

Effective implementation of l_0 -Regularised Compressed Sensing with Chaotic-Amplitude-Controlled Coherent Ising Machines

Mastiyage Don Sudeera Hasaranga Gunathilaka^{1*}, Satoshi Kako², Yoshitaka Inui², Kazushi Mimura^{1,4}, Masato Okada⁵, Yoshihisa Yamamoto^{2,3} and Toru Aonishi¹

^{1*}School of Computing, Tokyo Institute of Technology,
Yokohama, Kanagawa, Japan.

²Physics and Informatics Laboratories, NTT Research Inc., 940
Stewart Dr, Sunnyvale, CA, 94085, USA.

³E. L. Ginzton Laboratory, Stanford University, Stanford, CA,
94305, USA.

⁴Graduate School of Information Sciences, Hiroshima City
University, Hiroshima, Japan.

⁵Graduate School of Frontier Sciences, The University of Tokyo,
Kashiwa, Chiba, Japan.

*Corresponding author(s). E-mail(s):
mastiyage.s.aa@m.titech.ac.jp;

Abstract

Coherent Ising Machine (CIM) is a network of optical parametric oscillators that can solve large-scale combinatorial optimisation problems by finding the ground state of an Ising Hamiltonian. As a practical application of CIM, Aonishi *et al.*, proposed a quantum-classical hybrid system to solve optimisation problems of l_0 -regularisation-based compressed sensing. In the hybrid system, the CIM was an open-loop system without an amplitude control feedback loop. In this case, the hybrid system is enhanced by using a closed-loop CIM to achieve chaotic behaviour around the target amplitude, which would enable escaping from local minima in the energy landscape. Both artificial and magnetic resonance image

data were used for the testing of our proposed closed-loop system. Compared with the open-loop system, the results of this study demonstrate an improved degree of accuracy and a wider range of effectiveness.

Keywords: Coherent Ising machine, Compressed Sensing, Bayesian inference, Magnetic Resonance Imaging, Quantum-Classical hybrid system, LASSO, Zeeman term, Chaotic-Amplitude Control, Gaussian-approximation, Closed-loop system, Combinatorial optimisation

1 Introduction

Compressed sensing (CS) is a method of reconstructing a high-dimensional signal or image based on highly downsampled measurements.

There has been considerable interest in it across a wide range of fields and applications. Such as in the field of astronomy, a possible way to transmit data to Earth from spacecraft [1] has been attempted. And there are proposed methods with CS on astronomical image compression and in compression on remotely sensed data [2–4] as well. And in radar technologies for the reconstruction of the target image CS has been used [5]. On the other hand in the medical field using embedded compression using CS to improve energy efficiency in Electrocardiogram (ECG) machines has been proposed [6].

$$\hat{x} = \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \|x\|_p \quad \text{subject to } y = Ax. \quad (1)$$

The above equation shows an observed signal $y \in \mathbb{R}^M$, an observation matrix $A \in \mathbb{R}^{M \times N}$, and a source signal $x \in \mathbb{R}^N$. Hereafter, the ratio of the number of non-zero entries in x to N is defined as the sparseness a , and the ratio of M to N is defined as the compression ratio α . Since l_1 -norm CS is a convex optimisation problem, there are many efficient algorithms for optimisation of l_1 -norm CS that are widely applied in the real-world problems mentioned above. However, there has been a suggestion that l_0 -norm CS should outperform l_1 -norm CS since the l_1 -norm penalty does not lead to any solution shrinkage [7, 8]. In the thermodynamic limit $N, M \rightarrow \infty$ with $\alpha = M/N$ kept fixed, an l_0 -norm CS's threshold for a , determining whether or not the problem has a solution with no error, is larger than that of l_1 -norm CS's [7, 8]. Nonetheless, the optimisation in l_0 -norm CS is challenging since it involves combinatorial optimisation.

Numerous attempts have been made to overcome the issue in l_0 -norm CS optimisations. l_0 -norm CS can be formulated as a two-fold optimisation [9, 10].

$$(\hat{R}, \hat{\sigma}) = \underset{\sigma \in \{0,1\}^N}{\operatorname{argmin}} \underset{R \in \mathbb{R}^N}{\operatorname{argmin}} (\|y - A(\sigma \circ R)\|_2^2) \quad \text{subject to } \|\sigma\|_0 \leq \Omega. \quad (2)$$

Here $R \in \mathbb{R}^N$ and $\sigma \in \{0, 1\}^N$ correspond to the source signal and support vector, respectively. Especially, each entry in the support vector taking either 0 or 1 represents whether each entry in the source signal is zero or non-zero. The condition $\|\sigma\|_0 \leq \Omega$ is a sparsity-inducing prior for constraining the number of non-zero entries to be Ω . Therefore, the optimisation with respect to σ can be regarded as a quadratic-constrained binary optimisation problem to find a ground state of a two-state Potts Hamiltonian. Based on this formulation, simulated annealing (SA) algorithm has been attempted [7]. On the other hand, Aonishi *et al.*, attempted to solve optimisation problems of l_0 -norm CS with a quantum-classical hybrid approach. l_0 -norm CS implemented with the hybrid system is given as a regularisation form as follows [11].

$$(R, \sigma) = \underset{\sigma \in \{0, 1\}^N}{\operatorname{argmin}} \underset{R \in \mathbb{R}^N}{\operatorname{argmin}} \left(\frac{1}{2} \|y - A(\sigma \circ R)\|_2^2 + \lambda \|\sigma\|_0 \right). \quad (3)$$

The element-wise representation of Eq. (3) gives the following Hamiltonian.

$$\mathcal{H} = \sum_{r < r'}^N \sum_{k=1}^M A_r^k A_{r'}^k R_r R_{r'} \sigma_r \sigma_{r'} - \sum_{r=1}^N \sum_{k=1}^M y^k A_r^k R_r \sigma_r + \lambda \sum_{r=1}^N \sigma_r, \quad (4)$$

where an element A^k in A , an element y^k in y , an element R_r in R and an element σ_r in σ . Optimisation with respect to σ in Eq. (4) is a quadratic unconstrained binary optimisation (QUBO) problem, which is implementable with a quantum machine such as the coherent Ising machine (CIM) [11–14]. In the quantum-classical hybrid approach to conducting l_0 -regularised CS, σ is optimised by the CIM while R is optimised by a Classical Digital Processor (CDP) (see Fig. 1).

The CIM architecture in the hybrid approach was an open-loop (OL) CIM with the Zeeman term. The hybrid approach with the OL-CIM is hereafter referred to as OL-CIM-CDP. Note that the OL means the lack of feedback loop for amplitude control described below. It has been reported that the imbalance in the size of the interaction term and the Zeeman term degrades the system performance [15]. To balance these terms, for the local field, the measured-amplitudes were binarised. OL-CIM-CDP in this formulation outperformed SA on the regularisation form [11].

The close-loop CIM, in which the amplitudes of optical parametric oscillator (OPO) pulses are controlled to a target value, have been proposed to improve the performance of CIM's ground-state search [16, 17]. Especially, introducing auxiliary nonlinear dynamics forcefully trying to equalise to a target value results in chaotic behaviour around the target in the CIM which may result in escaping from local minima in the energy landscape. This chaotic method is referred to as chaotic amplitude control (CAC) [16–20]. Recently, Inui *et al.*, have proposed an approach to efficiently incorporate the Zeeman terms in CAC-CIM by scaling the Zeeman terms with target amplitude to match that of the interaction term [17].

In this paper, following Inui *et al.*'s approach, we modify the CAC-CIM for performing QUBO in l_0 -regularised CS and attempt to improve the performance of the hybrid CIM-CDP system by replacing the OL-CIM with the CAC-CIM with the Zeeman term (see Fig. 1). The hybrid system proposed here is hereafter referred to as CAC-CIM-CDP. Firstly, to demonstrate the effectiveness of CAC-CIM for performing QUBO in the support estimation, we compare the performance of CAC-CIM to those of OL-CIM and SA. Then, to demonstrate the effectiveness of CAC-CIM-CDP for performing an alternating minimisation, we compare the performance of CAC-CIM-CDP to that of OL-CIM-CDP on artificial random data, as well as magnetic resonance imaging (MRI) data.

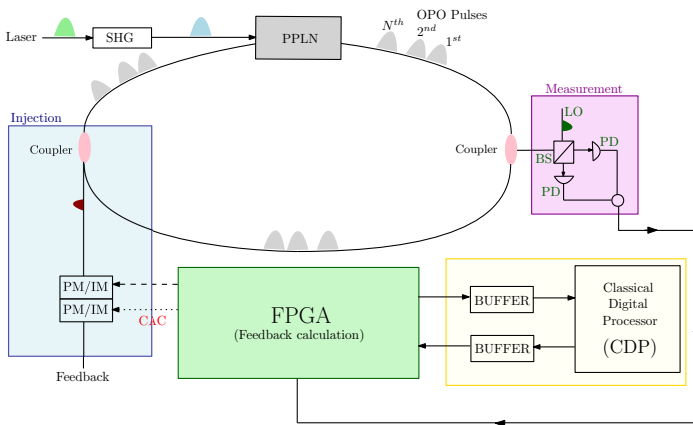


Fig. 1: CIM-CDP Architecture

Outline of the system architecture of the the feedback signal including CAC-loop is calculated in FPGA and is fed into the main ring cavity through a coupler. In this hybrid system, CIM optimises the support vector and CDP estimates the source signal in an alternating way. Without the CAC-loop, the architecture corresponds to the OL-CIM-CDP while with the CAC-loop, it is the CAC-CIM-CDP. SHG: second harmonic generation, PPLN: periodically poled lithium niobate, BS: beam splitter, PD: photon detector, PM/IM: phase modulator/intensity modulator, LO: local oscillator.

2 Results

2.1 Alternating minimisation algorithm

Alternating minimisation procedures on CAC-CIM-CDP and OL-CIM-CDP are summarised in Algorithm 1 and Algorithm 2, respectively. This type of minimisation suggests the back-and-forth optimisation performed between the CIM and CDP. CIM passes the optimisation results to the CDP after optimising the support, as shown in Fig. 1. The CDP then optimises the signal and

sends the resulting signal to the CIM for support optimisation. In Algorithm 1 and Algorithm 2, indicate the number of iterations of alternating minimisation, the initial values and the integration interval for stochastic differential equations (SDEs) of CIM and so on. The schedules of the pump rate, threshold and target amplitude are given in section 4.3.

Algorithm 1 Alternating minimisation for l_0 -regularised CS as a QUBO problem on CAC-CIM-CDP. The schedules of the pump rate and threshold are given in Section 4.3

Require: $M \times N$ observation matrix: A , M -dimensional observation signal: y ;

Ensure: N -dimensional support vector: σ , N -dimensional signal vector: r ;

1: Initialise $r = r_{init}$, $\eta = \eta_{init}$, $g^2 = 10^{-7}$, $K = 1$ and $\tau = 1$

2: **for** $i = 0$ to 51 **do**

3: Minimise \mathcal{H} with respect to σ by CIM:

4: $\sigma = \text{CIM-support-estimation}(R, \eta)$

 Initialise $\tilde{\mu} = 0$, $V = 0.5$ and $e = 1$ for CAC-CIM-CDP (Wigner) and

$\tilde{\mu} = 0$, $n = 0$, $m = 0$ and $e = 1$ for CAC-CIM-CDP (Positive- P).

 And we increase the photon's lifetime 20 times.

5: Minimise \mathcal{H} with respect to R by CDP using Conjugate Gradient Descent or Jacobi method:

6: Update η

7: **end for**

2.2 Outline of the CIM models and injection field for QUBO on support estimation

On CIM, l_0 -regularised CS is performed by updating the injection field dictated by the local field, which is determined by the gradient of the QUBO Hamiltonian Eq. (4) with respect to the spin coordinates. Aonishi *et al.*, proposed OL-CIM-CDP, which is based on an open-loop injection scheme [11]. They used the CIM model expressed as the Wigner stochastic differential equation (W-SDE) Eq. (13) and Eq. (14) (in Methods) with the following injection field.

$$\left(\frac{dc_r}{dt}\right)_{inj,r} = (|h_r| - \eta). \quad (5)$$

$$h_r = - \sum_{r'=1(\neq r)}^N \sum_{k=1}^M A_r^k A_{r'}^k R_{r'} H(c_{r'}) + \sum_{k=1}^M A_r^k y^k, \quad (6)$$

Here, h_r is the local field expressed as Eq. (6). R_r is the signal value estimated by the CDP. c_r is the in-phase amplitude of the r -th OPO pulse,

Algorithm 2 Alternating minimisation on OL-CIM-CDP. The schedules of the pump rate and threshold are given in Section 4.3

Require: $M \times N$ observation matrix: A , M -dimensional observation signal: y ;

Ensure: N -dimensional support vector: σ , N -dimensional signal vector: r ;

- 1: Initialise $\tilde{K} = 0.25$, $r = r_{init}$ and $\eta = \eta_{init}$
 - 2: **for** $i = 0$ to 51 **do**
 - 3: Minimise \mathcal{H} with respect to σ by CIM:
 - 4: $\sigma = \text{CIM-support-estimation}(R, \eta)$
 Initialise the in-phase amplitude as $c = 0$, and numerically integrate the W-SDE while increasing the normalised pump rate from 0 to 1.5 for five times the photon's lifetime when $g^2 = 10^{-7}$.
 - 5: Minimise \mathcal{H} with respect to R by CDP using Conjugate Gradient Descent or Jacobi method:
 - 6: Update η
 - 7: **end for**
-

and $H(c_r)$ is the binarised in-phase amplitude by the Heaviside step function as proposed in the discrete simulated bifurcation [21]. η is the threshold which is related to the l_0 -regularisation parameter λ by $\eta = \sqrt{2\lambda}$ according to the Maxwell rule (see [11] for a detailed explanation). In the local field Eq. (6), the mutual interaction is $\tilde{J}_{rr'} = -\sum_{k=1}^M A_r^k A_{r'}^k$ and the Zeeman term is $\sum_{k=1}^M A_r^k y^k$. Substituting the observation model Eq. (26) (in Section 4.4) into Eq. (6) when $w_{noise} = 0$ (no observation noise), the local field Eq. (6) can be expressed as follows.

$$h_r = - \sum_{r'=1(\neq r)}^N \sum_{k=1}^M A_r^k A_{r'}^k R_{r'} H(c_{r'}) + \sum_{r'=1}^N \sum_{k=1}^M A_r^k A_{r'}^k x_{r'} \xi_{r'}, \quad (7)$$

where x_r is the true signal value, ξ_r is the true support taking 1 or 0. The Zeeman term in the second term of Eq. (7) can be regarded as the matched filter, in which $A^T A$ is calculated. The mutual interaction term in the first term plays a role in removing off-diagonal elements ($r \neq r'$) corresponding to cross-talk noise in the Zeeman term, which are induced by the cross-correlation among the column vectors A_1, \dots, A_N in A . To obliterate the cross-talk noise, the in-phase amplitude c_r needs to be the same as the amplitude of ξ_r if $R_r = x_r$. Hence, c_r is binarised to either 1 or 0. In Fig. 2e, a typical evolution of c_r in the open-loop-type W-SDE is illustrated. c_r does not keep the same amplitude as that of ξ_r and increases with increasing the pump rate.

In this paper, we propose CAC-CIM-CDP, based on a closed-loop injection scheme with CAC. The idea of CAC for CIM was first introduced by Leleu *et al.*, [18]. It simply states that forcefully trying to equalise the amplitudes of the system to a specific value (in CAC, target amplitude τ) may result in a chaotic behaviour in the system which may result in escaping from local minima in

the energy landscape. In this paper, we used two CIM models expressed as W-SDE Eq. (15) and (16) and Positive- P stochastic differential equation (P-SDE) Eq. (17)-(19) (in Section 4.1) commonly having the following injection field with CAC feedback.

$$\left(\frac{d\mu_r}{dt}\right)_{inj,r} = j e_r \left(R_r h_r - \frac{\eta^2}{4} \sqrt{\frac{\tau}{g^2}} \right), \quad (8)$$

$$\frac{d}{dt} e_r = -\beta (g^2 \tilde{\mu}_r^2 - \tau) e_r, \quad (9)$$

$$\tilde{\mu}_r = \mu_r + \sqrt{\frac{1}{4j}} W_{R,r}, \quad (10)$$

$$h_r = - \sum_{r'=1(\neq r)}^N \sum_{k=1}^M A_r^k A_{r'}^k R_{r'} \frac{1}{2} \left(\tilde{\mu}_{r'} + \sqrt{\frac{\tau}{g^2}} \right) + \sum_{k=1}^M \sqrt{\frac{\tau}{g^2}} A_r^k y^k, \quad (11)$$

where h_r is the local field expressed as Eq. (11), e_r is the auxiliary variable for the error feedback in the CAC feedback loop, and τ indicates the target amplitude for the CAC. R_r is the signal value estimated by the CDP, which is the same as that of OL-CIM-CDP. η is the threshold given by $\eta = \sqrt{2\lambda}$, which is introduced to keep consistency with OL-CIM-CDP. As described in Section 4.1 in Methods, j is the normalised out-coupling rate for optical homodyne measurement, and g^2 is the nonlinear saturation parameter of the CIM which determines the abrupt jump of the photon number at the OPO threshold and the amplitude of the quantum noise present in CIM. $\tilde{\mu}_r$ implies the measured-amplitude, and $w(R, r)$ is the independent real Gaussian noise process, which is the same as that in W-SDE (15) and P-SDE (17). In the local field Eq. (11), the mutual interaction is $\tilde{J}_{rr'} = -\sum_{k=1}^M A_r^k A_{r'}^k$ and the Zeeman term is $h_r^z = \sqrt{\tau/g^2} \sum_{k=1}^M A_r^k y^k$. Substituting the observation model Eq. (26) into Eq. (11) when $w_{noise} = 0$ (no observation noise), the local field Eq. (11) can be expressed as follows.

$$h_r = - \sum_{r'=1(\neq r)}^N \sum_{k=1}^M A_r^k A_{r'}^k R_{r'} \frac{1}{2} \left(\tilde{\mu}_{r'} + \sqrt{\frac{\tau}{g^2}} \right) + \sum_{r'=1}^N \sum_{k=1}^M \sqrt{\frac{\tau}{g^2}} A_r^k A_{r'}^k x_{r'} \xi_{r'}. \quad (12)$$

In Fig. 2a and 2b, the typical evolution of normalised measured-amplitude $g\tilde{\mu}_r$ are shown. The corresponding error evolution is indicated in Fig. 2c and 2d. Due to the CAC feedback loop, as shown in Fig. 2a and 2b, if the squared-amplitude of DOPO is smaller than τ , e_r exponentially increases and vice-versa, and the measured-amplitude $\tilde{\mu}_{r'}$ is maintained around $\sqrt{\tau/g^2}$. Therefore, because $1/2(\tilde{\mu}_{r'} + \sqrt{\tau/g^2})$ in Eq. (12) can take around 0 or $\sqrt{\tau/g^2}$, the mutual interaction term and the Zeeman term scales are balanced, and crosstalk noise, i.e. off-diagonal elements, is eliminated from the Zeeman term as described in OL-CIM-CDP. Moreover, as shown in Fig. 2a and 2b, it is important to note that intermediate solutions are destabilised. By doing so, CAC introduced CIM

is able to keep searching for an answer until the maximum run-time has been reached. By taking the support vector that is generated by CIM at the end of each trajectory, we are evaluating the solution to estimate the support for the simulations in this paper.

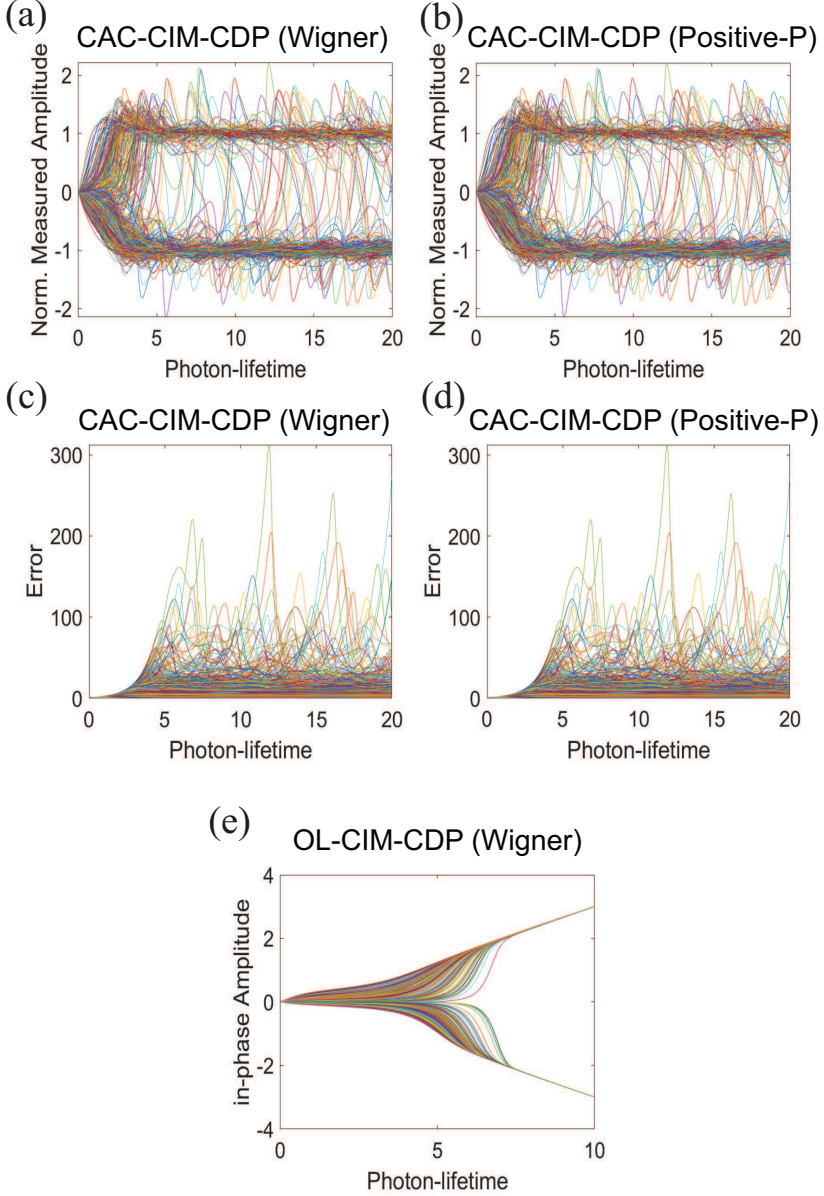


Fig. 2: Amplitude and Error evolution of each CIM model

(a) and (b) indicates the normalised amplitude $g\tilde{\mu}_r$ evolution of CAC-CIM-CDP (Wigner and Positive- P) where $\tau = 1$. With the introduction of CAC to the system, the chaotic behaviour is recognisable in the CAC-CIM-CDP models. (c) and (d) corresponds to the error e_r evolution of CAC-CIM-CDP (Wigner and Positive- P). (e) is the in-phase amplitude c_r evolution of OL-CIM-CDP. The system size was set as $N = 2000$ while the compression and the sparseness were 0.6 and 0.2 respectively for all the models.

2.3 Comparison with Simulated Annealing

Here our purpose is to demonstrate that CAC feedback is effective on CIM by comparing CAC-CIM to OL-CIM and SA. We follow the Metropolis algorithm for l_0 -regularised CS stated in [11]. As same as in [11], 1000 samples of the observation matrix and source signal and true support vector are randomly generated according to Section 4.5 under $N = 500$, $\alpha = a = 0.6$, $w_{noise} = 0$ (no observation noise). With the same observation matrices, source signals, and support vectors in all models, we statistically evaluate how well CAC-CIM estimates support in comparison to OL-CIM and SA when all R_r are fixed to be the source signal x_r . To measure the support estimation quality, we used the direction cosine defined as $\sum_{r=1}^N \xi_r \sigma_r / \sqrt{\sum_{r=1}^N \xi_r \sum_{r=1}^N \sigma_r}$ where (ξ_1, \dots, ξ_N) is the true support vector and $(\sigma_1, \dots, \sigma_N)$ is the estimated one. When the estimation is perfect, the direction cosine is equal to 1. We selected $\eta = 0.05$ corresponding to l_0 -regularisation parameter $\lambda = \eta^2/2 = 0.00125$ as in [11].

First, we evaluate the temporal profiles of the optimisation processes for the support estimation in CAC-CIM (Wigner), CAC-CIM (Positive- P), OL-CIM and SA. The upper three graphs (from left to right, CAC-CIM (Wigner), CAC-CIM (Positive- P) and OL-CIM respectively) in Fig. 3a show the change in the direction cosine of the three CIM models depending on the runtime on the CPU and the wall-clock time of physical CIM. For CAC-CIM (Wigner), and CAC-CIM (Positive- P) models, $20\times$ photon's lifetimes of integral interval (with 1000 time-steps) for the SDEs are about 105ms and 68ms of run-time respectively, and for OL-CIM, $5\times$ photon's lifetime of integral interval (with 50 time-steps) for the SDE is about 11ms. The physical CIM's wall-clock time for this optimisation is roughly estimated to be around 0.5ms, which can be estimated from the round-trip time of $N = 500$ and the time-steps-to-solution for the Sherrington-Kirkpatrick problem with $N = 500$ [20]. The direction cosine of these CIM models converged to about 1 by these run-times. The lower two graphs in Fig. 3a show the change in the direction cosine of SA depending on the runtime on CPU under constant temperature at $T = 0$ and exponential cooling scheduling from $T = 0.02$ to 0.00002. We adjusted the Monte-Carlo steps of SA (bottom two graphs of Fig. 3a) to accompany the wall-clock time of physical CIM (0.5ms) and the run-time of CAC-CIM (Wigner) (105ms). In our computational environment, the number of Monte Carlo steps for SA with runtimes of 0.5ms and 105ms is about 230 and 46000 steps, respectively. In SA, the direction cosine converged to about 1 by 105ms, while that did not by 0.5ms.

Next, we compare the histogram of the final states of direction cosines in CAC-CIM (Wigner), CAC-CIM (Positive- P), OL-CIM and SA. The upper three graphs in Fig. 3b indicate the histogram of the three CIM models (CAC-CIM (Wigner), CAC-CIM (Positive- P), OL-CIM, respectively), while the lower two graphs in Fig. 3b show the histograms of SA at run-times of 0.5ms and 105ms under zero temperature and exponential cooling schedules

respectively. Comparing these graphs, the proportion of the direction cosines of CAC-CIM (Wigner) and CAC-CIM (Positive- P) close to 1 is higher than those of OL-CIM and SA. The two-sample one-sided Kolmogorov-Smirnov test suggests that the histograms of the final direction cosines of CAC-CIM (Wigner) and CAC-CIM (Positive- P) are significantly biased toward 1 compared with all of those of OL-CIM and SA (P-value < 0.0001).

The above results thus demonstrate that CAC-CIM outperformed OL-CIM on support vector estimation and outperformed SA within the same run-time.

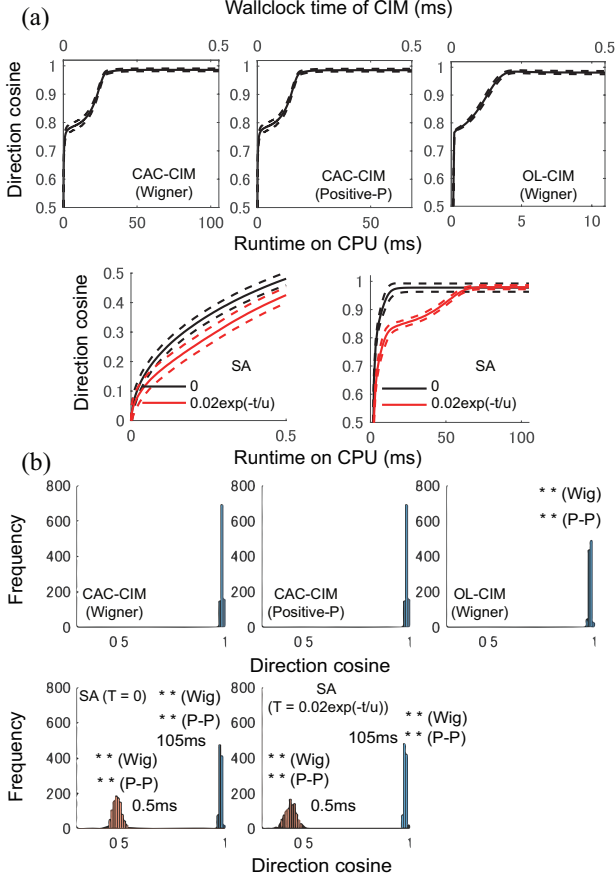


Fig. 3: Comparison of CAC-CIM with SA in support vector estimation when SA run-time is set to be the same as that of CAC-CIM (Wigner) model under the same computational environment

(a) comparison of the change of direction cosine is mapped for CAC-CIM (Wigner) (upper left), CAC-CIM (Positive- P) (upper middle), OL-CIM (upper right) and SA (bottom two). For CAC-CIM the photon's lifetime is increased for $20\times$ in 105ms of run time. For OL-CIM $5\times$ photon's lifetime is about 11ms. In SA 0: constant (zero). $0.02 \exp(t/u)$: exponential cooling scheduling were tested. All graphs show the mean (solid line) and standard deviation (dashed line) of 1000 samples. In our computational environment, the number of Monte Carlo steps for SA with a run time of 105ms is about 46000. (b) Histogram of 1000 final states of the direction cosine is shown (CAC-CIM (Wigner) (upper left), CAC-CIM (Positive- P) (upper middle), OL-CAC (upper right) and SA (bottom two)). ** (Wig) and ** (P-P) in the graphs means that cumulative histograms of these final states are significantly higher than Wigner and Positive- P models (P-value < 0.0001 on two-sample one-sided Kolmogorov-Smirnov test) and thus the final states of Wigner and Positive- P models are biased towards 1 compared to those final states. $N = 500$, $\alpha = a = 0.6$. In all CIM, $g^2 = 10^{-7}$.

2.4 Comparison with ground state predicted with statistical mechanics on alternating minimisation

We compare CAC-CIM-CDP’s capability to find the ground state with that of OL-CIM-CDP. In our previous study, we derived the macroscopic parameter equation (Eq. (26)-(28) in [11]) using a non-equilibrium statistical mechanics method to show the performance limit of OL-CIM-CDP. In the limit of the saturation parameter $g^2 \rightarrow 0$, the macroscopic parameter equation derived in the previous study is consistent with that for a two-state Potts spin system defined by the QUBO Hamiltonian Eq. (4). Therefore, the macroscopic parameter equation in this limit can predict the ground state of the Hamiltonian. Through a comparison of solutions of CAC-CIM-CDP and OL-CIM-CDP with a solution of the macroscopic parameter equation in the limit of $g \rightarrow 0$, we demonstrate the efficacy of CAC feedback on the alternating minimisation for optimising the Hamiltonian.

The precondition for applying statistical mechanics is that the values of all entries in the observation model Eq. (26), which is the premise of Eq. (3) and Eq. (4), are randomly determined as described in Section 4.5. To compare solutions of the models with the ground state predicted with statistical mechanics, 10 samples of the observation matrix and source signal and true support vector are randomly generated according to Section 4.5 under $N = 2000$ and various values of a , α and ν . Here ν indicates the standard deviation of the observation noise (w_{noise}). Then, we execute Algorithms 1 and 2 for the alternating minimisation in CAC-CIM-CDPs (Wigner and Positive- P) and OL-CIM-CDP sharing the same samples of observation matrices, source signals and support vectors. Here for Fig. 4, $\eta_{init} = 0.6$ and $\eta_{init} = 0.8$ was used for CAC-CIM-CDP models and OL-CIM-CDP respectively. η_{end} was set to 0.18 in Fig. 4a and Fig. 4b while in Fig. 4c and Fig. 4d η_{end} was set to 0.35.

The marks in Fig. 4 show the averaged root-mean-square-error (RMSE) calculated as $\sqrt{1/N \sum_{r=1}^N (R_r \sigma_r - x_r \xi_r)^2}$ of sampled solutions obtained from OL-CIM-CDP, Wigner and Positive- P of CAC-CIM-CDPs. Here σ_r is calculated as stated in Eq. (20). The black solid lines in Fig. 4 indicate RMSE at the ground state corresponding to successful signal retrieval, which is predicted with statistical mechanics. RMSEs of Wigner and Positive- P CAC-CIM-CDPs tend to keep a better consistency with that of the ground state compared to OL-CIM-CDP for various values of a , α and ν . Especially as shown in Figs. 4b and 4d, RMSE of OL-CIM-CDP tend to deviate gradually from that of the ground state as increasing a , while both Wigner and Positive- P CAC-CIM-CDPs keep up a better consistency with the theoretical prediction.

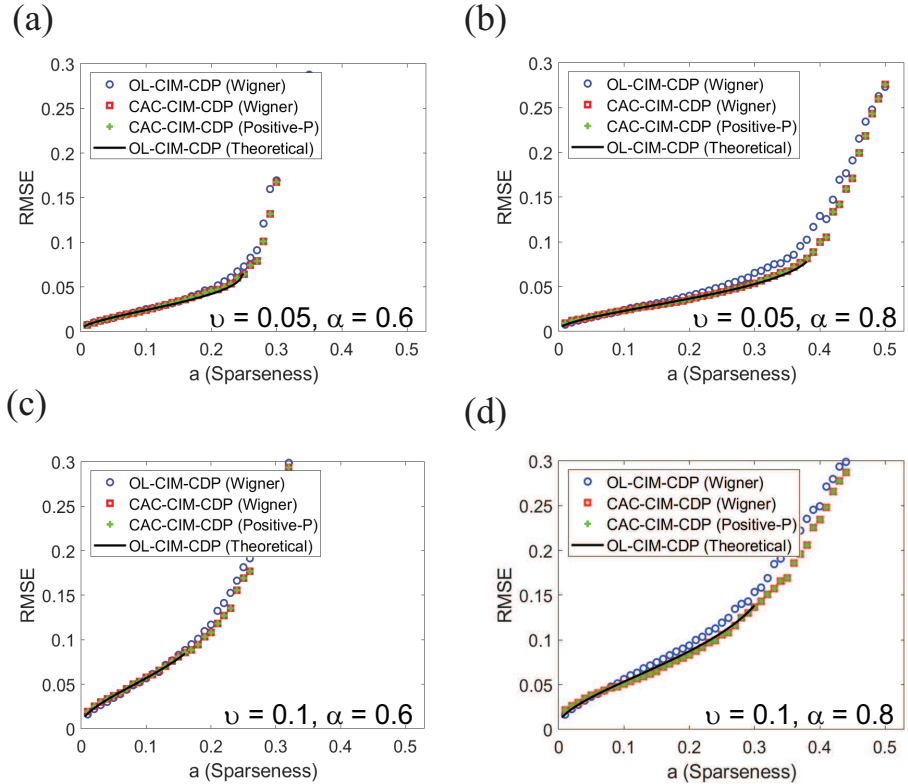


Fig. 4: Comparison of average RMSE of CAC-CIM-CDP models to the theoretical limit of OL-CIM-CDP when observation noise is present

(a) and (b) indicates the average performance for $N = 2000$ system where $\alpha = 0.6$ and $\alpha = 0.8$ respectively for $\nu = 0.05$. (c) and (d) states the average performance for $\nu = 0.1$. For all graphs $\eta_{init} = 0.8$ and $\eta_{init} = 0.6$ was used for CAC-CIM-CDP models and OL-CIM-CDP respectively. (a) and (b) η_{end} was set to 0.18. (c) and (d) η_{end} was set to 0.35.

2.5 Application to Sparse MRI

We evaluate the performance of CAC-CIM-CDP, OL-CIM-CDP and LASSO [22] on MRI data.

In the following numerical experiment, we used two different-sized sparse images (64×64 and 128×128 pixels) spanned by a Haar basis function. Detailed explanations of the two images we used as the source images are given in Section 4.6 in Methods. In accordance with our previous work [11], we sought to reconstruct the two images from the undersampled k -space data and by solving the optimisation problem defined in Eq. (27) (see Section 4.6). To realise the optimisation problem in Eq. (27) on CIM, the Haar wavelet

transform coefficients are estimated with the mutual interaction term and the Zeeman term constructed according to Eq. (28) and (29) in Section 4.6. The compression rate of the k -space data from the 64×64 and 128×128 images is 0.4 and 0.3 respectively. And the sparseness of the images is 0.212 and 0.178 respectively. As the solver for CDP, we used the Conjugate Gradient Descent method (further details on CDP optimisation refer to Section 4.2).

In Fig. 5a and Fig. 5b, for 10 simulations the average RMSE value is indicated for each threshold η for 64×64 and 128×128 images respectively. As for the minimum RMSE in the 64×64 case, LASSO (black line), OL-CIM-CDP (red), CAC-CIM-CDP (Wigner) (green) and CAC-CIM-CDP (Positive- P)'s (blue) can be stated as, 0.0292, 0.0216, 0.0182 and 0.0182 respectively (for the corresponding reconstructions see Fig. 6). In the 128×128 case, the minimum RMSE is 0.0276, 0.0242, 0.0209 and 0.0209 respectively (for the corresponding reconstructions see Fig. 7). Comparing the RMSE values acquired it is clear that CAC-CIM-CDP models have a better average performance compared to the other approaches in both image sizes. And even after reaching the optimal reconstruction for the given parameters, CAC-CIM-CDP tends to keep up a minimal error rate compared to LASSO and OL-CIM-CDP. This indicates that the effective range of CAC-CIM-CDP is much wider than OL-CIM-CDP. In both image sizes, the Wigner and Positive- P variations of CAC-CIM-CDP produce identical RMSE results.

In Fig. 6 and 7 the minimal RMSE constructions are shown for LASSO, OL-CIM-CDP, CAC-CIM-CDP (Wigner) and CAC-CIM-CDP (Positive- P). In Fig. 7, only CAC-CIM-CDP (Positive- P)'s reconstruction is shown because it is clear that both CAC-CIM-CDP (Wigner) and CAC-CIM-CDP (Positive- P)'s performance is identical. In the 64×64 image reconstruction when RMSE values are compared, CAC-CIM-CDP models have better reconstruction accuracy. The enlarged portions indicate the difference in pixel identification of each model compared to the initial resized image. Considering both simulations it is clear that even though the system size increases, proposing models have the upper hand in performing an accurate reconstruction compared to other models.

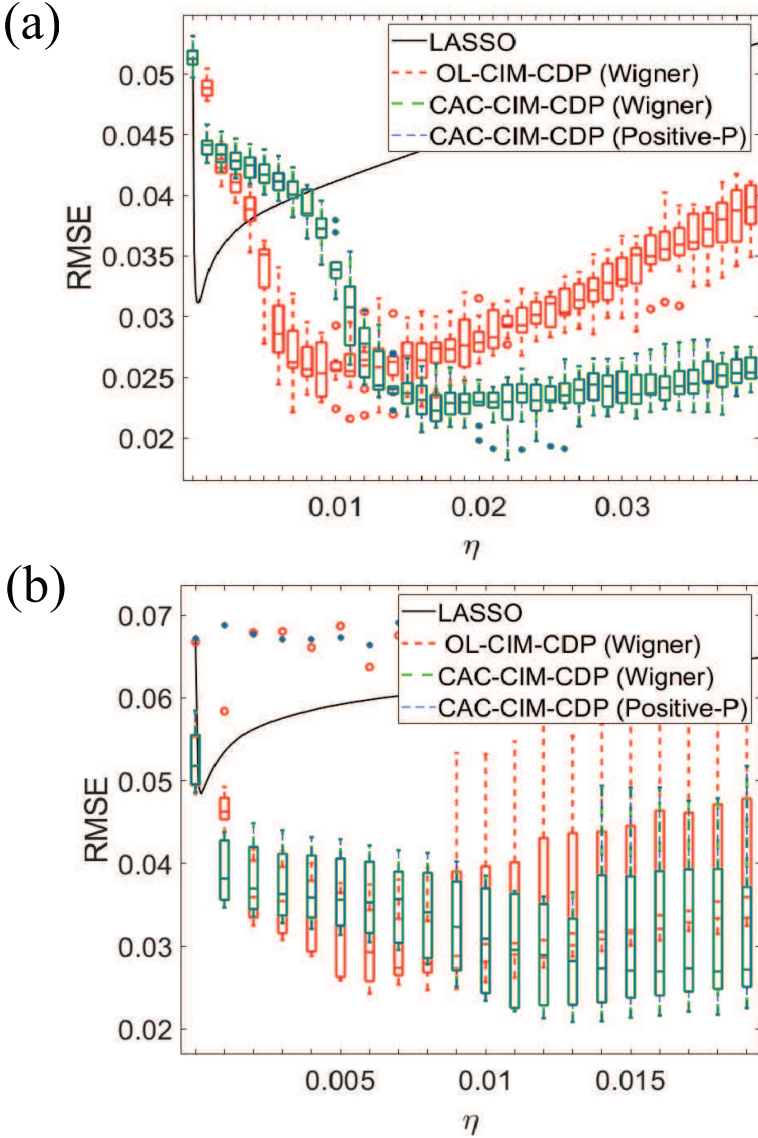


Fig. 5: Average performance of the models when l_0 -regularisation parameter varies for different image sizes

(a) Performance on 64×64 and (b) Performance on 128×128 . The black line indicates the performance on LASSO while the red boxes correspond to OL-CIM-CDP. Green and blue boxes indicate the performance on CAC-CIM-CDP Wigner and Positive- P respectively. For different threshold values, the graphs illustrate the maximum, minimum, 25-th percentile (bottom edge), 75-th percentile (top edge), and median (central horizontal line) of RMSEs for each model with box plots. The markers indicate the outliers. The compression and sparseness for (a) were 0.4 and 0.212 respectively while for (b) were 0.3 and 0.178.

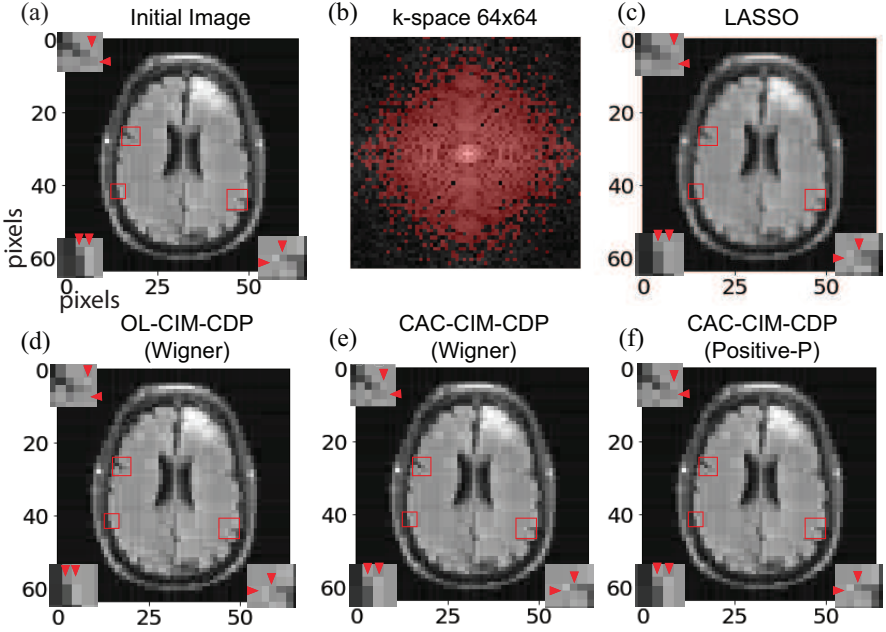


Fig. 6: Reconstructed Images for 64×64

(a) Resized 64×64 initial image. The compression and sparseness were 0.4 and 0.212 respectively. (b) Undersampled k-space data (random red points). (c), (d), (e), and (f) correspond to the reconstructions obtained from LASSO, OL-CIM-CDP, CAC-CIM-CDP (Wigner), and CAC-CIM-CDP (Positive- P) with RMSE values 0.0292, 0.0216, 0.0182 and 0.0182 respectively. The enlarged image portions indicate the pixel-wise differences between the reconstructions. For (d) 31 alternating minimisation processes were performed. For (e) and (f) 11 alternating minimisations were performed. And for (c), (d), (e), and (f) $\eta_{init} = \eta_{end}$ was 0.0003, 0.011, 0.022, and 0.022 respectively.

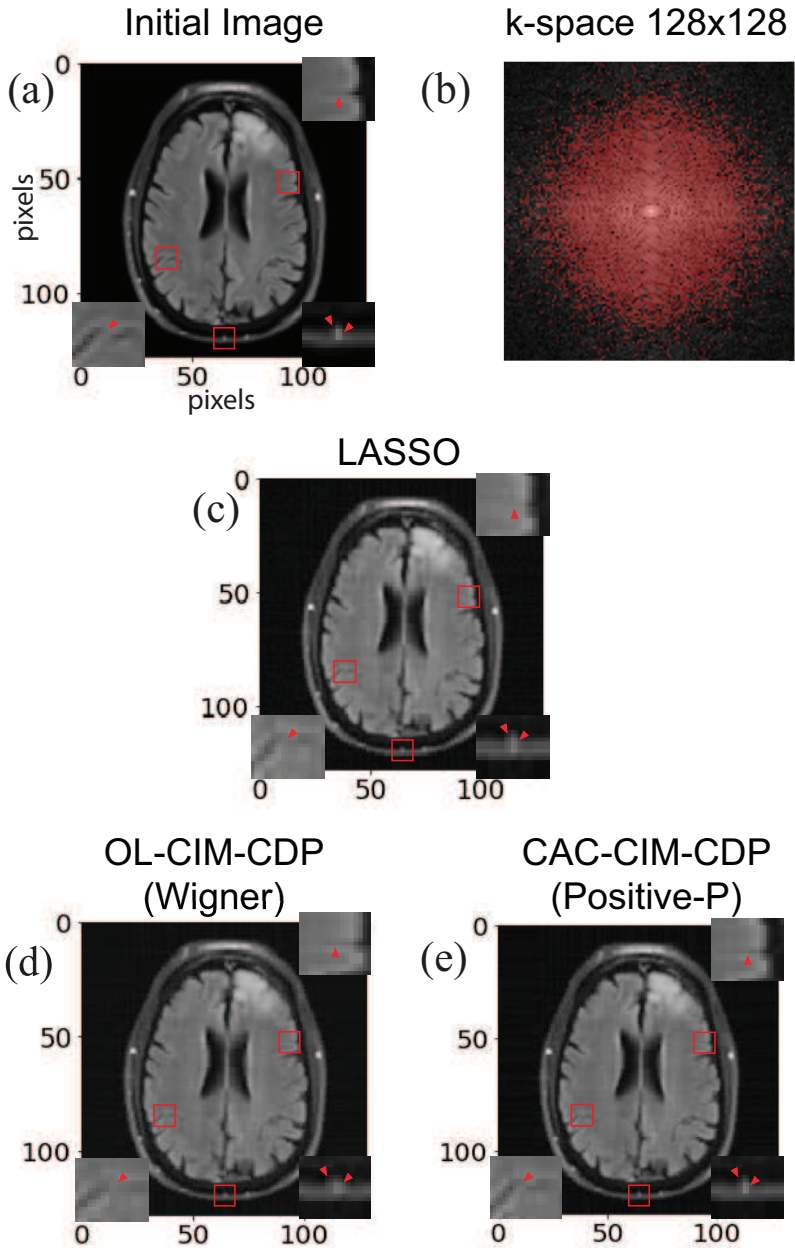


Fig. 7: Reconstructed Images for 128×128

(a) Resized 128×128 initial image. The compression and sparseness were 0.3 and 0.178 respectively (a) Red dots indicate the sampled k-space from the (a)'s k-space. (c), (d), and (e) correspond to the reconstructions obtained from LASSO, OL-CIM-CDP, and CAC-CIM-CDP (Positive- P) with RMSE values 0.0276, 0.0243, and 0.0209 respectively. The enlarged image portions indicate the pixel-wise differences between the reconstructions. For (d) and (e), 31 and 11 alternating minimisation processes were performed respectively. And for (c), (d), and (e) $\eta_{init} = \eta_{end}$ was 0.0001, 0.006, and 0.013 respectively.

3 Discussion

In this paper, we have proposed an improved CIM approach to solve l_0 -regularised compressed sensing problems. The proposed algorithm has shown that it can outperform the previously proposed algorithm accuracy-wise in all the simulations performed. With the OL-CIM algorithm, the CIM model in use was lacking the CAC feedback for chaotically exploring solutions. Therefore, CAC-CIM has been able to provide convergence to a better solution than OL-CIM. One factor to emphasise here is that CAC does not guarantee convergence to the ground state. Even the ground state is reached, due to the forceful equalisation to τ may prevent from stopping there. Even though this is the case in this paper, CAC has been shown to be effective especially when the problem instances are relatively harder in both artificial random data and MRI data.

3.1 Effect of system size on performance

The introduction of CAC has previously been shown to have better performance with small-scale frustrated Ising problem instances [17]. In this manuscript, we have demonstrated the applicability of CAC for real-world combinatorial optimisation problems (in this case Compressed sensing) where the problem instances with a Zeeman term are mapped to a QUBO formulation that is large-scale. The simulations with random artificial data on various system sizes are illustrated in Supplementary note 1. Even though the performance increase is present, in very large system sizes such as in 128×128 , it is clear that the RMSE gap between CAC-CIM-CDP and OL-CIM-CDP is smaller compared to 64×64 . This poses the question that whether there is a system-size threshold for CAC-CIM-CDP in the very-large-scale regime. Considering the MRI-based simulations require 4096 and 16384 DOPO pulses to operate (compared to 16 DOPOs in theoretical simulations in [17]), the system size of CAC's applicability is largely improved. Yet the system-size-wise dependency is yet to be explored.

3.2 Advantages of CAC-CIM architecture

With the use of CDP, the problem which involves quadratic optimisation has been solved in this hybrid system. As shown in the schematic illustration of the CAC-CIM-CDP in Fig. 1, proposing approach performs an alternating minimisation between the CIM and CDP. It is clear considering the results stated in Section 2.5 that CAC-CIM-CDP has outperformed OL-CIM-CDP and the generally used approach LASSO which is an l_1 -regularised method for solving compressed sensing problems. It is interesting to see that advancements in CIM architecture can offer better results in real-world problem instances.

3.3 CAC-CIM-CDP (Wigner) Vs. CAC-CIM-CDP (Positive- P)

Even though this paper introduces two variants (Wigner and Positive- P) of CAC-CIM-CDP, the performances have been almost identical between the models. However, we encountered a deviation when the problem instances become harder i.e. sparseness/compression ratio becomes higher when $w_{noise} = 0$. The results are presented in Supplementary note 2. As the models approach a threshold point for optimal reconstruction (a critical sparseness/compression ratio), beyond that the producing RMSE values are somewhat different between the models. Performance-wise it is hard to state that one model is better than the other. Because the significance of Wigner and Positive- P lies in the density matrix approximation and how it behaves with a large quantum noise presence. We discuss this in Supplementary note 2.

3.4 Future improvements to the CAC-CIM-CDP

3.4.1 Simultaneous minimisation

One of the major bottlenecks the proposed model (CAC-CIM-CDP) has is the alternating minimisation process between the CIM and CDP. This is a time-consuming operation. As a future direction to this model, we plan to improve the CIM system to accommodate quadratic optimisation problems and perform simultaneous minimisation using only the CIM to solve compressed sensing problems. We believe that the use of "CIM-only" will have a positive effect on accuracy as well.

3.4.2 CAC-CIM-CDP with large quantum noise

While this manuscript solely focuses on combining CAC with CIM for solving CS problems more accurately, the considered quantum noise present in the CIM is very low ($g^2 = 10^{-7}$). This opens up a problem of whether CAC-CIM-CDP can keep up the performance with a large quantum noise presence. For small-scale frustrated Ising hamiltonians, this has been previously explored in [17] ($N = 16$) where it has shown a decrease in success probability for larger g^2 terms. This result is consistent with CAC-CIM-CDP as well as shown in Supplementary note Fig. ?? for MRI simulations. Recent advances in CIM research have led to the introduction of a method known as Negative Parametric Gain (NPG), which accommodates higher quantum noise and at the same time as maintaining a higher probability of success [23]. This method considers a negative starting pump rate with large injection field feedback. NPG has shown promising results in the theoretical simulations [23]. We are planning to improve the endurance of the CAC-CIM-CDP with NPG for a larger quantum noise presence.

3.4.3 CAC-CIM-CDP with the mean-field CIM model

As it is obvious from the perspective of numerical simulations, CAC-CIM-CDP SDEs are computationally costly to simulate. Even though the shown results are acquired using a GPU implementation of the SDEs, as a digital simulator, field-programmable gate arrays (FPGAs) are more suitable (less energy cost, faster processing etc). As a future direction, we plan on implementing the mean-field CIM SDEs [18, 20] with CAC on an FPGA to perform compressed sensing simulations. Due to the fact that CAC-CIM-CDP has relatively low noise present in the system, we believe that the mean-field SDEs will have approximately the same or better results but with faster simulation times. This is mainly due to the simplicity and the negligence of the noise terms in the mean-field CIM SDEs.

4 Methods

4.1 Stochastic Differential equation in OL-CIM-CDP and CAC-CIM-CDP

4.1.1 Wigner-type

The CIM model based on the Wigner formulation was introduced in [24, 25]. The c -number Heisenberg Langevin equation [24] was used to overcome the higher computational cost of simulating the direct density matrix formulation of CIM and it has been found to be equivalent to the truncated Wigner SDEs. The density operator master equation expanded by the Wigner function results in the Kramers-Moyal series including third-order terms. In order to derive the Langevin equation, we neglect third-order terms [17]. Then, we can formulate the following Wigner SDEs used for OL-CIM-CDP.

$$\frac{d}{dt}c_r = [-1 + p - (c_r^2 + s_r^2)]c_r + \tilde{K} \left(\frac{dc_r}{dt} \right)_{inj,r} + g\sqrt{(c_r^2 + s_r^2) + \frac{1}{2}}W_{r,1}, \quad (13)$$

$$\frac{d}{dt}s_r = [-1 - p - (c_r^2 + s_r^2)]s_r + g\sqrt{(c_r^2 + s_r^2) + \frac{1}{2}}W_{r,2}. \quad (14)$$

Here, in-phase and quadrature-phase normalised amplitudes are represented as c_r and s_r respectively. p is the normalised pump rate. If p is above the oscillation threshold ($p > 1$), each of the OPO pulses is either in the 0-phase state or π -phase state. The last terms of the upper and lower equations express the vacuum fluctuations injected from external reservoirs and the pump fluctuations coupled to the OPO system via gain saturation [11]. $W_{r,1}$ and $W_{r,2}$ are independent real Gaussian noise processes satisfying $\langle W_{r,k}(t) \rangle = 0$ and $\langle W_{r,k}(t)W_{r',l}(t') \rangle = \delta_{rr'}\delta_{kl}\delta(t-t')$. g indicates the saturation parameter. $(dc_r/dt)_{inj,r}$ is the optical injection field, which only considers the in-phase

amplitudes for the calculations. The injection field is defined in Eq. (5) and Eq. (6). \tilde{K} indicates the normalised feedback strength.

Focusing on the behaviour of the OPO pulses only in the in-phase direction, the Wigner-type SDE, which is used for CAC-CIM-CDP, can be stated as,

$$\frac{d}{dt}\mu_r = -(1-p+j)\mu_r - g^2\mu_r^3 + \sqrt{j}\left(V_r - \frac{1}{2}\right)W_{R,r} + K\left(\frac{d\mu_r}{dt}\right)_{inj,r}, \quad (15)$$

$$\frac{d}{dt}V_r = -2(1-p+j)V_r - 6g^2\mu_r^2V_r + 1 + j + 2g^2\mu_r^2 - 2j\left(V_r - \frac{1}{2}\right)^2. \quad (16)$$

Here μ_r and V_r are the mean-amplitudes and the variance of the r -th DOPO pulse. $(d\mu_r/dt)_{inj,r}$ is the optical injection field defined in Eq. (8)-(11). $W_{R,r}$ is independent real Gaussian noise processes satisfying $\langle W_{R,r}(t) \rangle = 0$ and $\langle W_{R,r}(t)W_{R,r'}(t') \rangle = \delta_{rr'}\delta(t-t')$. g , p , j and K indicate the saturation parameter, pump rate, the normalised out-coupling rate for optical homodyne measurement and the feedback strength, respectively.

4.1.2 Positive- P -type

Positive- P (P-P) representation [26] is a generalised form of Glauber-Sudarshan P representation. When the density operator master equations are expanded using the P-P distribution function, the resulting Kramers-Moyal series only consists of first and second-order terms. Due to this factor, there is no truncation needed to derive the Langevin equation. Because of this one can argue that P-P SDEs might be a better candidate for density operator approximations. The effectiveness of P-P SDEs has been demonstrated on CIMs with higher quantum noise presence [17]. We can formulate the P-P-type SDEs we used for CAC-CIM-CDP.

$$\frac{d}{dt}\mu_r = -(1-p+j)\mu_r - g^2\mu_r(\mu_r^2 + 2n_r + m_r) + \sqrt{j}(m_r + n_r)W_{R,r} + K\left(\frac{d\mu_r}{dt}\right)_{inj,r}, \quad (17)$$

$$\frac{d}{dt}n_r = -2(1+j)n_r + 2pm_r - 2g^2\mu_r^2(2n_r + m_r) - j(m_r + n_r)^2, \quad (18)$$

$$\begin{aligned} \frac{d}{dt}m_r = & -2(1+j)m_r + 2pn_r - 2g^2\mu_r^2(2m_r + n_r) + p \\ & - g^2(\mu_r^2 + m_r) - j(m_r + n_r)^2. \end{aligned} \quad (19)$$

Here μ_r corresponds to the mean-amplitude, m_r and n_r represent variances of quantum fluctuations of the r -th DOPO pulse. $(d\mu_r/dt)_{inj,r}$ is the optical injection field defined in Eq. (8)-(11). $W_{R,r}$ is independent real Gaussian noise processes satisfying $\langle W_{R,r}(t) \rangle = 0$ and $\langle W_{R,r}(t)W_{R,r'}(t') \rangle = \delta_{rr'}\delta(t-t')$. g , p , j and K are the same as those for the Wigner model.

4.2 Optimisation in CDP

The CDP performs the optimisation of the Hamiltonian (Eq. 4) with respect to R_r for a support vector σ given by CIM. σ is obtained by binarising the measured-amplitude ($\tilde{\mu}_r$) defined in Eq. (10) (CAC-CIM-CDP) or in-phase amplitude c_r (OL-CIM-CDP) with the Heaviside function stated as,

$$\sigma_r = \text{Heaviside}(x_r) = \begin{cases} 1, & (x_r > 0) \\ 0, & (x_r \leq 0). \end{cases} \quad (20)$$

The CDP solve the following system of equations, which is satisfied the stationary point that minimises \mathbb{H} with respect to r .

$$R_r \sum_{k=1}^M (A_r^k)^2 = \sigma_r \mathbb{H}_r, \quad (21)$$

$$\mathbb{H}_r = - \sum_{r'=1(\neq r)}^N \sum_{k=1}^M A_r^k A_{r'}^k R_{r'} \sigma_{r'} + \sum_{k=1}^M A_r^k y^k. \quad (22)$$

Here, \mathbb{H}_r in Eq. (22) is the local field of the CDP, which is the same as Eq. (4) and (11). For the simulations, we used the Jacobi method or Conjugate Gradient Descent (CGD) method as the CDP optimiser. During the optimisation in the CDP, all σ_r are fixed.

4.3 Schedule of pump rate, threshold and target amplitude for optimisation in CIM

A rough parameter search was used to determine the schedules for each of the following parameters in the experiments. The pump rate p for both Wigner and P-P type CAC-CIM-CDPs was scheduled depending on the time t as follows.

$$p = (p_{thr} - d) + \frac{2d}{1 + e^{-\left(\frac{t-4}{2}\right)}}. \quad (23)$$

Here, $p_{thr} = 1$ for all simulations of both Wigner and P-P type CAC-CIM-CDPs. For artificial random data and MRI data simulations, d was set at 0.6 and 0.4 respectively.

In accordance with [11], the pump rate p for OL-CIM-CDP was scheduled depending on the time t as follows.

$$p = 1.5 \times \left(\frac{t}{5}\right)^2. \quad (24)$$

The pump rate becomes equal to 1.5 when $t = 5$. We used this pump rate schedule for all simulations of OL-CIM-CDP. In both CAC-CIM-CDP and OL-CIM-CDP, the threshold η was scheduled depending on the alternating iteration time i as follows.

$$\eta_i = \max \left[\eta_{init} \left(1 - \frac{i}{velo} \right), \eta_{end} \right]. \quad (25)$$

Here $velo = 51$ for all simulations of both CAC-CIM-CDP and OL-CIM-CDP in artificial random data. For the MRI data, $velo = 31$ and $velo = 11$ were used in OL-CIM-CDP and CAC-CIM-CDP respectively. For synthesised random data (Figs. 3, 4, and Supplementary note Fig. ??), the threshold η was linearly lowered from η_{init} to η_{end} as the alternating minimisation proceeds. η_{init} and η_{end} are adjusted to maximise the performance of those models. On the other hand, for MRI data (Figs. 5, 6, 7, Supplementary note Fig. ?? and Supplementary note Fig. ??), the threshold η was constant by setting as $\eta_{init} = \eta_{end}$. The values of η_{init} and η_{end} used for each simulation are shown in the figure captions.

In both Wigner and P-P type CAC-CIM-CDPs, the target amplitude of CAC, τ , was constant with respect to the time t . For the simulations in Fig. 4a and Fig. 4b, $\tau = 0.21$ was used while in Fig. 4c and Fig. 4d τ was set to 0.15. For other simulations, τ was 1.

4.4 Observation model for Compressed Sensing

The observation model that is the premise of Eq. (3) and Eq. (4) is defined as follows.

$$\begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^M \end{bmatrix} = \begin{bmatrix} A_1^1 & A_2^1 & \cdots & A_N^1 \\ A_1^2 & A_2^2 & \cdots & A_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ A_1^M & A_2^M & \cdots & A_N^M \end{bmatrix} \begin{bmatrix} \xi_1 x_1 \\ \xi_2 x_2 \\ \vdots \\ \xi_N x_N \end{bmatrix} + \begin{bmatrix} w_{noise}^1 \\ w_{noise}^2 \\ \vdots \\ w_{noise}^M \end{bmatrix}. \quad (26)$$

Here, $A \in \mathbb{R}^{N \times M}$ is the observation matrix, $y \in \mathbb{R}^M$ implies the observation signal, $x \in \mathbb{R}^N$ and $\xi \in (0, 1)^N$ are the true source signal and true support, respectively. $w_{noise} \in \mathbb{R}^M$ indicates the observation noise satisfying $\langle w_{noise}^k \rangle = 0$ and $\langle w_{noise}^k w_{noise}^{k'} \rangle = \nu^2 \delta_{kk'}$. ν^2 is the variance of the observation noise.

4.5 Artificial random data

To verify the performance of the proposed models statistically and moreover compare those results with ground states predicted with statistical mechanics [11], we used many samples of artificial random data $y \in \mathbb{R}^M$ synthesised from the observation model Eq. (26) in which the values of all entries were randomly determined as follows. Each entry of the observation matrix $A \in \mathbb{R}^{M \times N}$ is randomly generated from an independent and identical normal distribution with the variance of $1/M$, which satisfies $\langle A_r^k \rangle = 0$ and $\langle A_r^k A_{r'}^{k'} \rangle = 1/M \delta_{rr'} \delta(kk')$.

Each entry of the true source signal $x \in \mathbb{R}^N$ is randomly generated from an independent and identical normal distribution with the variance of 1, which satisfies $\langle x_r \rangle = 0$ and $\langle x_r x_{r'} \rangle = \delta_{rr'}$. $a \times N$ elements of $\xi \in (0, 1)^N$ are randomly selected and assigned 1 while others are assigned 0. a is the sparseness defined in the Introduction.

4.6 Simulations with MRI data

To evaluate the performance of the proposed models on realistic data, we used MRI data provided from the fastMRI datasets [27]. The initial brain MRI used here was a 320×320 image. To reduce the problem size, we resized the image to 64×64 and 128×128 images with the BILINEAR interpolation method. We applied the Haar-wavelet transform (HWT) to the two different-sized images and in Fig. 6 and Fig. 7 we set 78.8% and 82.2% of the HWT coefficients to zero to create two different-sized sparse images (64×64 and 128×128 pixels) spanned by Haar basis functions with a sparseness of 0.212 and 0.178, respectively. Then, we applied the discrete Fourier transform (DFT) to the two different-sized sparse images to obtain 64×64 and 128×128 k -space data, respectively. Finally, we undersampled 1638 and 4915 points from the 64×64 and 128×128 k -space data at random red points (Fig. 6b and Fig. 7b) to create two observation signals with a compression rate of 0.4 and 0.3 respectively.

In accordance with our previous work, we sought to reconstruct the source signals from the undersampled k -space data by solving the following optimisation problem with CAC-CIM-CDP and OL-CIM-CDP.

$$x = \operatorname{argmin}(\|y - SFx\|_2^2 + \frac{1}{2}\gamma\|\Delta_v x\|_2^2 + \frac{1}{2}\gamma\|\Delta_h x\|_2^2 + \lambda\|\Psi x\|_0). \quad (27)$$

Here, x is a source signal, and y is the observation signal constructed through the above steps. F indicates the DFT matrix and Ψ is the HWT matrix. F and Ψ are orthogonal matrices and their transpose matrices correspond to inverse DFT and inverse HWT, respectively. S is an undersampling matrix executing undersampling at random red points shown in Fig. 6b and Fig. 7b. Δ_v and Δ_h are the matrices discretely representing the vertical and horizontal second-order derivative operators, respectively. γ and λ are the l_2 and l_0 regularisation parameters.

To implement the optimisation problem in Eq. (27) on CIM, we estimate the HWT coefficients instead of the pixel values of the image. Applying the HWT $r = \Psi x$ to Eq. (27), the mutual interaction matrix J and the Zeeman term vector h^z for CIM are given as

$$h^z = SF\Psi^T y, \quad (28)$$

$$\tilde{J} = \Psi F^T S^T S F \Psi^T + \gamma \Psi \Delta_v^T \Delta_v \Psi^T + \gamma \Psi \Delta_h^T \Delta_h \Psi^T. \quad (29)$$

Here, the observation matrix is given as $A = SF\Psi^T$. The second and third terms in \tilde{J} are from the l_2 regularisation terms. After the alternating minimisation, the output of the CDP, r , is transformed to the image, x , with the inverse HWT $x = \Psi^T r$. γ is set to 0.0001. \tilde{K} for OL-CIM-CDP was set to 0.25 while K for CAC-CIM-CDP was 0.01. Here we use LASSO's solution as the initial condition for the CIM simulation.

Authors' contributions. M.D.S.H.G., and T.A., modelled the system, performed the numerical simulations for the proposed models and wrote the

manuscript. M.D.S.H.G., T.A., and S.K., worked on the evaluation of the models. K.M., and M.O., provided feedback on numerical simulations. S.K., Y.I., and Y.Y., helped with the physics of CIM and provided feedback.

Availability of data and materials. The data generated and/or analysed during this study are not publicly available for legal/ethical reasons. But M.D.S.H.G. can provide the raw data if formally requested.

Competing Interests. The authors declare no competing interests.

Funding. This work is supported by the Japan Science and Technology Agency through its ImPACT program, NTT Research Inc. And Authors acknowledges the support of the NSF CIM Expedition award (CCF-1918549).

References

- [1] Bobin, J., Starck, J.-L. & Ottensamer, R. Compressed sensing in astronomy. *IEEE Journal of Selected Topics in Signal Processing* **2** (5), 718–726 (2008). <https://doi.org/10.1109/JSTSP.2008.2005337> .
- [2] Zhang, Y., Jiang, J. & Zhang, G. Compression of remotely sensed astronomical image using wavelet-based compressed sensing in deep space exploration. *Remote Sensing* **13** (2) (2021). URL <https://www.mdpi.com/2072-4292/13/2/288>. <https://doi.org/10.3390/rs13020288> .
- [3] Zhang, Y., Jiang, J. & Zhang, G. Compression of remotely sensed astronomical image using wavelet-based compressed sensing in deep space exploration. *Remote Sensing* **13** (2) (2021). URL <https://www.mdpi.com/2072-4292/13/2/288>. <https://doi.org/10.3390/rs13020288> .
- [4] Zhou, W.-P., Li, Y., Liu, Q.-S., Wang, G.-D. & Liu, Y. Fast compression and reconstruction of astronomical images based on compressed sensing. *Research in Astronomy and Astrophysics* **14** (9), 1207 (2014) .
- [5] Herman, M. A. & Strohmer, T. High-resolution radar via compressed sensing. *IEEE Transactions on Signal Processing* **57** (6), 2275–2284 (2009). <https://doi.org/10.1109/TSP.2009.2014277> .
- [6] Mamaghanian, H., Khaled, N., Atienza, D. & Vandergheynst, P. Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes. *IEEE Transactions on Biomedical Engineering* **58** (9), 2456–2466 (2011). <https://doi.org/10.1109/TBME.2011.2156795> .
- [7] Obuchi, T., Nakanishi-Ohno, Y., Okada, M. & Kabashima, Y. Statistical mechanical analysis of sparse linear regression as a variable selection problem. *Journal of Statistical Mechanics: Theory and Experiment* **2018** (10), 103401 (2018). URL <https://doi.org/10.1088/1742-5468/aae02c>. <https://doi.org/10.1088/1742-5468/aae02c> .

- [8] Kabashima, Y., Wadayama, T. & Tanaka, T. A typical reconstruction limit for compressed sensing based on l_p -norm minimization. *Journal of Statistical Mechanics: Theory and Experiment* **2009** (09), L09003 (2009). URL <https://doi.org/10.1088/1742-5468/2009/09/109003>. <https://doi.org/10.1088/1742-5468/2009/09/109003> .
- [9] Louizos, C., Welling, M. & Kingma, D. P. Learning sparse neural networks through l_0 regularization (2017). URL <https://arxiv.org/abs/1712.01312>.
- [10] Nakanishi-Ohno, Y., Obuchi, T., Okada, M. & Kabashima, Y. Sparse approximation based on a random overcomplete basis. *Journal of Statistical Mechanics: Theory and Experiment* **2016** (6), 063302 (2016). URL <https://dx.doi.org/10.1088/1742-5468/2016/06/063302>. <https://doi.org/10.1088/1742-5468/2016/06/063302> .
- [11] Aonishi, T., Mimura, K., Okada, M. & Yamamoto, Y. L0 regularization-based compressed sensing with quantum-classical hybrid approach (2021). [2102.11412](https://arxiv.org/abs/2102.11412).
- [12] Mohseni, N., McMahon, P. L. & Byrnes, T. Ising machines as hardware solvers of combinatorial optimization problems. *Nature Reviews Physics* **4** (6), 363–379 (2022) .
- [13] Matsumoto, N., Hamakawa, Y., Tatsumura, K. & Kudo, K. Distance-based clustering using qubo formulations. *Scientific reports* **12** (1), 1–10 (2022) .
- [14] Tanahashi, K., Takayanagi, S., Motohashi, T. & Tanaka, S. Application of ising machines and a software development for ising machines. *Journal of the Physical Society of Japan* **88** (6), 061010 (2019). URL <https://doi.org/10.7566/JPSJ.88.061010>. <https://doi.org/10.7566/JPSJ.88.061010>, <https://doi.org/10.7566/JPSJ.88.061010> .
- [15] Aonishi, T., Mimura, K., Okada, M. & Yamamoto, Y. Statistical mechanics of cdma multiuser detector implemented in coherent ising machine. *Journal of Applied Physics* **124** (23), 233102 (2018). URL <https://doi.org/10.1063/1.5041998>. <https://doi.org/10.1063/1.5041998>, <https://doi.org/10.1063/1.5041998> .
- [16] Kako, S. *et al.* Coherent ising machines with error correction feedback. *Advanced Quantum Technologies* **3** (11), 2000045 (2020). URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.202000045>. <https://doi.org/https://doi.org/10.1002/qute.202000045>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202000045> .
- [17] Inui, Y., Gunathilaka, M. D. S. H., Kako, S., Aonishi, T. & Yamamoto, Y. Control of amplitude homogeneity in coherent ising machines with

- artificial zeeman terms. *Communications Physics* **5** (1), 154 (2022). URL <https://doi.org/10.1038/s42005-022-00927-x>. <https://doi.org/10.1038/s42005-022-00927-x>.
- [18] Leleu, T., Yamamoto, Y., McMahon, P. L. & Aihara, K. Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity. *Phys. Rev. Lett.* **122**, 040607 (2019). URL <https://link.aps.org/doi/10.1103/PhysRevLett.122.040607>. <https://doi.org/10.1103/PhysRevLett.122.040607>.
- [19] Leleu, T., Yamamoto, Y., Utsunomiya, S. & Aihara, K. Combinatorial optimization using dynamical phase transitions in driven-dissipative systems. *Phys. Rev. E* **95**, 022118 (2017). URL <https://link.aps.org/doi/10.1103/PhysRevE.95.022118>. <https://doi.org/10.1103/PhysRevE.95.022118>.
- [20] Reifenstein, S., Kako, S., Khoystate, F., Leleu, T. & Yamamoto, Y. Coherent ising machines with optical error correction circuits. *Advanced Quantum Technologies* **4** (11), 2100077 (2021). URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.202100077>. <https://doi.org/https://doi.org/10.1002/qute.202100077>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202100077>.
- [21] Goto, H. *et al.* High-performance combinatorial optimization based on classical mechanics. *Science Advances* **7** (6), eabe7953 (2021). URL <https://www.science.org/doi/abs/10.1126/sciadv.abe7953>. <https://doi.org/10.1126/sciadv.abe7953>, <https://www.science.org/doi/pdf/10.1126/sciadv.abe7953>.
- [22] Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58** (1), 267–288 (1996). URL <http://www.jstor.org/stable/2346178>.
- [23] Ng, E. *et al.* Efficient sampling of ground and low-energy ising spin configurations with a coherent ising machine. *Phys. Rev. Research* **4**, 013009 (2022). URL <https://link.aps.org/doi/10.1103/PhysRevResearch.4.013009>. <https://doi.org/10.1103/PhysRevResearch.4.013009>.
- [24] Wang, Z., Marandi, A., Wen, K., Byer, R. L. & Yamamoto, Y. Coherent ising machine based on degenerate optical parametric oscillators. *Phys. Rev. A* **88**, 063853 (2013). URL <https://link.aps.org/doi/10.1103/PhysRevA.88.063853>. <https://doi.org/10.1103/PhysRevA.88.063853>.
- [25] Inui, Y. & Yamamoto, Y. Noise correlation and success probability in coherent ising machines (2020). [2009.10328](https://arxiv.org/abs/2009.10328).

- [26] Drummond, P. D. & Gardiner, C. W. Generalised p-representations in quantum optics. *Journal of Physics A: Mathematical and General* **13** (7), 2353 (1980). URL <https://dx.doi.org/10.1088/0305-4470/13/7/018>. <https://doi.org/10.1088/0305-4470/13/7/018> .
- [27] Zbontar, J. *et al.* fastmri: An open dataset and benchmarks for accelerated mri (2018). URL <https://arxiv.org/abs/1811.08839>.

Effective implementation of l_0 -Regularised Compressed Sensing with Chaotic-Amplitude-Controlled Coherent Ising Machines

Supplementary Information

Mastiyage Don Sudeera Hasaranga Gunathilaka, Satoshi Kako, Yoshitaka Inui, Kazushi Mimura, Masato Okada, Yoshihisa Yamamoto, and Toru Aonishi

Supplementary Note 1: Performance dependence on system size

This section compares how the performance of each model differs as the system size increases. In this case, N is in the range of $[200, 5000]$, with compression $\alpha = 0.6$ and sparseness $a = 0.2$. Fig. S1 illustrates the acquired results. Black, blue and red marks indicate the log average RMSE values acquired for 10 random hamiltonian simulations for OL-CIM-CDP, CAC-CIM-CDP (Wigner) and CAC-CIM-CDP (Positive- P) respectively.

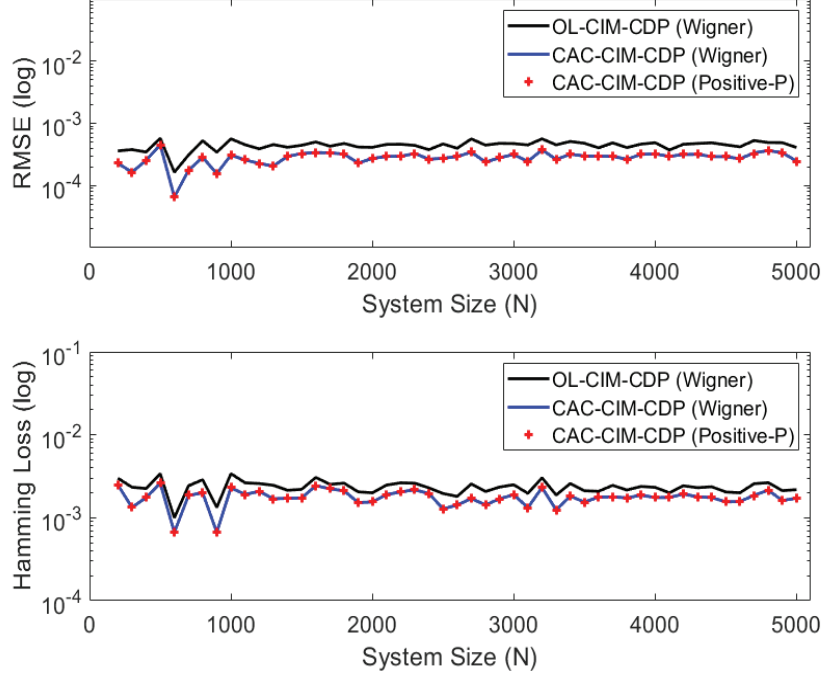


Figure S1: **Performance on different System Sizes**

(a) Log average RMSE values for 10 random Hamiltonians for each system size. For CAC-CIM-CDP (Wigner and Positive- P) $j = 1.0$, and $\beta = 1.0$ were used. η_{init} and η_{end} assigned as 0.9 and 0.01 respectively. For OL-CIM-CDP η_{init} and η_{end} were set as 0.6 and 0.01 respectively. (b) Log average Hamming loss values for 10 random Hamiltonians for each system size calculated according to Eq. (S1) which calculates the average support estimation accuracy.

As a measure of model performance, we have used average RMSE (Root Mean Squared Error) and average Hamming loss of support estimation from CIM. Using the following formula (S1), we calculated the hamming loss.

$$Hamming\ Loss = \frac{1}{N} \left(\sum_{r=0}^N |\sigma_r^{CIM} - \xi_r| \right). \quad (S1)$$

Here ξ_r corresponds to the correct support while σ_r^{CIM} is the estimation produced by CIM models. Based on the RMSE values in the top graph of Fig. S1, we can state that even when the size of the system changes, CAC-CIM-CDP models maintain a lower average RMSE value than OL-CIM-CDP. In addition, the corresponding hamming loss values (bottom graph of Fig. S1) are consistent also with the RMSE results. This suggests that CAC-CIM-CDP has performed a more accurate sparse reconstruction.

Thus, we can conclude that CAC-CIM-CDP provides better results than OL-CIM-CDP, regardless of system size.

Supplementary Note 2: Performance difference on relatively harder problem instances

Here we discuss the observation we had on the amplitude evolution of CAC-CIM-CDP models when the a/α ratio is higher (after exceeding the critical threshold).

The difference in performance is clear in normalised measured-amplitude $g\tilde{\mu}_r$ and error e_r as well. In Fig. S2 and Fig. S3 $N = 2000$ system is considered with $\alpha = 0.6$. Fig. S2a and Fig. S2b $a = 0.2$ is used while Fig. S2c and Fig. S2d use $a = 0.4$. It is clear that in the $a = 0.4$ case, the amplitude evolution is more chaotic compared to $a = 0.2$ in both models. The corresponding e_r evolution is illustrated in Fig. S3. Even though $a = 0.2$ has identical evolution for CAC-CIM-CDP, with $a = 0.4$ the evolution is slightly different.

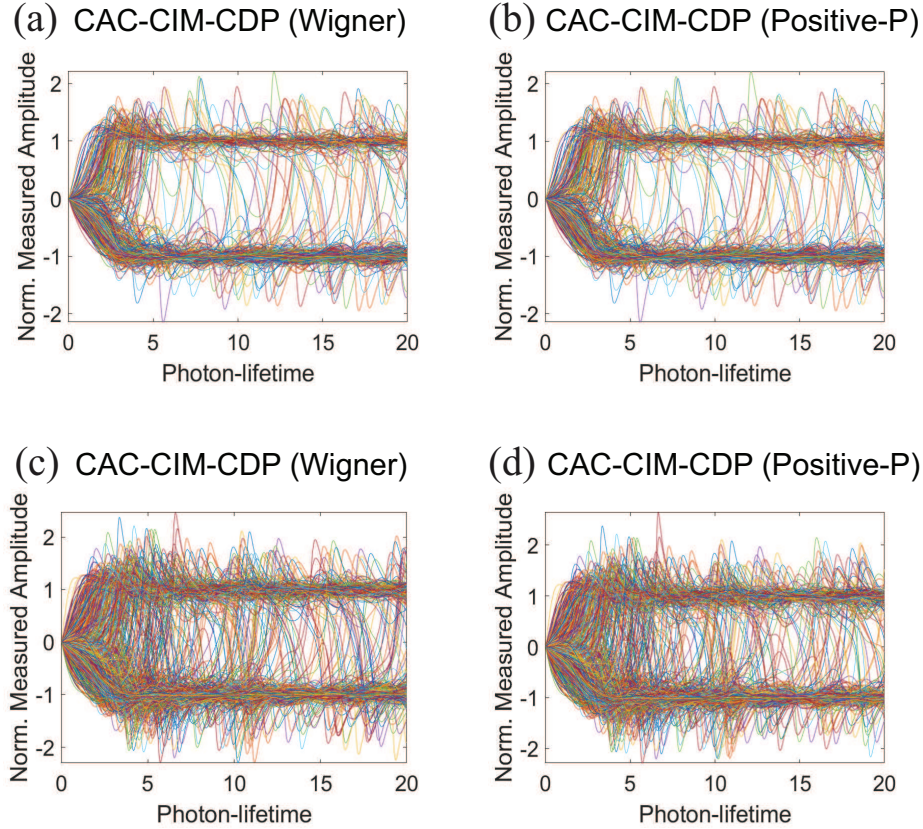


Figure S2: Normalised measured amplitude $g\tilde{\mu}_r$ evolution for easy and harder problem instances

(a) and (b) indicates the normalised amplitude $g\tilde{\mu}_r$ evolution of CAC-CIM-CDP (Wigner and Positive- P) with $a = 0.2$ and $\alpha = 0.6$. (c) and (d) indicates the normalised amplitude $g\tilde{\mu}_r$ evolution of CAC-CIM-CDP (Wigner and Positive- P) with $a = 0.4$ and $\alpha = 0.6$.

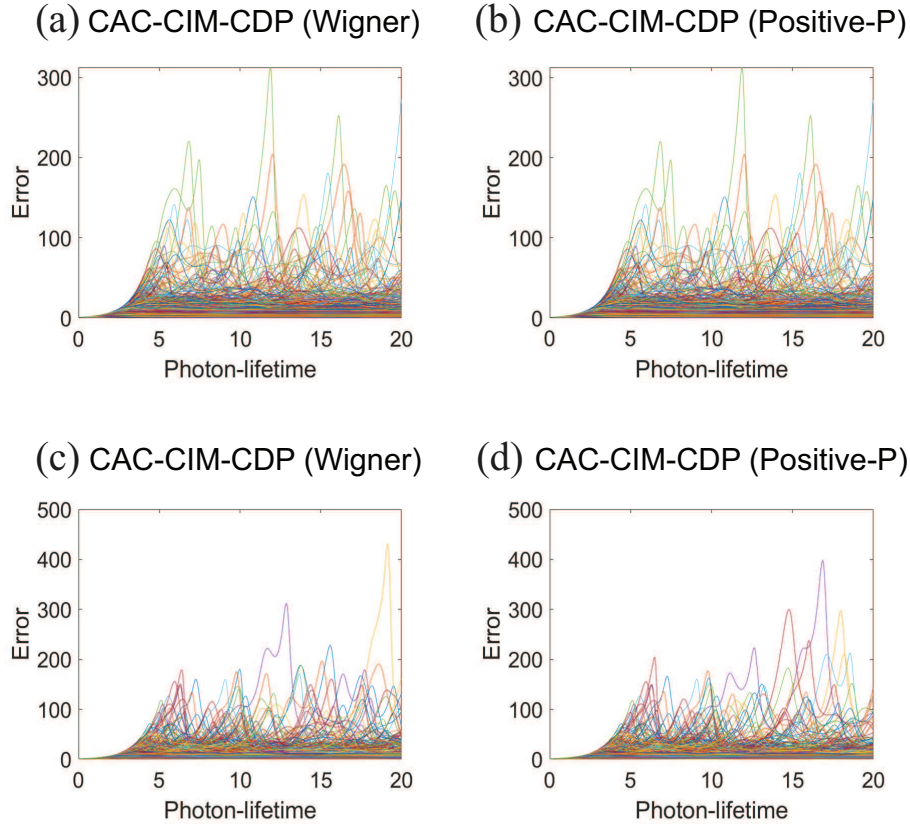


Figure S3: **Error e_r evolution for easy and harder problem instances**

(a) and (b) indicates the error amplitude e_r evolution of CAC-CIM-CDP (Wigner and Positive- P) with $a = 0.2$ and $\alpha = 0.6$. (c) and (d) indicates the error amplitude e_r evolution of CAC-CIM-CDP (Wigner and Positive- P) with $a = 0.4$ and $\alpha = 0.6$.

Supplementary Note 3: Performance with larger saturation parameters

With the same parameters used for CAC-CIM-CDP (Wigner and Positive- P), here we conducted an experiment on how the performance varies when the saturation parameter g^2 changes. Parameter g^2 corresponds directly to the quantum noise present in the CIM system. The higher the g^2 , the more quantum noise is present. In the MRI experiments mentioned previously, $g^2 = 10^{-7}$ was used. Which is quite small. In Fig. S4, we compare CIMs performance with $g^2 = 10^{-2}$, $g^2 = 10^{-3}$, and $g^2 = 10^{-7}$. Red lines indicate the CAC-CIM-CDP (Wigner) and the blue dashed line corresponds to the CAC-CIM-CDP (Positive- P) model. Vertical lines are the error bars for each threshold η value.

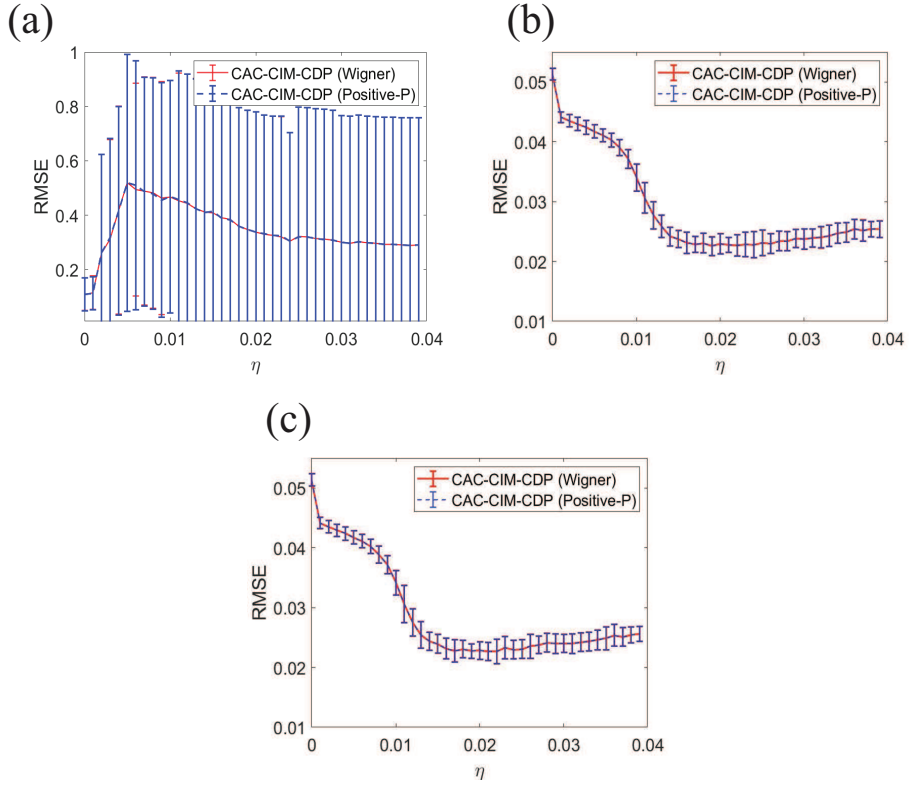


Figure S4: **Performance of CAC-CIM-CDP models with large quantum noise present**

(a), (b), and (c) illustrates the performance of CAC-CIM-CDP models when g^2 is 10^{-2} , 10^{-3} and 10^{-7} respectively for 64×64 image reconstruction.

Results indicate that both CAC-CIM-CDP models tend to perform similarly in all the cases. Yet when $g^2 = 10^{-2}$, reconstruction fails. Compared to other g^2 values (which have better reconstruction), it implies that for CIM performance there is a threshold saturation parameter value.

Supplementary Note 4: Performance change in sparseness

Here we consider different sparseness a values. For the simulations 64×64 images were used.

Reconstruction becomes harder when the a/α ratio becomes larger. In Fig. S5, we show-case the results on how the performance of LASSO, OL-CIM-CDP and the two CAC-CIM-CDP models change when a is increased for a constant α value with respect to η . The green boxplot indicates the CAC-CIM-CDP (Wigner) and the blue boxplot corresponds to the CAC-CIM-CDP (Positive- P) model. The black line and red boxplot indicate the LASSO algorithm and OL-CIM-CDP respectively. Based on different threshold values, box plots illustrate the maximum, minimum, 25th percentile (bottom edge), 75th percentile (top edge), and median (central horizontal line) of RMSEs for each model. In the markers, outliers are identified.

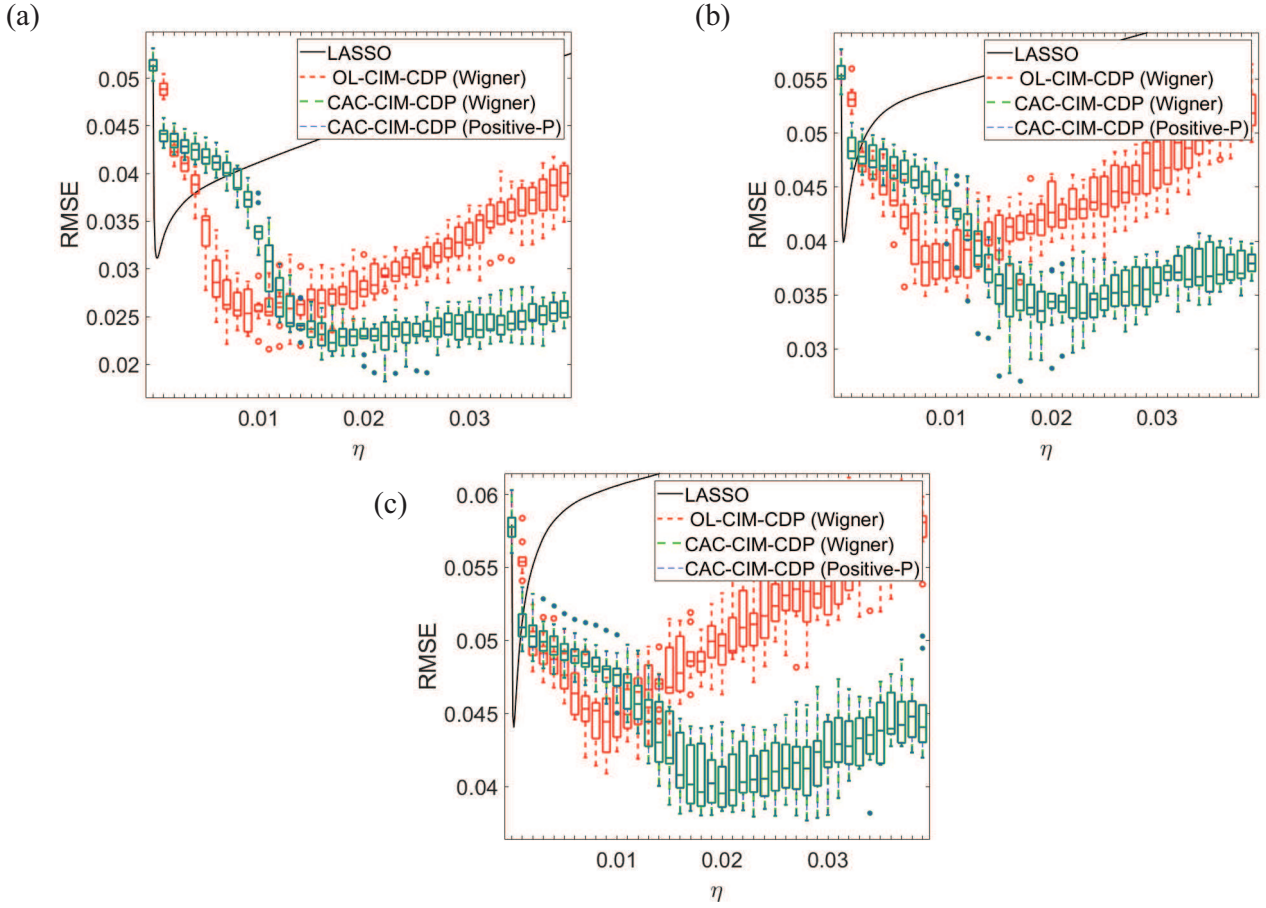


Figure S5: **Average performance of LASSO, OL-CIM-CDP, CAC-CIM-CDP (Wigner) and CAC-CIM-CDP (Positive- P) models when sparseness varies for 64×64 MRI image**

(a), (b), and (c) correspond to the reconstructions obtained from LASSO, OL-CIM-CDP, and CAC-CIM-CDP (Positive- P) with sparseness values 0.212, 0.255, and 0.280 respectively. For every model compression ratio was 0.4. The black line indicates the performance on LASSO while the red boxes correspond to OL-CIM-CDP. Green and blue boxes indicate the performance on CAC-CIM-CDP Wigner and Positive- P respectively. For different threshold values, the graphs illustrate the maximum, minimum, 25-th percentile (bottom edge), 75-th percentile (top edge), and median (central horizontal line) of RMSEs for each model with box plots. The markers indicate the outliers.

With the above results, it is clear that CAC-CIM-CDP models tend to perform better than OL-CIM-CDP. However, when a increases the reconstruction accuracy worsens in all three CIM-CDP models and LASSO. Even though the accuracy deteriorates, CAC-CIM-CDP models still perform better than OL-CIM-CDP. This indicates that even though the problem becomes more challenging, CAC-CIM-CDP is still a better CIM model than the OL-CIM-CDP for l_0 -regularised CS.