i)\right)\right] \qquad (S.12a)$$

$$|\dot{y}_i|^2 = \left|\dot{B}_i\right|^2\left[1 + \sin\left(2(\mu t + \phi_i)\right)\right] \qquad (S.12b)$$

Furthermore, these forms of $|\dot{x}_i|^2$ and $|\dot{y}_i|^2$ permit to express $\frac{\partial F_C}{\partial t}$ as Eq. $(S.13)$:

$$\frac{\partial F_C}{\partial t} = -2\sum_i^N \left[\left|\dot{A}_i\right|^2 + \left|\dot{B}_i\right|^2 + \left(\left|\dot{B}_i\right|^2 - \left|\dot{A}_i\right|^2\right)\sin\left(2(\mu t + \phi_i)\right)\right] \leq 0 \qquad (S.13)$$

In the limit where $\mu, \phi_i \to 0$ (being the conditions for the period-doubling regime), $\frac{\partial F_C}{\partial t}$ simplifies to $-2\sum_i^N\left[\left|\dot{A}_i\right|^2 + \left|\dot{B}_i\right|^2\right]$ and η simplifies to $\frac{1}{2}\sum_{i,j}\left[A_i J_{i,j} A_j\right]$, so the system's modulations vanish. In other words, η exactly matches η_{PO}, so the system minimizes its energy as if it were a PO-based IM. Thus, the AFS implements a gradient descent optimization when driven in the period-doubling regime [3], [6], [7], [15].

On the other hand, when the AFS is driven such that its POs exhibit activated Floquet states, η_1 (equal to $\frac{1}{4}\sum_{i\neq j}^N\left[J_{i,j}A_{i,1}A_{j,1}\cos\left(2\mu t + \left(\phi_{i,1} + \phi_{j,1}\right)\right)\right]$) is endowed with a modulation of frequency 2μ, which is twice the characteristic Floquet exponent describing the modulation frequency of the modes x and y. In this regard, we have plotted in Fig. S2 the normalized frequency spectra of η and the normalized frequency spectra of a randomly selected mode amplitude x_i for the AFS computation conducted in Fig. 2(a) and Fig. 2(c). As can be seen, the mode x_i only contains various Floquet harmonics evenly spaced by μ, which has been numerically determined to have a value of $0.00285\frac{\omega_p}{2}$. On the other hand, the spectrum of η features a strong peak at DC, representing the energy minimization process arbitrated via gradient descent or η_0, while its second largest frequency peak occurs at $0.0057\frac{\omega_p}{2}$, representing the modulation process at 2μ appearing in η_1.

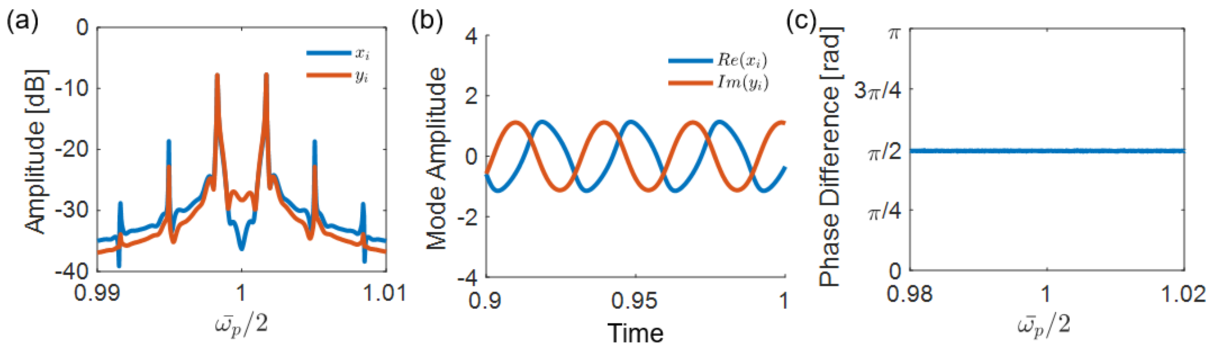

Figure S1 a) Frequency domain spectra of $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$ when driven with $\boldsymbol{r_2 = 0.7\kappa}$. **b)** Time domain waveforms of $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$ that demonstrates the different solution forms of the two modes. **c)** The phase difference between $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$ across a range of frequencies normalized by $\boldsymbol{\frac{\omega_p}{2}}$. As can be seen, the modes are separated with a phase of $\boldsymbol{\frac{\pi}{2}}$ throughout the entire spectrum due to the dispersive coupling between $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$ mediated by $\boldsymbol{g}$.

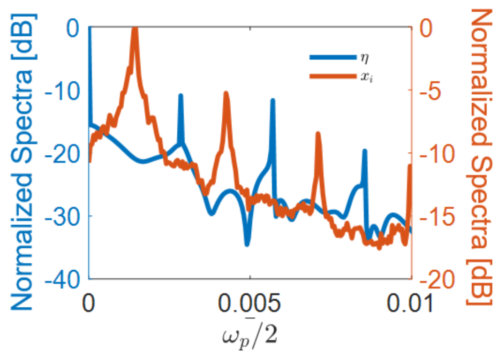

Figure S2 Frequency spectrum of η for the 20-node 3-regular graph solved in the main manuscript, as well as the frequency spectrum of an amplitude mode x_i. Both quantities are normalized such that their highest peak corresponds to a value of 0 dB.

II. Solution Identification

A. Analog and Digital Hamiltonians

In this section, we expound the nuances between η and η_{PO} and H_{Ising}. Specifically, we focus on the relationship between the minimization of a spin-solver's interaction term and the minimization of the Ising Hamiltonian to show how the analog nature of the AFS can help it reach the ground state of H_{Ising} whereas the analog nature of PO-based IMs can harm their ability to determine the ground state of H_{Ising}.

η and η_{PO}, which map to the interaction term in the systems' cost functions, consider each node's continuous amplitude dynamics rather than just the discretized spin state, σ. Furthermore, interactions between nodes i and j are governed not only by their coupling term $J_{i,j}$, but also by the relative difference in amplitude between the spins, which is commonly referred to as amplitude heterogeneity [16]. On the other hand, the Ising Hamiltonian, H_{Ising}, deals only with discretized spin states in $\{+1, -1\}$ and is completely agnostic to heterogeneous amplitudes, as intended in the Ising Model.

While IMs are meant to minimize a QUBO problem mapped to H_{Ising}, they instead minimize their continuous or analog interaction term (η or η_{PO}), which dynamically embeds an uncontrolled distortion of the mapped QUBO problem due to the system's heterogeneous amplitudes. As such, the minimization of η_{PO} in PO-based IMs can lead to suboptimal solutions to the problem graph mapped by H_{Ising} when the heterogeneity between amplitudes is sufficiently large (as η_{PO} no longer properly maps to H_{Ising}) [3]. Yet, proper mapping between η and H_{Ising} of the QUBO problem is not strictly necessary in the AFS, as long as the modulations in η permit the system to somehow reach the ground state of H_{Ising} at some point along the computation. In this regard, the

AFS is not confined to relying on strategies for amplitude heterogeneity reduction to solve QUBO problems, like conventional PO-based IMs. Instead, the AFS even benefits from the amplitude heterogeneities introduced by periodic modulations since these modulations enable deeper landscape explorations, permitting the system to identify several degenerate ground state or near-ground state solutions throughout a single computation.

To demonstrate impact of the analog computing platform on the minimization of H_{Ising}, we plot in Figs. S3(a) and S3(b) $\bar{\eta}_{PO}$ and H_{Ising} of the PO-based IM as well as $\bar{\eta}$ and H_{Ising} of the AFS, as extracted from the simulations of the 20-node Max-Cut problem solved in Fig. 2(a). $\bar{\eta}$ and $\bar{\eta}_{PO}$, which are the normalized interaction terms of the AFS and the PO-based IM, respectively, are computed by setting the average envelope of x_i to 1 (to match the amplitude of σ_i) in η and η_{PO}. As can be seen, most of the reduction of H_{Ising} when using the PO-based IM occurs when the nodes' amplitudes are very small [thus, η_{PO} is very close to 0; Fig. S3(a)]. Consequently, small perturbations in the mode amplitudes at that stage in the computation can lead to drastic changes in the system's overall ability to identify the ground state. When the node amplitudes are already very small, these perturbations introduce significant amplitude heterogeneity and distort the mapping between H_{Ising} (what is intended to be minimized) and η_{PO} (what is indeed minimized). Similarly, the AFS also undergoes a majority of the minimization of its corresponding H_{Ising} during the early stages of its computation, while the nodes' amplitudes are still small. Yet, the eventual emergence of amplitude modulations assist the AFS to identify a lower value of H_{Ising} by oscillating the energetic landscape, leading the AFS to benefit from its continuous node amplitudes rather than be harmed by them, as is often the case with PO-based IMs.

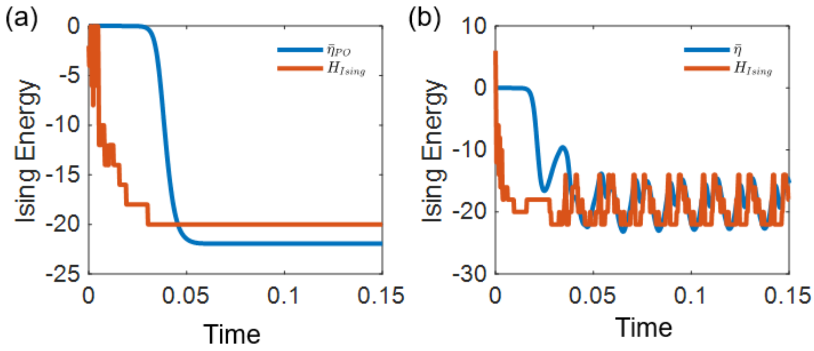

Figure S3 Numerically extracted H_{Ising} and $\bar{\eta}$ for the 20-node 3-regular problem graph solved in the main manuscript when considering **a)** a PO-based IM and **b)** an AFS. One can intuitively see how distortions in problem mapping generated by amplitude heterogeneity leads to more adverse computation outcomes when implementing only gradient descent to minimize the problem energy.

B. Recurrent Sampling

The reliability of the AFS to reach the best solutions is dependent on the sampling rate f_s at which the system's waveforms are sampled and H_{Ising} is computed. In our analysis, we have been computing H_{Ising} at each time-step of the numerical simulation, leading to a value of f_s that is approximately equal to 180μ. With such high f_s, we have tracked the H_{Ising} identified by the AFS with very high fidelity. However, as is analogous in the realm of signal processing, lower f_s can lead to degradations in the resolution of H_{Ising}, which can lead to a less complete evaluation of the values of H_{Ising} identified by the AFS. To help visualize this point, Fig. S4(a) shows the impact of f_s on the solution quality extracted by the AFS by comparing the sampled reconstruction of H_{Ising} when the AFS tackles the 200-node 3-regular problem from Fig. 3(a) when $f_s = 180\mu$ and when $f_s = 9\mu$. Unsurprisingly, the lower f_s leads to a degradation in the resolution of H_{Ising} and even causes the AFS to miss certain values of H_{Ising} that were extracted when $f_s = 180\mu$.

Furthermore, we decided also to study the impact of f_s on the solution quality of the AFS to understand the necessary sampling requirements to ensure that the AFS outperforms the PO-based IM. To do this, we computed the average minimum identified value of H_{Ising} over 100 runs, $\overline{H}_{min}$, when considering a range of f_s between $\mu/10$ and 180μ for both the AFS and the PO-based IM. As expected, $\overline{H}_{min}$ deteriorates as f_s decreases when using the AFS but it is unchanged when using the PO-based IM. Yet, for all $f_s > \frac{\mu}{4}$, the AFS outperforms the PO-based IM.

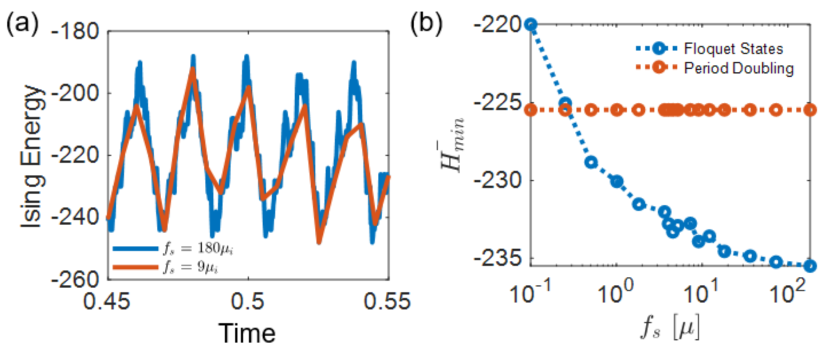

Figure S4 a) Numerically extracted H_{Ising} when considering $f_s = 180\mu$ and $f_s = 9\mu$ for the 200-node 3-regular problem graph solved in Fig. 3(a) in the main manuscript. Specifically, H_{Ising} is centered around the middle of the computation time. As can be seen, the resolution of H_{Ising} dramatically worsens with a twentyfold reduction in f_s. **b)** Numerically extracted relationship between $\overline{H}_{min}$ and f_s for the PO-based IM and the AFS. As f_s decreases, the solution quality of the AFS goes through a commensurate degradation. For $f_s > \frac{\mu}{4}$, the AFS reliably accesses better quality solutions than the PO-based IM.

III. Impact of r_2 on Computational Performance

In this section, we study the impact of the pump power, r_2 on the computational performance of the AFS to determine optimal operating conditions. Changes in r_2 result in commensurate shifts in both the characteristic Floquet exponent and each Floquet harmonic's amplitude, as seen in Fig. S5, which shows the frequency spectra of a single Floquet node when driven by three different r_2

values (namely, $r_2 = 0.7\kappa, 0.8\kappa,$ and 0.9κ). As elucidated in [1], the dependence between r_2 and μ can be interpreted as a pre-synchronization between modes, which finally occurs once the period-doubling regime is triggered.

In the context of computation, the value of r_2 therefore sets the frequency of the modulation in η and H_{Ising}, thus effectively configuring the system's evolution and exploration throughout the landscape. Furthermore, to characterize the impact of r_2 on the solution quality obtained by the AFS when solving QUBO problems, we solved a set of Max-Cut problems with varying size and coupling density for the same r_2 as those reported in Fig. S5 ($r_2 = 0.7\kappa, 0.8\kappa,$ and 0.9κ). We solved a set of randomly generated 3-regular, 5-regular, and 7-regular Max-Cut problem graphs with N ranging from 40 to 200 in steps of 40 nodes. Each problem graph was solved 100 times for each r_2 when considering different randomly generated initial conditions (ranging between -10^{-5} to 10^{-5}). The system's H_{Ising} is continuously monitored throughout each computation and the lowest attained H_{Ising}, which we deem as the solution of the highest quality, is extracted for each run. Fig. S6(a)-(e) show the distribution of the lowest identified H_{Ising} when the AFS tackles 3-regular graphs of varying sizes. Particularly for $N = 40$ and $N = 80$, there seems to be no discernible difference in the distribution of H_{Ising}. For higher N, it seems that $r_2 = 0.7\kappa$ identifies marginally lower energy levels. On the other hand, Fig. S7(a)-(e) and Fig. S8(a)-(e), which respectively present the distribution of H_{Ising} against 5-regular and 7-regular graph topologies with varying sizes indicate very limited differences in the solution quality achieved with respect to r_2.

Next, we conducted a similar study characterizing the impact of r_2 on the solution quality of the PO-based IM. In this regard, we minimized the same set of problem graphs as before, but this time using Eq. $(S.4)$ driven with the same values of r_2 and considering a different set of randomly generated initial conditions (but within the same range as before). We again continuously monitor the system's energy and identify the lowest energy attained during each run. Figs. S9(a)-(e), Figs. S10(a)-(e), and Figs. S11(a)-(e) show the distribution of solutions for each combination of problem graph dimension, coupling density, and r_2. In this study, we noted a more clearly defined relation between the quality of the extracted solutions and r_2, with $r_2 = 0.7\kappa$ identifying the lowest values of H_{Ising}. Nonetheless, the solutions identified by the PO-based IM were, on average, worse than the solutions obtained by the AFS and driven by the same r_2, as seen in Fig. 3 [17], [18]. These extracted trends of solution quality v. r_2 indicate that the AFS is more agnostic to r_2 than its PO-based IM counterparts.

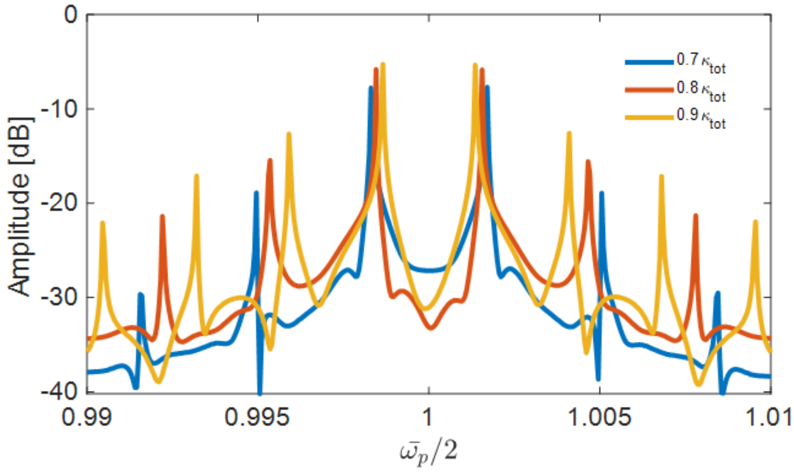

Figure S5 Numerically extracted frequency domain spectra of a the $\boldsymbol{x}$ mode of a single Floquet node when driven with $\boldsymbol{r_2} = \mathbf{0.7}\boldsymbol{\kappa}, \mathbf{0.8}\boldsymbol{\kappa}$, and $\mathbf{0.9}\boldsymbol{\kappa}$. As can be seen, changes in $\boldsymbol{r_2}$ lead to commensurate shifts in $\boldsymbol{\mu}$ as well as slight changes in each Floquet harmonic's amplitude.

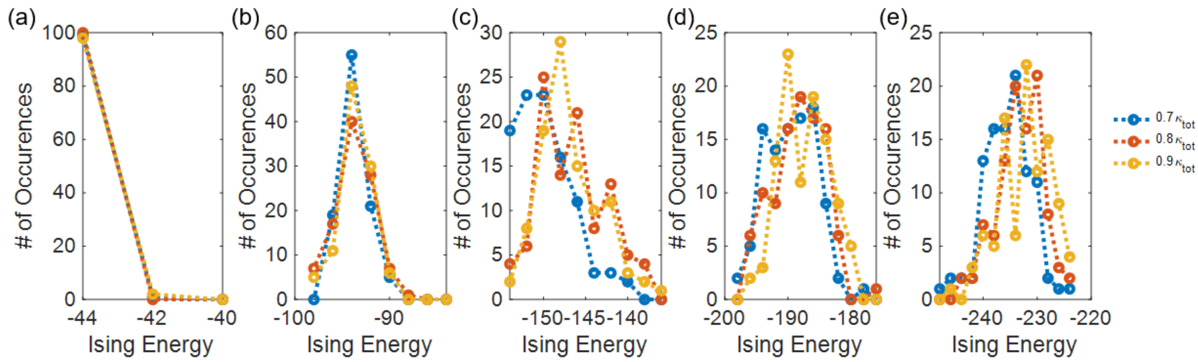

Figure S6 Numerically extracted energy values over 100 runs for a: **a)** 40-node 3-regular problem graph, **b)** 80-node 3-regular problem graph, **c)** 120-node 3-regular problem graph, **d)** 160-node 3-regular problem graph, and **e)** 200-node 3-regular problem graph solved by the AFS. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2} = \mathbf{0.7}\boldsymbol{\kappa}, \mathbf{0.8}\boldsymbol{\kappa}$, and $\mathbf{0.9}\boldsymbol{\kappa}$.

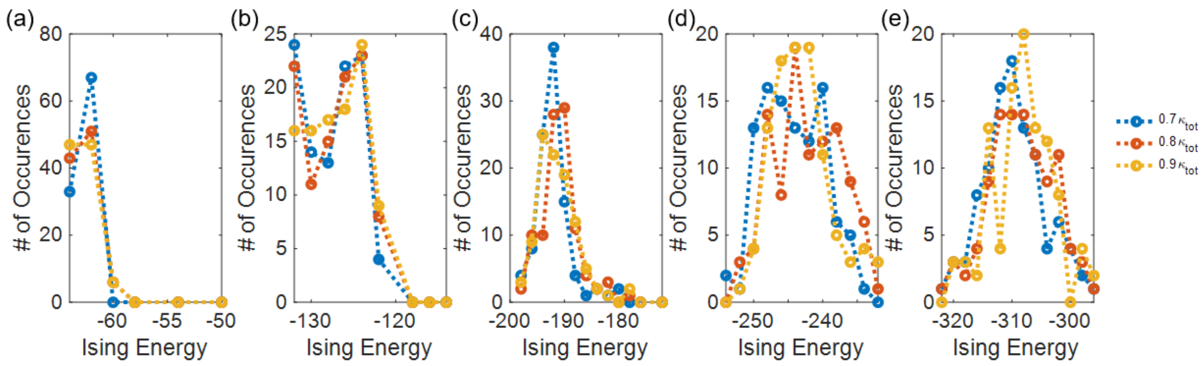

Figure S7 Numerically extracted energy values over 100 runs for a: **a)** 40-node 5-regular problem graph, **b)** 80-node 5-regular problem graph, **c)** 120-node 5-regular problem graph, **d)** 160-node 5-regular problem graph, and **e)** 200-node 5-regular problem graph solved by the AFS. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2 = 0.7\kappa, 0.8\kappa}$, and $\boldsymbol{0.9\kappa}$.

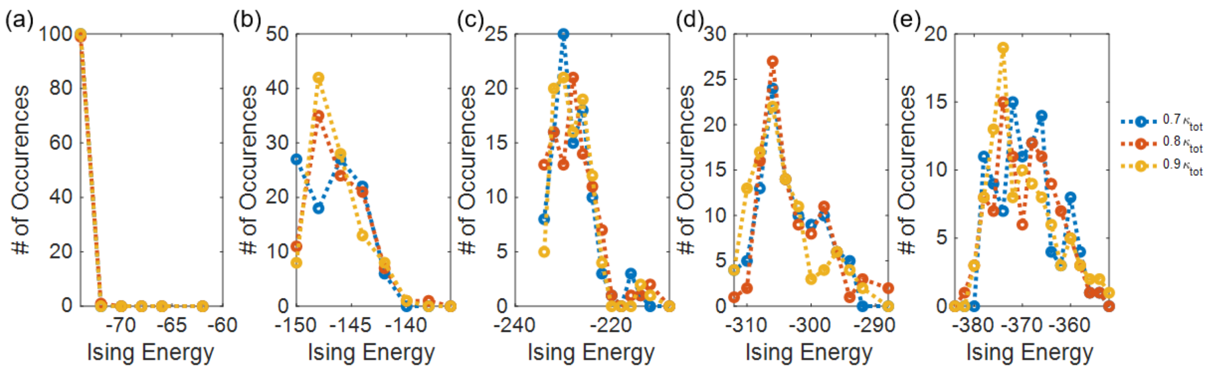

Figure S8 Numerically extracted energy values over 100 runs for a: **a)** 40-node 7-regular problem graph, **b)** 80-node 7-regular problem graph, **c)** 120-node 7-regular problem graph, **d)** 160-node 7-regular problem graph, and **e)** 200-node 7-regular problem graph solved by the AFS. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2 = 0.7\kappa, 0.8\kappa}$, and $\boldsymbol{0.9\kappa}$.

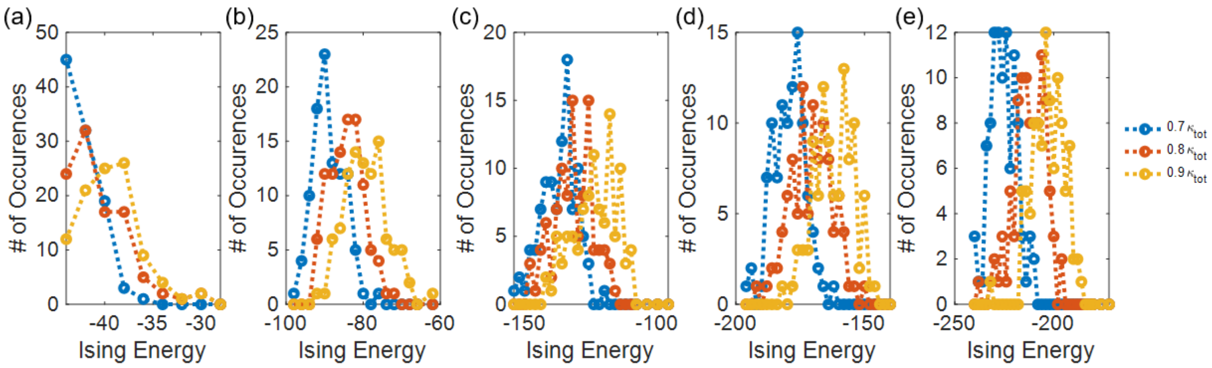

Figure S9 Numerically extracted energy values over 100 runs for a: **a)** 40-node 3-regular problem graph, **b)** 80-node 3-regular problem graph, **c)** 120-node 3-regular problem graph, **d)** 160-node 3-regular problem graph, and **e)** 200-node 3-regular problem graph solved by a PO-based IM. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2} = \mathbf{0.7}\boldsymbol{\kappa}, \mathbf{0.8}\boldsymbol{\kappa}$, and $\mathbf{0.9}\boldsymbol{\kappa}$.

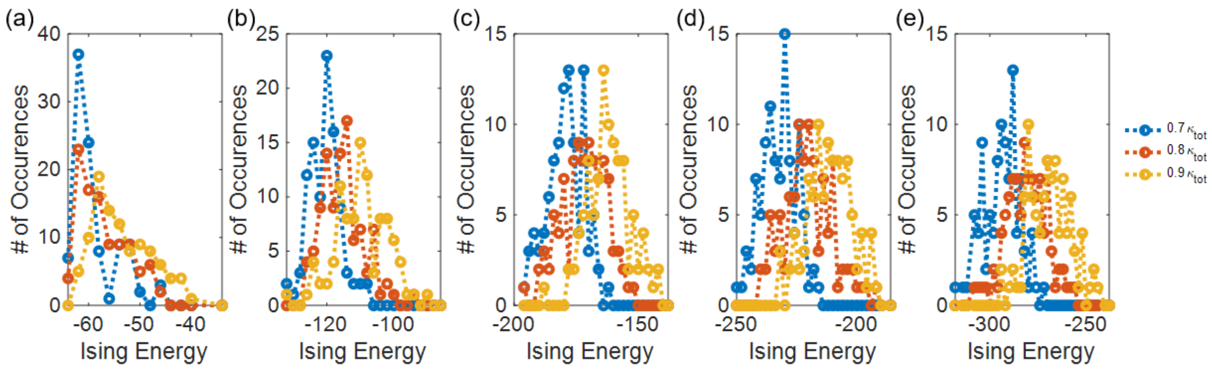

Figure S10 Numerically extracted energy values over 100 runs for a: **a)** 40-node 5-regular problem graph, **b)** 80-node 5-regular problem graph, **c)** 120-node 5-regular problem graph, **d)** 160-node 5-regular problem graph, and **e)** 200-node 5-regular problem graph solved by a PO-based IM. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2} = \mathbf{0.7}\boldsymbol{\kappa}, \mathbf{0.8}\boldsymbol{\kappa}$, and $\mathbf{0.9}\boldsymbol{\kappa}$.

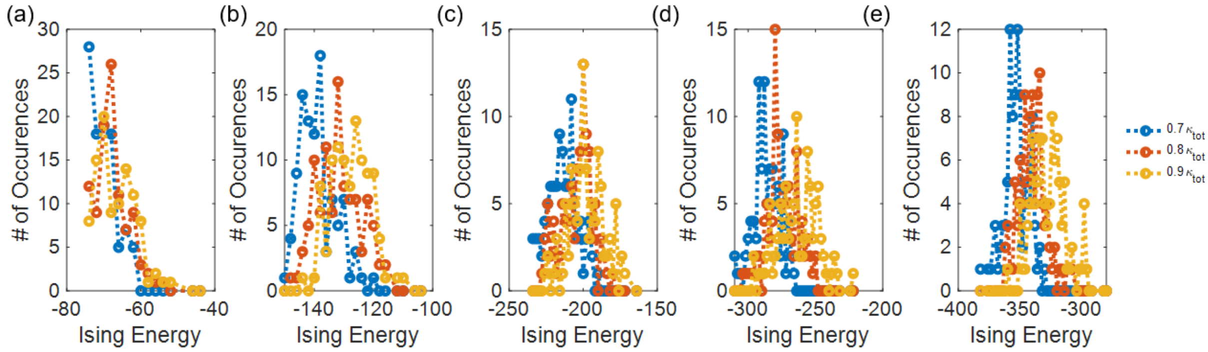

Figure S11 Numerically extracted energy values over 100 runs for a: **a)** 40-node 7-regular problem graph, **b)** 80-node 7-regular problem graph, **c)** 120-node 7-regular problem graph, **d)** 160-node 7-regular problem graph, and **e)** 200-node 7-regular problem graph solved by a PO-based IM. All these problem graphs are generated randomly, and each graph is solved with $\boldsymbol{r_2} = \mathbf{0.7}\boldsymbol{\kappa}, \mathbf{0.8}\boldsymbol{\kappa}$, and $\mathbf{0.9}\boldsymbol{\kappa}$.

References

[1] H. M. E. Hussein, S. Kim, M. Rinaldi, A. Alù, and C. Cassella, "Passive frequency comb generation at radiofrequency for ranging applications," *Nat. Commun.*, vol. 15, no. 1, p. 2844, Apr. 2024, doi: 10.1038/s41467-024-46940-2.

[2] C. Cassella *et al.*, "Ising Dynamics for Programmable Threshold Sensing in Wireless Devices," Jul. 01, 2024. doi: 10.21203/rs.3.rs-4530072/v1.

[3] N. Casilli, T. Kaisar, L. Colombo, S. Ghosh, P. X.-L. Feng, and C. Cassella, "Parametric Frequency Divider Based Ising Machines," *Phys. Rev. Lett.*, vol. 132, no. 14, p. 147301, Apr. 2024, doi: 10.1103/PhysRevLett.132.147301.

[4] M. Calvanese Strinati and C. Conti, "Multidimensional hyperspin machine," *Nat. Commun.*, vol. 13, no. 1, p. 7248, Nov. 2022, doi: 10.1038/s41467-022-34847-9.

[5] M. Calvanese Strinati, L. Bello, A. Pe'er, and E. G. Dalla Torre, "Theory of coupled parametric oscillators beyond coupled Ising spins," *Phys. Rev. A*, vol. 100, no. 2, p. 023835, Aug. 2019, doi: 10.1103/PhysRevA.100.023835.

[6] L. Q. English, A. V. Zampetaki, K. P. Kalinin, N. G. Berloff, and P. G. Kevrekidis, "An Ising machine based on networks of subharmonic electrical resonators," *Commun. Phys.*, vol. 5, no. 1, p. 333, Dec. 2022, doi: 10.1038/s42005-022-01111-x.

[7] T. L. Heugel, O. Zilberberg, C. Marty, R. Chitra, and A. Eichler, "Ising machines with strong bilinear coupling," *Phys. Rev. Res.*, vol. 4, no. 1, p. 013149, Feb. 2022, doi: 10.1103/PhysRevResearch.4.013149.

[8] Y. Yamamoto *et al.*, "Coherent Ising machines—optical neural networks operating at the quantum limit," *Npj Quantum Inf.*, vol. 3, no. 1, p. 49, Dec. 2017, doi: 10.1038/s41534-017-0048-9.

[9] R. Hamerly *et al.*, "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer," *Sci. Adv.*, vol. 5, no. 5, p. eaau0823, May 2019, doi: 10.1126/sciadv.aau0823.

[10] M. Honari-Latifpour and M.-A. Miri, "Optical Potts machine through networks of three-photon down-conversion oscillators," *Nanophotonics*, vol. 9, no. 13, pp. 4199–4205, Jul. 2020, doi: 10.1515/nanoph-2020-0256.

[11] M. Honari-Latifpour, M. S. Mills, and M.-A. Miri, "Combinatorial optimization with photonics-inspired clock models," *Commun. Phys.*, vol. 5, no. 1, p. 104, Apr. 2022, doi: 10.1038/s42005-022-00874-7.

[12] J. Roychowdhury, "A global Lyapunov function for the coherent Ising machine," *Nonlinear Theory Its Appl. IEICE*, vol. 13, no. 2, pp. 227–232, 2022, doi: 10.1587/nolta.13.227.

[13] M. Calvanese Strinati, I. Aharonovich, S. Ben-Ami, E. G. Dalla Torre, L. Bello, and A. Pe'er, "Coherent dynamics in frustrated coupled parametric oscillators," *New J. Phys.*, vol. 22, no. 8, p. 085005, Aug. 2020, doi: 10.1088/1367-2630/aba573.

[14] P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics*, 1st ed. Cambridge University Press, 1990. doi: 10.1017/CBO9781139167994.

[15] Q. Cen *et al.*, "Large-scale coherent Ising machine based on optoelectronic parametric oscillator," *Light Sci. Appl.*, vol. 11, no. 1, p. 333, Nov. 2022, doi: 10.1038/s41377-022-01013-1.

[16] T. Albash, V. Martin-Mayor, and I. Hen, "Analog errors in Ising machines," *Quantum Sci. Technol.*, vol. 4, no. 2, p. 02LT03, Apr. 2019, doi: 10.1088/2058-9565/ab13ea.

[17] M. Calvanese Strinati, L. Bello, E. G. Dalla Torre, and A. Pe'er, "Can Nonlinear Parametric Oscillators Solve Random Ising Models?," *Phys. Rev. Lett.*, vol. 126, no. 14, p. 143901, Apr. 2021, doi: 10.1103/PhysRevLett.126.143901.
[18] F. Böhm *et al.*, "Understanding dynamics of coherent Ising machines through simulation of large-scale 2D Ising models," *Nat. Commun.*, vol. 9, no. 1, p. 5020, Nov. 2018, doi: 10.1038/s41467-018-07328-1.