Enhancing the performance of coherent Ising machines in the large-noise regime with a fifth-order nonlinearity

RUQI SHI[1,*], FABIAN BÖHM[2], THOMAS VAN VAERENBERGH[3] AND PETER BIENSTMAN[1]

[1] *Photonics Research Group, Department of Information Technology, Ghent University-imec, Ghent, Belgium*
[2] *Hewlett Packard Labs, Boeblingen, Germany*
[3] *Hewlett Packard Labs, Diegem, Belgium*
[*] *Ruqi.Shi@UGent.be*

Abstract:

Coherent Ising Machines (CIMs), leveraging the bistable physical properties of coherent light to emulate Ising spins, exhibit great potential as hardware accelerators for tackling complex combinatorial optimization problems. Recent advances have demonstrated that the performance of CIMs can be enhanced either by incorporating large random noise or higher-order nonlinearities, yet their combined effects on CIM performance remain mainly unexplored. In this work, we develop a numerical CIM model that utilizes a tunable fifth-order polynomial nonlinear dynamic function under large noise levels, which has the potential to be implemented in all-optical platforms. We propose a normal form of a CIM model that allows for both supercritical and subcritical pitchfork bifurcation operational regimes, with fifth-order nonlinearity and tunable hyperparameters to control the Ising spin dynamics. In the benchmark studies, we simulate various sets of MaxCut problems using our fifth-order polynomial CIM model. The results show a significant performance improvement, achieving an average of 59.5% improvement in median time-to-solution (TTS) and an average of 6 times improvement in median success rate (SR) for dense Maxcut problems in the BiqMac library, compared to the commonly used third-order polynomial CIM model with low noise. The fifth-order polynomial CIM model in the large-noise regime also shows better performance trends as the problem size scales up. These findings reveal the enhancements on the computational performance of Ising machines in the large-nose regime from fifth-order nonlinearity, showing important implications for both simulation and hardware perspectives.

1. Introduction

As a versatile framework of both theoretical and practical significance, the Ising model is a cornerstone of statistical mechanics, offering valuable insights into the behavior of interacting particles in a variety of physical systems. Moreover, the cost function of many optimization problems can be mapped to the Ising model, such that the optimal solution corresponds to the Ising ground state [1]. This model, represented by an undirected graph, encompasses a set of binary Ising spin nodes ($x_i = \pm 1$) connected by Ising coupling interactions denoted as J_{ij}. The Ising Hamiltonian is defined as $H = -\sum_{i<j} J_{ij} x_i x_j$. The objective of the optimization problem usually lies in determining the spin configurations that minimize the Ising Hamiltonian for a given coupling matrix. Ising machines are physical systems designed to find low-energy solutions of the Ising model. Their core operational principle is to mimic an Ising model, by letting its state evolve to minimize the Ising Hamiltonian over time.

Ising machines have a diversity of hardware realizations, including quantum annealers [2, 3], electrical oscillator-based Ising solvers [4–8], polariton lattices [9], spatial Ising machines [10], and coherent Ising machines (CIMs) [11–16]. Since their invention, considerable attention has

been paid to CIMs as a promising next generation of hardware accelerators [17], designed to tackle hard combinatorial optimization problems. CIMs usually consist of a network of optical Ising nodes, offering fast and energy-efficient optimization capabilities for generic quadratic unconstrained binary optimization (QUBO) problems.

Various optical systems have been proposed to build CIMs, among which the most widely representative platform is the degenerate optical parametric oscillator (DOPO) [13, 15, 18]. The DOPO system maps the Ising spins into bistable states of coherent light pulses (e.g., 0 and π phases), or into the in-phase amplitude of the coherent light pulses [13]. Most commonly for the existing DOPO platforms, the following stochastic differential equation (SDE) can well approximate the in-phase amplitude dynamics in a third-order nonlinear dynamic form [11, 13, 15, 18]:

$$dx_i = [(r-1)x_i - \eta x_i^3 + \beta \sum_{j=1}^{N} J_{ij} x_j]dt + \gamma dW_t, i = 1, 2, \cdots, N. \tag{1}$$

Here, x_i represents the in-phase amplitude of the i^{th} light pulse and its sign represents the binary spin direction, while dt is the integration time step. N is the number of spins, r is the linear gain of the laser pump, the term -1 in $(r-1)$ is the normalized loss term, η is the third-order polynomial coefficient and β is the mutual coupling strength. γdW_t is the stochastic Gaussian noise, with γ the amplitude of the noise or noise level, and W_t standing for the Wiener process [19] with expectation value 0 and variance $\sqrt{dt}$. The SDE function consists of a nonlinear dynamic part $(r-1)x_i - \eta x_i^3$ that generates bistability for encoding the Ising node state from a third-order supercritical pitchfork bifurcation, and a linear coupling part $\beta \sum_{j=1}^{N} J_{ij} x_j$ that serves as the negative gradient to minimize the Ising Hamiltonian.

Incorporating higher-order interactions or nonlinearities in the spin dynamics may enhance the performance of an Ising machine, as suggested in a recent study [20]. Ising machine models based on sigmoidal, periodic, and clipped spin dynamic functions have been investigated, which bring higher-order interactions to the Ising node spin dynamic function. The third-order polynomial CIM model without any higher-order interactions has been found to be less powerful in computational performance compared with other Ising machine models that have higher-order nonlinearities. This is because a more distinct spin amplitude inhomogeneity exists in third-order models, which might lead to a faulty mapping of the original combinatorial optimization problems [20–23]. However, the previously studied nonlinearities are with fixed nonlinear terms that originate from currently known physical Ising systems, without further engineering to optimize the nonlinear terms. Inspired by the work in [24–28] with programmable nonlinearities in optical neural networks, we assume that a tunable nonlinear term also has the potential to further optimize the Ising machine.

A recent study in [29] has shown the fabricable possibility of an all-optical CIM based on a microrings. Here, the Ising nodes are realized with Kerr nonlinear rings, which naturally give rise to a tunable fifth-order nonlinearity in the spin dynamics. The incorporation of tunable higher-order terms can be understood intuitively as a result of the presence of bistable optical ring cavities and symmetry breaking in the two microring coupled MZI system, which leads to a higher-order nonlinear spin dynamic function. As stated in the supplementary materials, tunability of the higher-order nonlinearity can be achieved by controlling the operational optical parameters, e.g., the detuning of the ring resonance and the pump power. In this work, based on the proposed device, we present a generic Ising machine model with tunable higher-order nonlinear terms to engineer the nonlinearity, such that it optimizes the CIM's performance in optimization tasks. It is worth mentioning that the microring-based Ising machine in [29] is not unique in realizing a higher-order nonlinearity. Other optical CIM platforms also have the potential of incorporating higher-order nonlinearities, such as OEO-based Ising machines [20] or Ising machines using spatial light modulators (SLMs) [30, 31].

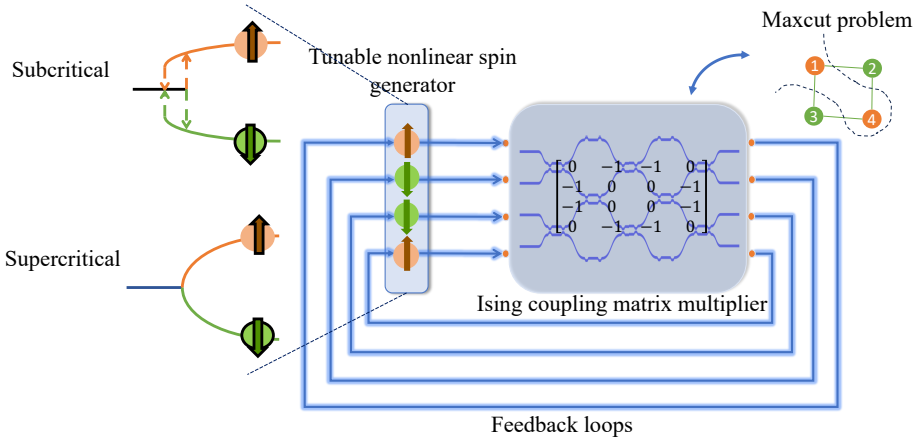


Fig. 1. Schematic of the proposed CIM model, allowing subcritical and supercritical operational regimes. There are three main building blocks: a tunable nonlinear spin generator that produces the nonlinear dynamics of the spins, an Ising coupling matrix multiplier that computes the multiplication of the spin vector and the Ising coupling matrix, and feedback loops that update the spin configuration. The building blocks can be built in photonic integrated circuits as stated in [29]. The tunability explanation can be seen in the supplementary materials.

A schematic illustration of the proposed CIM model can be found in Figure 1, which shows a 4-spin Ising system with spin configuration $[1, -1, -1, 1]$ at time t_0. The system encodes the Ising spins into the bistable states of the in-phase amplitude of light in nonlinear optical devices. This system has the potential to be energy-efficient when operated under ultra-low photon number limits. The nonlinear optical device in the system of [29] has an approximate spin dynamic function, derived from its optical properties, of the form:

$$dx_i = [(r-1)x_i - \eta x_i^3 + \zeta x_i^5 + \beta \sum_{j=1}^{N} J_{ij} x_j]dt + \gamma dW_t, i = 1, 2, \cdots, N. \tag{2}$$

Compared with Eq. (1), a fifth-order term ζx_i^5 is included in the dynamic function of the fifth-order CIM model. Notably, r, η, ζ, β and γ are tunable hyperparameters in our fifth-order CIM model. Uniquely, the fifth-order model enables operation in a regular supercritical regime in addition to a subcritical regime, which induces an additional hysteresis effect that can interact with the noise. The fifth-order model can be further engineered with the tunability of ζ and η.

The noise level parameter γ in Eq. (1) and Eq. (2) plays an important role in Ising machines. On one hand, to achieve minimal power consumption in the proposed system, which means a low photon number associated with each detected signal, it is crucial to consider the negative impact of the large noise inherent in the experimental setup when operating the Ising machine. The signal-to-noise ratio can be approximately 1 or even smaller when utilizing an ultra-low photon number as investigated in [32, 33]. On the other hand, however, the large intrinsic noise can also serve as a resource to enhance the computational performance of Ising machines. The importance of noise in Ising machines' performance has been reported in [34–36]. Indeed, for some hard optimization problems, injecting a substantial amount of noise is necessary for CIMs to find the ground state (GS) of hard optimization problems, as this can be beneficial in exploring the configuration space, particularly when the energy landscape has many local minima. Sufficient random noise which is uncorrelated to the current state of the Ising machine can help escape

from these large numbers of local minima. Therefore, finding the GSs requires a relatively large amount of noise in these circumstances. [35, 36] suggests that leveraging noise amplitude adequately can also facilitate and speed up the convergence to the GS.

However, High noise levels drive Ising machines into a thermal equilibrium, where the probability of reaching the ground state decays exponentially with the random noise level [10, 35, 36]. These conflicting effects need to be carefully balanced. In this sense, the incorporation of hysteresis in the fifth-order CIM spin dynamic function holds the promise of effectively mitigating the negative influence of noise, while still being able to benefit from the exploration it provides. Broadly speaking, hysteresis refers to the lag or delay in the response of a system to external changes in its input or operating conditions, and results in a "memory effect", where the state depends on its history. It has been commonly applied in electronic engineering [37, 38] and can also help in artificial neural networks to tolerate as well as control the noise [39–41]. As an example, the Schmitt trigger uses the hysteresis loop to help suppress noise and signal fluctuations. Recently, memristor-based Hopfield neural networks equipped with tunable hysteretic thresholds have been shown to be able to reduce the influence of random noise from experimental setups, and are able to perform simulated annealing [6]. These implementations are all in the digital domain or hybrid analog-digital domain, where using hysteresis to control the noise is a relatively mature technique. However, we anticipate that hysteresis can enhance the capability to mitigate the effects of large noise levels in analog CIM systems as well, particularly when the mean saturated analog spin amplitude and the Gaussian noise level are comparable in magnitude. This is one of the main points are be investigated in this paper.

To the best of our knowledge, previous studies of CIM models have neither systematically explored a tunable fifth-order nonlinearity in the spin dynamic function nor the subcritical pitchfork bifurcation regime with hysteresis. As suggested above, this could help to interact with the noise in the Ising system, facilitating the balance between the exploitation and exploration processes. In this paper, the subcritical pitchfork bifurcation regime model is compared with the supercritical regime model to reveal the performance enhancement from utilizing hysteresis. In order to build a general framework, we focus on the numerical computational performance of a fifth-order polynomial CIM model, without taking into account the original physical constraints of hyperparameters in the proposed hardware. We show that the fifth-order nonlinearity results in a promising enhancement in computational performance for the MaxCut benchmark sets BiqMac [42], some randomly-generated MaxCut problems, and Gset [12]. Additionally, we find that the subcritical regime with hysteresis is less sensitive to variations of hyperparameter settings. This is particularly important for accelerating hyperparameter optimization and addressing hardware imperfections.

Below is a brief overview of the paper's organization. In section 2, the steady-state analysis of the fifth-order polynomial CIM model in the supercritical and subcritical pitchfork regime is presented. In section 3, the optimization of the numerical setup that we use is discussed. In section 4, the benchmark results on MaxCut problems are presented and discussed. Finally, we conclude the findings from this research and discuss the future directions in section 5.

2. Steady state analysis of fifth-order polynomial CIM model

A third-order polynomial CIM model 1 can only be operated in a supercritical pitchfork regime to allow bistability for encoding the spins. A fifth-order polynomial CIM model however, depending on the sign of the third-order coefficient η in Eq. (2), can be operated in either a supercritical (i.e. $\eta > 0$) or a subcritical (i.e. $\eta < 0$) pitchfork regime [43]. In this section, we study the different bifurcation diagrams in these two operational regimes that the Ising system exhibits, characterized by different types of bifurcation points and stable branch shapes. These are important in practice, as they affect the dynamic evolution of the Ising system, leading to

diverse outcomes in performance.

2.0.1. Supercritical model ($\eta > 0$)

To visualize the differences in the supercritical and subcritical bifurcation diagrams, we focus on a single uncoupled Ising spin from the Ising system. An uncoupled-node bifurcation of the Ising node in the supercritical pitchfork bifurcation regime (i.e. $\eta > 0$) can be described by Eq. 2 without the coupling terms and the noise in the following ordinary differential equation (ODE):

$$\dot{x}_{uncoupled} = (r-1)x - \eta x^3 + \zeta x^5, \eta > 0. \tag{3}$$

The corresponding bifurcation diagram and corresponding potential energy curves are shown in Figure 2. For comparison, a third-order (i.e. $\zeta = 0$) supercritical pitchfork bifurcation has two branches of stable fixed points. A negative fifth-order coefficient ζ in Eq. (2) can modify the shape of the stable branch after reaching the same pitchfork bifurcation point as a third-order spin dynamic function. The stable fixed point as a function of the parameters can be calculated from Eq. 3 as:

$$x^2_{stable} = \frac{\eta - \sqrt{\eta^2 - 4(r-1)\zeta}}{2\zeta} \tag{4}$$

The fifth-order coefficient ζ can be fine-tuned as a hyperparameter in the CIM model to improve performance results, as we show in the benchmark studies in section 4.

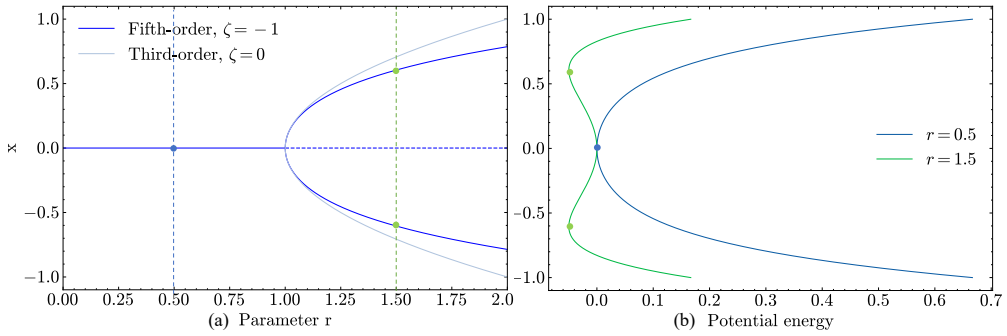


Fig. 2. (a): bifurcation diagram for $\dot{x} = (r-1)x - x^3 + \zeta x^5$ in the supercritical (i.e. $\eta > 0$) pitchfork regime. (b): potential energy curves corresponding to points below or above the supercritical pitchfork bifurcation. The potential energy is obtained by the integration of the ODE over x, providing a view for the stability.

To further illustrate the steady states at different bifurcation parameter values, the potential curves for different bifurcation parameters r are also shown in Figure 2. When r is below the pitchfork bifurcation point $r_p = 1$, e.g. $r = 0.5$, the minimum of the potential curve only appears at $x = 0$, indicating that the spin amplitude is only stable at $x = 0$. When r is above the pitchfork bifurcation point, e.g. $r = 1.5$, two minima appear at $x = \pm\sqrt{\frac{\eta - \sqrt{\eta^2 - 4(r-1)\zeta}}{2\zeta}}$, indicating that the spin amplitude is bistable at these values.

2.0.2. Subcritical model ($\eta < 0$)

In this section, we focus on the spin dynamics of the fifth-order CIM model in a subcritical ($\eta < 0$) pitchfork regime without coupling terms and noise:

$$\dot{x}_{uncoupled} = (r-1)x - \eta x^3 + \zeta x^5. \tag{5}$$

The bifurcation diagram and corresponding potential energy curves are shown in Figure 3. When $\eta < 0$ and $\zeta < 0$, as shown in Figure 3, a pair of saddle-node bifurcation points leads to the emergence of two stable steady states. In the third-order model ($\zeta = 0$), the absence of a non-trivial stable steady state precludes the existence of bistability for encoding Ising spins. Therefore, the fifth-order nonlinearity plays a stabilizing role in this scenario to encode the Ising spin states. The bifurcation diagram shows two symmetric hysteretic loops, delineated by the purple arrows. As hysteresis exists, the CIM model possesses the capability to ensure fewer fluctuations of spin states even with random noise. Moreover, it holds the potential to augment the uniformity of analog spin amplitudes because of a flatter amplitude trend compared with the supercritical regime.

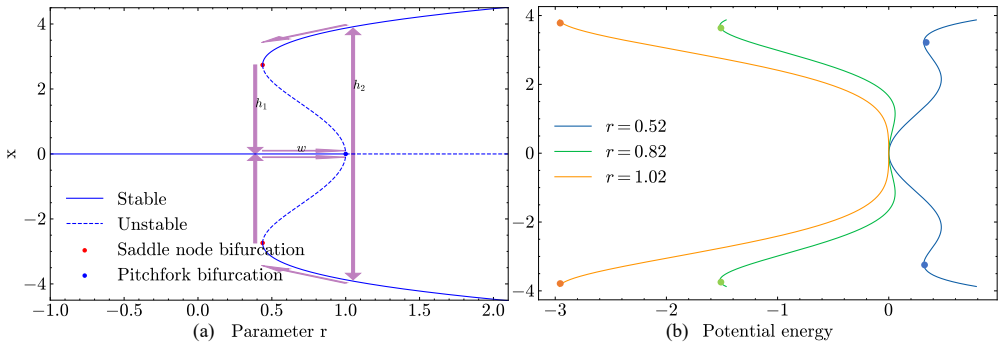


Fig. 3. (a): Bifurcation diagram for $\dot{x} = (r-1)x + 0.15x^3 - 0.01x^5$ in subcritical pitchfork regime ($\eta < 0$ and $\zeta < 0$). (b): Three potential energy curves corresponding to three r values below or above the subcritical pitchfork bifurcation.

The hysteretic loops are characterized by their width, represented by w, and their heights, denoted as h_1 and h_2. To derive these values, we observe that for the stable subcritical regime ($\eta < 0, \zeta < 0$) in Figure 3, the non-trivial stable branch shape can be obtained as:

$$x^2 = \frac{\eta + \sqrt{\eta^2 - 4(r-1)\zeta}}{-2\zeta}. \tag{6}$$

Here, r, η and ζ are coefficients from the fifth-order polynomial equation Eq. 2. The hysteretic width can be derived as $w = -\frac{\eta^2}{4\zeta}$, and the hysteretic heights are $h_1 = \sqrt{\frac{\eta}{2\zeta}}$, $h_2 = \sqrt{\frac{\eta}{\zeta}}$.

Apart from being potentially beneficial because of its stabilizing effect, hysteresis could be used to perform simulated annealing, by adjusting its parameters such as width w, height h_1, and h_2 as the annealing progresses. However, achieving precise control or measurement over these parameters in an analog Ising system with coupled nodes poses a significant challenge. Indeed, due to the inhomogeneity of analog spin amplitudes, the resulting values of w, h_1, and h_2 might deviate from node to node.

Finally, the corresponding potential curves for the stable subcritical regime with multiple bifurcation parameters r are shown in Figure 3: by controlling the bifurcation parameter r, the energy barriers between the energy minima and maxima are tuned, leading to different locations of the energy minima (stable fixed points).

3. Optimization setup

In this section, we present the techniques of linear annealing and the hyperparameter optimization workflow for optimizing the performance of the Ising machine setup.

3.1. Linear annealing scheme

The choice of the gain parameter is typically important in the performance of Ising machines [13, 44, 45]. In our simulations, we use a linear annealing scheme for the linear gain parameter r, as was shown to be efficient in [13, 45]:

$$r(t) = \begin{cases} r_0 + \frac{r_{end} - r_0}{T_{ann}} \times t & t < T_{ann} \\ r_{end} & t \geq T_{\text{ann}}. \end{cases} \tag{7}$$

Here, r_0 is an initial linear gain from which we linearly increase to the ending value r_{end} in the annealing time span T_{ann}. This scheme allows the Ising machine to explore both the regime near the bifurcation threshold and the regime high above it. If the annealing time span T_{ann} is sufficiently long, the linear gain can be considered slowly varying during one run. The Ising machine can thus explore different parameter settings, increasing the likelihood of achieving the optimal parameter setting that correctly maps the analog Ising energy to the real Ising Hamiltonian.

3.2. Hyperparameter optimization

To find the optimal combination of all seven hyperparameters (i.e. $r_0, r_{end}, \zeta, \eta, \beta, T_a, T_{run}$) for the polynomial CIM models, we follow a workflow based on Bayesian optimization [46] per set of problems with the same problem size and then grid search per instance:

1. Set the runtime T_{run} to a sufficiently large value, then do a rough grid search for the other six parameters in a large hyperparameter space to find a smaller space where the SR is non-zero (or where the distance to the best-known solution is under 10%, for Gset problems where the CIM models failed to reach the best-known solution).

2. Run Bayesian optimization on this smaller hyperparameter space, and find a near-optimal setting for all six hyperparameters to maximize the median SR per set of problems with the same problem size if it's non-zero or maximize the best-obtained cut value per instance.

3. With the other four hyperparameters fixed to the optimal values found by Bayesian optimization, run a two-dimensional scan per instance of what we found to be the most performance-relevant hyperparameters, i.e. linear gain r_{end} and coupling strength β.

4. (for the problems with non-zero SR after the above optimization process) Set the linear gain r_{end} and coupling strength β to the optimal values from the previous steps, and perform a one-dimensional scan per instance of T_{run} to find the optimal TTS.

Note that in this paper, the SR is obtained from 200 trials of simulation. The SR is defined as the transient success rate as in [20], which marks a trial as a success if the GS is reached once or more than once during the run time or integration time T_{run}, i.e.:

$$P_s(T_{run}) = \frac{\#Successful\_trials}{\#All\_trials} \tag{8}$$

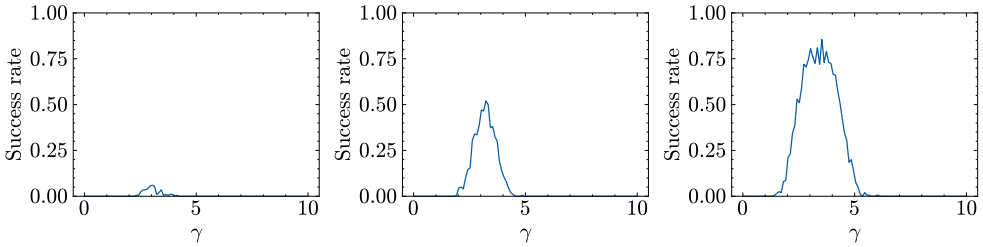
4. Benchmark studies on MaxCut problems

To evaluate the computational performance and scalability of the fifth-order polynomial CIM model with large noise, after the optimization of CIM models in section 3.2, we solve benchmark MaxCut problems from the BiqMac benchmark [42] of size $N = 60, 80, 100$ with 50% edge density, our own MaxCut tasks with randomly generated edges with $N = 30, 40, 50, 120, 150, 200, 300$ with 50% edge density, and Gset instances of $N = 800, 2000$ from the Suite Sparse Matrix Collection [12]. In this section, we first analyze the influence of the noise

level and show time-domain simulations to illustrate the effect of hysteresis. Then, we show the computational performance using the third-order polynomial Ising machine model with low noise as a baseline.

4.1. Noise level scan

We first examine the impact of random Gaussian noise on the performance of CIM models, to investigate whether large noise can improve performance. As mentioned before, some previous studies show that, when the noise intensity is orders of magnitude lower than the stable fixed point values of the Ising node amplitudes, the performance of a CIM is unlikely to be influenced by the random fluctuations stemming from the noise [20, 35]. Yet different studies [34, 35, 47] show that large noise improves the Ising machine performance by helping the escape from local minima, especially basin attractors, by reducing the autocorrelation time for the applications of Boltzmann sampling [47] and enlarging the operational area in hyperparameter spaces for hard optimization problems [34, 35]. To analyze the impact of noise on our CIM models, we conduct a noise level scan. Representative SR curves as a function noise level γ for different models, as depicted in Figure 4, offer insights into the enhanced performance resulting from hysteresis.



(a) Third-order supercritical model (b) Fifth-order supercritical model (c) Fifth-order subcritical model with hysteresis

Fig. 4. Representative SR as a function of noise level γ for BiqMac instance g05_100_4 for third-order and two fifth-order polynomial models. The fifth-order subcritical model with hysteresis achieves over 30% higher peak SR and encompasses an operational range of γ with a non-zero SR broader than the supercritical models. The other model parameters are optimized by the hyperparameter optimization workflow.

The SR curves in Figure 4 of all models show low values in the low-noise regime, followed by a rise to an optimal SR and then a drop back to zero. This can be understood because, in the low-noise regime, the Ising machine can be easily trapped in local minima of a non-convex energy landscape which is typical for the hard MaxCut instances. On the other hand, in the large-noise regime, the signal-to-noise ratio becomes too small such that the random spin flips induced by the noise prevent convergence to the ground state, causing a decrease in SR. Figure 4 also shows that the subcritical model outperforms the other models in achieving a higher optimal SR and a wider operational window. This suggests that the subcritical model is less sensitive to noise level variations, leading to improved robustness for different experimental conditions.

Note that the trends of SR curves of instance g_05_100.4 have been observed in 25 out of 30 instances in the BiqMac library simulation. The optimal noise of all three models per BiqMac instances can be seen in the supplementary materials. For all the models, the optimal noise for the maximal SR appears larger than 1 for most of the instances, indicating that the large noise can improve SR performance in BiqMac tasks.

4.2. *Time-domain simulations*

In order to further validate the hypothesis that the hysteresis in the subcritical regime can help reduce the fluctuations and make the Ising machine more robust to the deleterious effects of noise, we present the time-domain evolution of the spin amplitude and energy for BiqMac instance $g\_05\_100.6$, as a representative example.

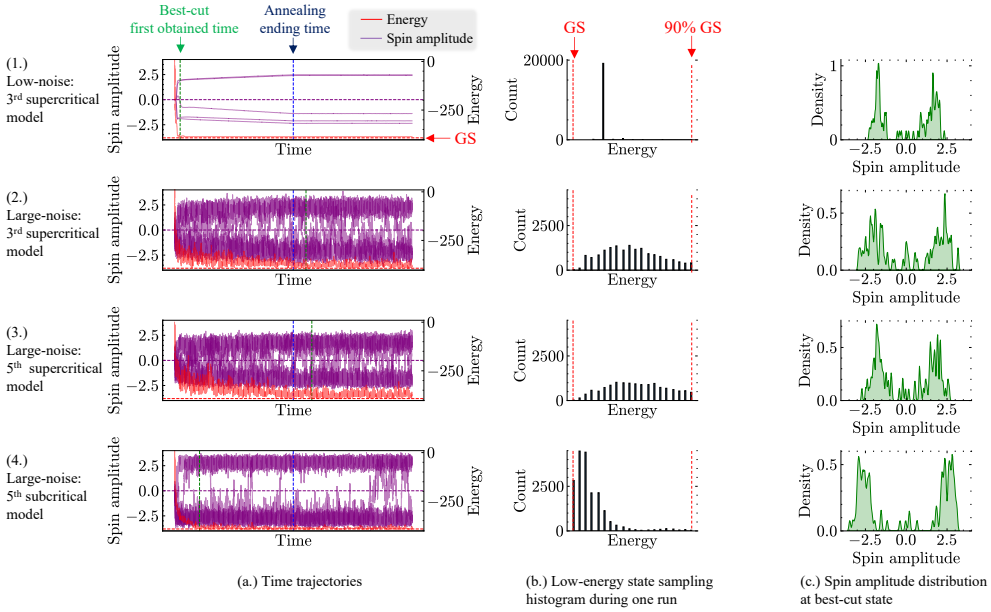


Fig. 5. Time trajectories and related statistical analysis of CIM models under low and large-noise profiles for BiqMac instance $g\_05\_100.6$ with $N = 100$ Ising nodes. Row (1.) shows the low-noise third-order supercritical model operated and optimized under $\gamma = 0.01$. Row (2-4) show the large-noise simulation operated and optimized under $\gamma = 2$ for the third-order supercritical model, fifth-order supercritical model, and fifth-order subcritical model, respectively. Column (a.) presents the evolution of the energy (red curve) and spin amplitude (purple curves) (only 5 randomly selected nodes out of 100 nodes are displayed for clarity). Column (b.) displays the histogram of the low-energy state samplings (from the GS energy to 90% GS energy) during one simulation. Column (c.) shows the spin amplitude density distribution of all 100 Ising nodes when the best cut is first obtained, demonstrating the analog heterogeneity of each model when reaching the optimal solution during one simulation.

Figure 5 displays time trajectories with corresponding statistics of energy and spin-amplitude distribution for different models. Row (1.) shows a low-noise simulation of the third-order model as a baseline. A large-noise third-order model is depicted in row (2.), a supercritical fifth-order model in row (3.), and finally a subcritical fifth-order model with hysteresis in row (4.).

The time trajectories in column (a.) show that the third-order model under a low-noise level exhibits clean time traces but quickly gets trapped in a local minimum. The large-noise simulation shows much noisier behavior in the spin dynamics, as expected. Between the three models, the subcritical model in the bottom row reaches the GS the fastest and generates fewer spin flips afterward, compared to the other two models. The supercritical model with the same fifth-order coefficient as the subcritical model (third row) behaves the worst in terms of solving speed and robustness against random spin flipping.

Turning our attention now to the histogram of the low-energy state samplings for the entire

runtime series in column (b.), we see that the third-order model under a low-noise level in the top row spends most of the time in a single low-energy state near the GS, but is unable to actually reach it. On the other hand, the high-noise subcritical model in the bottom row visits a large number of low-energy states during the simulation, indicating that it samples this region well, when looking for the GS. On the other hand, the remaining two models in the middle columns have a higher probability of sampling the less useful higher-energy states. This trend further supports the superiority of the subcritical model in terms of low-energy state sampling efficiency.

Finally, column (c.) illustrates the spin amplitude distribution for the state when the best cut is first obtained. The results suggest that the large-noise subcritical model has a smaller width of the spin amplitudes distribution for both negative and positive spin directions, with fewer amplitudes around the flipping threshold compared to the other two models, even better than the low-noise simulation of a third-order model.

In summary, the time domain simulations demonstrate that noise plays a crucial role in allowing the Ising system to escape from local minima. Under a large noise level, the fifth-order subcritical CIM model is able to reach the GS more quickly and has a higher probability of sampling the GS or low-energy states over time compared with the other polynomial CIM models. Additionally, the inhomogeneity of the analog amplitudes is also reduced, indicating that we can harness the positive exploration effect of noise, while not suffering from its negative effects in terms of amplitude noise.

4.3. Computational performance

TTS and SR are chosen to be the figures of merit in our work. The SR has been defined in section 3.2. The TTS in our simulations is defined in the same manner as in [48]:

$$TTS = \begin{cases} T_{run} \times \frac{\log(0.01)}{\log(1-P_s(T_{run}))} & 0 < P_s(T_{run}) \leq 0.99 \\ T_{run} & P_s(T_{run}) > 0.99, \end{cases} \tag{9}$$

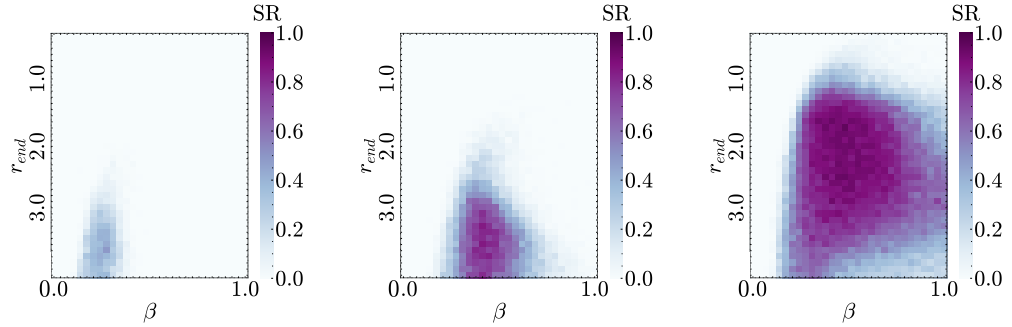
where $T_{run} = \#time\_interval \times \#time\_steps$ denotes the run time or integration time for one simulation trial, and $P_s(T_{run})$ is the SR for a certain run time as defined before. Note that due to a computation resource limitation for the simulation, the maximal $\#time\_steps$ is 25000 in our simulation. Thus TTS has an upper bound value of 55273.067, which means that when the true TTS is equal to or larger than 55273.067, the simulation will return this value.

4.3.1. BiqMac benchmark results ($N = 60 \cdots 100$)

In this section, we present the simulation results for the BiqMac benchmark problems, including an analysis of the sensitivity to hyperparameter variations. We also compare the TTS and SR performance for the third-order and fifth-order polynomial CIM models.

Sensitivity to hyperparameter variations

Figure 6 illustrates a representative two-dimensional scan of SR versus different combinations of (r_{end}, β) under large noise levels. The other hyperparameters are fixed at the best value for each nonlinear model. The subcritical fifth-order CIM model not only achieves a higher peak SR but also encompasses larger operational windows with high SR, compared with the other two models. This implies that the subcritical fifth-order CIM model can maintain high performance, even without fine-tuning the hyperparameters which can simplify the time-intensive parameter optimization. Besides, this signifies increased robustness against hyperparameter imperfection in real physical hardware. Additionally, we have performed simulations of adding spin-to-spin variance to the fifth-order term ζ, with a standard deviation $\sigma = 0.1\zeta$ on the hard MaxCut instance g05_100.6. We observe a minor impact on the SR for the supercritical model (SR changes from 49% to 45%) and subcritical model (SR changes from 96% to 95%). This indicates that the fifth-order models appear to be robust to spin-to-spin variance of the fifth-order coefficient.



(a) Third-order CIM model.

(b) Fifth-order CIM model: supercritical.

(c) Fifth-order CIM model: subcritical.

Fig. 6. SR for a 2-dimensional scan of different (β, r_{end}) settings for BiqMac problem g05_100.4, in the large-noise regime.

TTS and SR performance

Table 1 shows the median TTS and SR values for third-order and fifth-order supercritical models, as well as the fifth-order subcritical model. All results are obtained after hyperparameter scans and using a linear annealing scheme for the linear gain parameter r.

Benchmark sets	Third-order model in low noise		Third-order model in large noise		Fifth-order model supercritical in large noise		Fifth-order model subcritical in large noise	
	Median TTS	Median SR	Median TTS	Median SR	Median TTS	Median SR	Median TTS	Median SR
g_05_60	621.0	19.5%	602.2	77%	**365.8**	82%	384.7	**85%**
g_05_80	891.1	9.5%	354.9	**95%**	**352.7**	94%	464.9	**95%**
g_05_100	1511.0	16%	1352.5	78%	423.2	**96%**	**348.7**	91%

Table 1. BiqMac library MaxCut benchmark performance. The best performance among the models of SR and TTS for each problem set is highlighted in bold. The fifth-order models outperform the others in median TTS or median SR for the benchmark sets.

Table 1 reveals that compared with the commonly-used third-order CIM model with low noise, the other three large-noise models reach a much higher SR for the benchmark instances. The median SR improvement ranges from 65.5% to 85.5% for different N. Furthermore, compared with the supercritical models, the subcritical fifth-order models with hysteresis have a better median TTS as the graph size scales up to $N = 100$.

Figure 7 presents the TTS and SR for each problem in the different BiqMac benchmark sets. Overall, compared with the third-order low-noise model, the median performance of TTS and SR are improved for the large-noise models. The improvement is more distinct for hard instances (i.e., achieving less than 10% SR for the third-order CIM model in low noise) in the BiqMac library, which we indicate by having a red instance index. 9 out of 15 hard instances gain an

average SR increase of 67% for the large-noise models. However, the TTS improvement is less pronounced with a large-noise profile. This might be explained by the time trajectory results in Figure 5: the third-order model with a low-noise profile first reaches its best-cut much faster than models with a large-noise profile. Noise can serve as a random sampling in the energy landscape of the optimization problem, thus a large noise can improve the probability of reach the GS at the cost of taking longer time for the sampling process.

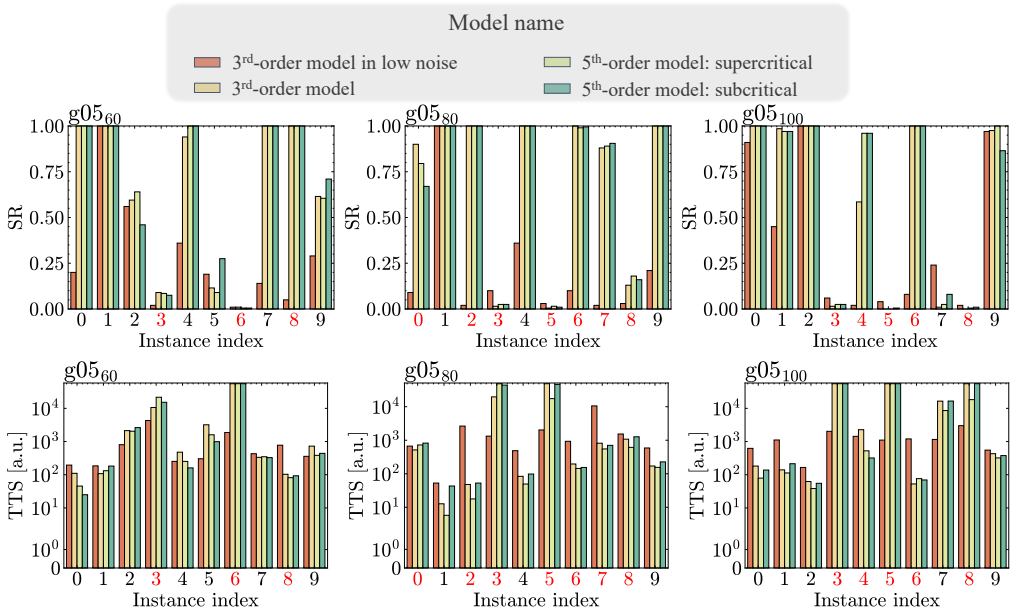


Fig. 7. TTS and SR of different models for instances from BiqMac benchmark sets (if not otherwise specified, the models are operated under large-noise levels). The indices of hard instances are highlighted in red.

4.3.2. Custom benchmark results ($N = 30 \cdots 300$)

To further show the scalability of the models, we augment the BiqMac set with our own custom MaxCut problems with sizes varying from $N = 30$ to 300, using randomly generated edges with 50% edge density. The models' results are computed after the optimization process outline above. The outcomes corresponding to TTS scalability analysis are presented in Figure 8. Although large noise can reduce the median TTS, it tends to increase the spread of TTS data. Also, if N increases, the spread of TTS data tends to increase for all models. However, between all large-noise models, the spread of TTS data from the fifth-order subcritical model is the smallest, at least for $N < 150$. Note that we are unable to accurately determine the spread for $N \geq 150$, due to the upper bound limit of TTS. Indeed, this upper bound in our simulations is 55273.067 due to a limited maximal integration time T_{run}. A TTS larger than the upper bound is considered to be equal to the upper bound itself. Note that for $N = 300$, only the fifth-order subcritical model achieves multiple TTS values below this limit, suggesting a promising TTS trend for even larger N.

4.3.3. Gset benchmark results ($N = 800, 2000$)

Finally, we also run experiments on eight selected Gset instances including $N = 800$ spins for graph G18, G19, G20, G21 and $N = 2000$ spins for graph G39, G40, G41, G42. These instances

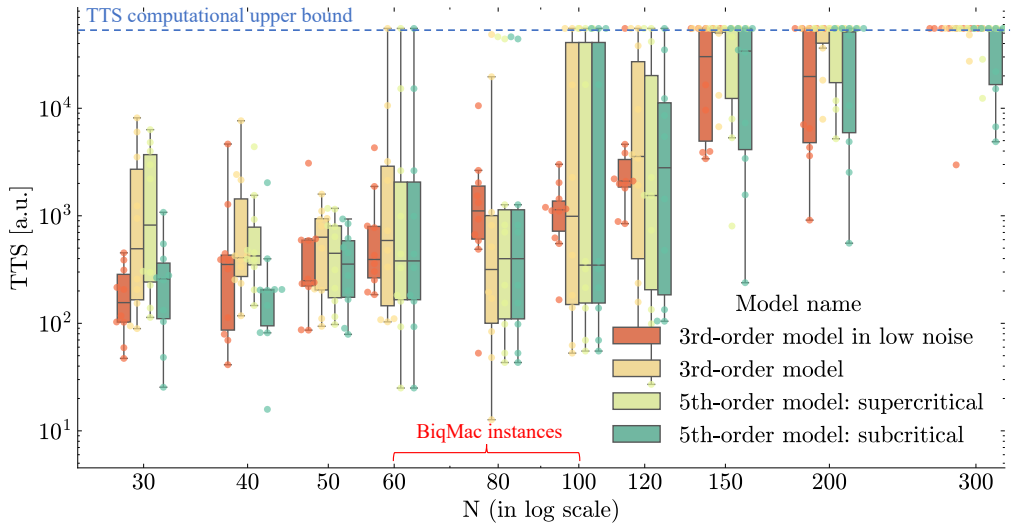


Fig. 8. TTS scaling curves for CIM models: each problem size N contains 10 problem instances. The upper bound of the TTS computation time is indicated by the dashed line, which is given by the maximal simulation time.

have been observed to be relatively hard for the third-order polynomial supercritical model to solve [20], i.e., they have a gap of a few percentages compared to the best-known solutions. Thus we choose the best-obtained cut per model, as the performance.

The Gset benchmark results can be seen in the table 2. The subcritical fifth-order model obtains a median improvement of 3.85% in the best cut for the eight hard instances in Gset library, compared with the supercritical third-order model, suggesting that the computational advantage of a fifth-order model remains, as the problem size scales up.

5. Conclusions

We develop fifth-order polynomial CIM models which can be operated either in the supercritical pitchfork regime or in the subcritical pitchfork regime with hysteresis in its bifurcation diagram. The fifth-order polynomial CIM models can be potentially implemented on integrated photonic hardware, which can be operated under ultra-low photon number regime with large intrinsic noise. It's important to note that the fifth-order nonlinearity in spin dynamics is derived from the microring-based photonic Ising node structure as in [29], which is built using materials with third-order nonlinearities such as Kerr and thermal nonlinearities, rather than materials with fifth-order nonlinearity. Thus, there is no need for additional power to drive the fifth-order nonlinearity and the power requirements for the fifth-order polynomial model should be similar to those for operating under the third-order polynomial model. We investigate the effect of the noise level on the SR and TTS for different MaxCut problems, and we find that a relatively large noise can be beneficial for improving the overall SR, and the SR improvement is more prominent for hard instances. We have observed that the subcritical model with hysteresis efficiently filters out undesired spin flipping after reaching low-energy states, as evidenced by the time trajectories. This suggests that hysteresis allows a better balance between exploitation and exploration under a large-noise profile, countering the issue of becoming ensnared in local minima and failing to reach a GS. Also, the large-noise subcritical model achieves enhanced robustness against hyperparameter variations. Besides, the fifth-order models hold the potential of being implemented on other Ising machines, e.g., by using physically programmable nonlinearities [24–28].

Gset instance index	N	Best-known cut [49]	Third-order model in low noise best obtained cut	Subcritical model in large noise best obtained cut	Improvement [%]
18	800	992	945	981	3.8
19	800	906	867	889	2.5
20	800	941	899	929	3.3
21	800	931	889	917	3.1
39	2000	2408	2233	2323	4.0
40	2000	2400	2224	2320	4.3
41	2000	2405	2234	2300	3.0
42	2000	2481	2304	2395	3.9

Table 2. Best obtained cut comparisons of third-order model in low noise and subcritical model. In this simulation, we obtained the results by performing Bayesian optimization for the model hyperparameters under low noise and large noise, respectively.

However, although the advantage in SR for hard instances is distinct, the results also show that large noise can yield a longer time to find the optimal configuration due to its sampling process and a larger spread in the SR and TTS performance, making a minority of instances harder to solve. The subcritical model can mitigate the issue relatively, but it still has a larger spread compared with the model operated under low noise. It's worth noting that the current optimization flow is not perfect and there remains potential space to improve the model either by extending the explored hyperparameter spaces, or by using more advanced hyperparameter optimization tools [50], for example. In the future, it might be necessary to further optimize the model parameters and investigate the working principle of the hysteresis in the subcritical model to address this phenomenon in order to ensure a more reliable performance. We will also focus on the physical realization of the proposed system. It is important to note that achieving full optimization of all hyperparameters may be challenging in hardware due to the extensive number of combinations required in the optimization process. However, the subcritical fifth-order model reduces parameter sensitivity, resulting in less need for parameter tuning and a significantly reduced overhead in parameter optimization.

Funding

This project was funded by the Belgian EOS project PINCH.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

Data availability

Data underlying the results presented in this paper can be obtained from the authors upon reasonable request.

Supplement

See Supplement 1 for supporting content.

References

1. A. Lucas, "Ising formulations of many np problems," Front. physics **2**, 5 (2014).
2. A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, "Quantum annealing: A new method for minimizing multidimensional functions," Chem. physics letters **219**, 343–348 (1994).
3. P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz *et al.*, "Architectural considerations in the design of a superconducting quantum annealing processor," IEEE Trans. on Appl. Supercond. **24**, 1–10 (2014).
4. M. Jiang, K. Shan, X. Sheng, C. Graves, J. P. Strachan, and C. Li, "An efficient synchronous-updating memristor-based ising solver for combinatorial optimization," in *2022 International Electron Devices Meeting (IEDM),* (IEEE, 2022), pp. 22–2.
5. T. Wang and J. Roychowdhury, "Oim: Oscillator-based ising machines for solving combinatorial optimisation problems," in *Unconventional Computation and Natural Computation: 18th International Conference, UCNC 2019, Tokyo, Japan, June 3–7, 2019, Proceedings 18,* (Springer, 2019), pp. 232–256.
6. 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**, 409–418 (2020).
7. S. Dutta, A. Khanna, A. Assoa, H. Paik, D. G. Schlom, Z. Toroczkai, A. Raychowdhury, and S. Datta, "An ising hamiltonian solver based on coupled stochastic phase-transition nano-oscillators," Nat. Electron. **4**, 502–512 (2021).
8. R. Afoakwa, Y. Zhang, U. K. R. Vengalam, Z. Ignjatovic, and M. Huang, "Brim: Bistable resistively-coupled ising machine," in *2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA),* (IEEE, 2021), pp. 749–760.
9. N. G. Berloff, M. Silva, K. Kalinin, A. Askitopoulos, J. D. Töpfer, P. Cilibrizzi, W. Langbein, and P. G. Lagoudakis, "Realizing the classical xy hamiltonian in polariton simulators," Nat. materials **16**, 1120–1126 (2017).
10. D. Pierangeli, G. Marcucci, and C. Conti, "Large-scale photonic ising machine by spatial light modulation," Phys. review letters **122**, 213902 (2019).
11. P. L. McMahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara *et al.*, "A fully programmable 100-spin coherent ising machine with all-to-all connections," Science **354**, 614–617 (2016).
12. S. Utsunomiya, K. Takata, and Y. Yamamoto, "Mapping of ising models onto injection-locked laser systems," Opt. express **19**, 18091–18108 (2011).
13. 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).
14. 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**, 937–942 (2014).
15. T. Inagaki, K. Inaba, R. Hamerly, K. Inoue, Y. Yamamoto, and H. Takesue, "Large-scale ising spin network based on degenerate optical parametric oscillators," Nat. Photonics **10**, 415–419 (2016).
16. T. Honjo, T. Sonobe, K. Inaba, T. Inagaki, T. Ikuta, Y. Yamada, T. Kazama, K. Enbutsu, T. Umeki, R. Kasahara *et al.*, "100,000-spin coherent ising machine," Sci. advances **7**, eabh0952 (2021).
17. N. Mohseni, P. L. McMahon, and T. Byrnes, "Ising machines as hardware solvers of combinatorial optimization problems," Nat. Rev. Phys. **4**, 363–379 (2022).
18. Y. Haribara, Y. Yamamoto, K.-i. Kawarabayashi, and S. Utsunomiya, "A coherent ising machine with quantum measurement and feedback control," arXiv preprint arXiv:1501.07030 (2015).
19. S. T. Huang and S. Cambanis, "Stochastic and multiple wiener integrals for gaussian processes," The Ann. Probab. pp. 585–614 (1978).
20. F. Böhm, T. V. Vaerenbergh, G. Verschaffelt, and G. Van der Sande, "Order-of-magnitude differences in computational performance of analog ising machines induced by the choice of nonlinearity," Commun. Phys. **4**, 149 (2021).
21. T. Leleu, Y. Yamamoto, S. Utsunomiya, and K. Aihara, "Combinatorial optimization using dynamical phase transitions in driven-dissipative systems," Phys. Rev. E **95**, 022118 (2017).
22. 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).
23. S. Kako, T. Leleu, Y. Inui, F. Khoyratee, S. Reifenstein, and Y. Yamamoto, "Coherent ising machines with error correction feedback," Adv. Quantum Technol. **3**, 2000045 (2020).
24. I. A. Williamson, T. W. Hughes, M. Minkov, B. Bartlett, S. Pai, and S. Fan, "Reprogrammable electro-optic nonlinear activation functions for optical neural networks," IEEE J. Sel. Top. Quantum Electron. **26**, 1–12 (2019).
25. A. Jha, C. Huang, and P. R. Prucnal, "Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics," Opt. letters **45**, 4819–4822 (2020).
26. J. R. R. Campo and D. Pérez-López, "Reconfigurable activation functions in integrated optical neural networks," IEEE J. Sel. Top. Quantum Electron. **28**, 1–13 (2022).

27. V. Nikkhah, M. J. Mencagli, and N. Engheta, "Reconfigurable nonlinear optical element using tunable couplers and inverse-designed structure," Nanophotonics (2023).
28. Y. Rah, Y. Jeong, S. Han, and K. Yu, "Low power coherent ising machine based on mechanical kerr nonlinearity," Phys. Rev. Lett. **130**, 073802 (2023).
29. N. Tezak, T. Van Vaerenbergh, J. S. Pelc, G. J. Mendoza, D. Kielpinski, H. Mabuchi, and R. G. Beausoleil, "Integrated coherent ising machines based on self-phase modulation in microring resonators," IEEE J. Sel. Top. Quantum Electron. **26**, 1–15 (2019).
30. M. C. Strinati, D. Pierangeli, and C. Conti, "All-optical scalable spatial coherent ising machine," Phys. Rev. Appl. **16**, 054022 (2021).
31. S. Kumar, H. Zhang, and Y.-P. Huang, "Large-scale ising emulation with four body interaction and all-to-all connections," Commun. Phys. **3**, 108 (2020).
32. C. Santori, J. S. Pelc, R. G. Beausoleil, N. Tezak, R. Hamerly, and H. Mabuchi, "Quantum noise in large-scale coherent nonlinear photonic circuits," Phys. Rev. Appl. **1**, 054005 (2014).
33. S.-Y. Ma, T. Wang, J. Laydevant, L. G. Wright, and P. L. McMahon, "Quantum-noise-limited optical neural networks operating at a few quanta per activation," arXiv preprint arXiv:2307.15712 (2023).
34. J. Roychowdhury, "A global lyapunov function for the coherent ising machine," Nonlinear Theory Its Appl. IEICE **13**, 227–232 (2022).
35. C. Roques-Carmes, Y. Shen, C. Zanoci, M. Prabhu, F. Atieh, L. Jing, T. Dubček, C. Mao, M. R. Johnson, V. Čeperić *et al.*, "Heuristic recurrent algorithms for photonic ising machines," Nat. communications **11**, 249 (2020).
36. D. Pierangeli, G. Marcucci, D. Brunner, and C. Conti, "Noise-enhanced spatial-photonic ising machine," Nanophotonics **9**, 4109–4116 (2020).
37. I. Filanovsky and H. Baltes, "Cmos schmitt trigger design," IEEE Trans. on Circuits Syst. I: Fundam. Theory Appl. **41**, 46–49 (1994).
38. A. Pfister, "Novel cmos schmitt trigger with controllable hysteresis," Electron. Lett. **7**, 639–641 (1992).
39. J. D. Keeler, E. E. Pichler, and J. Ross, "Noise in neural networks: thresholds, hysteresis, and neuromodulation of signal-to-noise." Proc. National Acad. Sci. **86**, 1712–1716 (1989).
40. S. Bharitkar and J. M. Mendel, "The hysteretic hopfield neural network," IEEE Trans. on neural networks **11**, 879–888 (2000).
41. M. Sun, L. Zhao, W. Cao, Y. Xu, X. Dai, and X. Wang, "Novel hysteretic noisy chaotic neural network for broadcast scheduling problems in packet radio networks," IEEE Trans. on Neural Networks **21**, 1422–1433 (2010).
42. F. Rendl, G. Rinaldi, and A. Wiegele, "A branch and bound algorithm for max-cut based on combining semidefinite and polyhedral relaxations," in *IPCO,* vol. 4513 (Springer, 2007), pp. 295–309.
43. S. H. Strogatz, *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering* (CRC press, 2018).
44. F. Böhm, T. Inagaki, K. Inaba, T. Honjo, K. Enbutsu, T. Umeki, R. Kasahara, and H. Takesue, "Understanding dynamics of coherent ising machines through simulation of large-scale 2d ising models," Nat. communications **9**, 5020 (2018).
45. S. Kiesewetter and P. D. Drummond, "Coherent ising machine with quantum feedback: The total and conditional master equation methods," Phys. Rev. A **106**, 022409 (2022).
46. P. I. Frazier, "A tutorial on bayesian optimization," arXiv preprint arXiv:1807.02811 (2018).
47. F. Böhm, D. Alonso-Urquijo, G. Verschaffelt, and G. Van der Sande, "Noise-injected analog ising machines enable ultrafast statistical sampling and machine learning," Nat. Commun. **13**, 5847 (2022).
48. 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," Sci. Adv. **7**, eabe7953 (2021).
49. F. Ma and J.-K. Hao, "A multiple search operator heuristic for the max-k-cut problem," Ann. Oper. Res. **248**, 365–403 (2017).
50. M. Parizy, N. Kakuko, and N. Togawa, "Fast hyperparameter tuning for ising machines," in *2023 IEEE International Conference on Consumer Electronics (ICCE),* (IEEE, 2023), pp. 1–6.