

## Article

# Boltzmann Sampling by Degenerate Optical Parametric Oscillator Network for Structure-Based Virtual Screening

Hiromasa Sakaguchi <sup>1,\*</sup>, Koji Ogata <sup>2</sup>, Tetsu Isomura <sup>3</sup>, Shoko Utsunomiya <sup>4</sup>, Yoshihisa Yamamoto <sup>5</sup> and Kazuyuki Aihara <sup>1,6</sup>

<sup>1</sup> Department of Electrical Engineering and Information Systems, The University of Tokyo, Tokyo 113-8654, Japan; aihara@sat.t.u-tokyo.ac.jp

<sup>2</sup> RIKEN Innovation Center, Saitama 351-0198, Japan; ogatak@riken.jp

<sup>3</sup> The KAITEKI Institute, Inc., Tokyo 100-8251, Japan; isomura.tetsu@mx.kaiteki-institute.com

<sup>4</sup> National Institute of Informatics, Tokyo 101-8403, Japan; shoko@nii.ac.jp

<sup>5</sup> ImPACT Program, The Japan Science and Technology Agency, Tokyo 102-0076, Japan; yoshihisa@stanford.edu

<sup>6</sup> Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan

\* Correspondence: sakaguchi@sat.t.u-tokyo.ac.jp; Tel.: +81-3-5452-6693

Academic Editor: Jay Lawrence

Received: 3 September 2016; Accepted: 11 October 2016; Published: 13 October 2016

**Abstract:** A structure-based lead optimization procedure is an essential step to finding appropriate ligand molecules binding to a target protein structure in order to identify drug candidates. This procedure takes a known structure of a protein-ligand complex as input, and structurally similar compounds with the query ligand are designed in consideration with all possible combinations of atomic species. This task is, however, computationally hard since such combinatorial optimization problems belong to the non-deterministic nonpolynomial-time hard (NP-hard) class. In this paper, we propose the structure-based lead generation and optimization procedures by a degenerate optical parametric oscillator (DOPO) network. Results of numerical simulation demonstrate that the DOPO network efficiently identifies a set of appropriate ligand molecules according to the Boltzmann sampling law.

**Keywords:** degenerate optical parametric oscillator (DOPO); structure-based drug design; coherent Ising machine; combinatorial optimization problem

## 1. Introduction

Structure-based drug design (SBDD) is a powerful tool for identifying lead compounds which can be further optimized or filtered to generate drug candidates [1–3]. When details of a 3D structure of a target protein are known, SBDD can identify suitable compounds, which have high binding affinity and selectivity with a target protein. Several structure-based virtual screening methods based on the lock-and-key model for protein-ligand interactions have succeeded in identifying hit compounds [4–6]. A lead optimization procedure from given fixed geometry has also been proposed [7,8]. In this method, all possible combinations of atomic species to given geometry are considered and ranked based on the score function which is an interaction energy between protein and ligand molecules. Since this procedure is a notorious combinatorial optimization problem, an exhaustive search is used to rank the combinations of atomic species in order to discover drug candidates.

Recently, there has been several attempts to solve computationally hard combinatorial optimization problems by using real physical systems rather than a conventional digital computer [9–11]. Such systems take the approach of physical simulation of Ising models, to which various combinatorial optimization problems can be mapped [12–14]. These systems are expected to solve combinatorial

optimization problems as a random sampling machine like Markov chain Monte Carlo (MCMC) simulation which is often used for machine learning [15–17]. The degenerate optical parametric oscillator (DOPO) network is one of such systems, which is recently proposed and demonstrated as a coherent Ising machine (CIM) [18–20]. The advantage of CIM is that any graph topology, including all to all connections, can be implemented without a time-consuming task of graph transformation. The performance of DOPO-based CIM is evaluated against NP-hard MAX-CUT problems which are equivalent to find the ground state of the Ising Hamiltonian without Zeeman term [18,21]. However, applications of CIM to real-world problems have not been proposed yet. Here, we show the mapping of the structure-based lead optimization problem onto CIM and present the benchmark simulation results by the DOPO network.

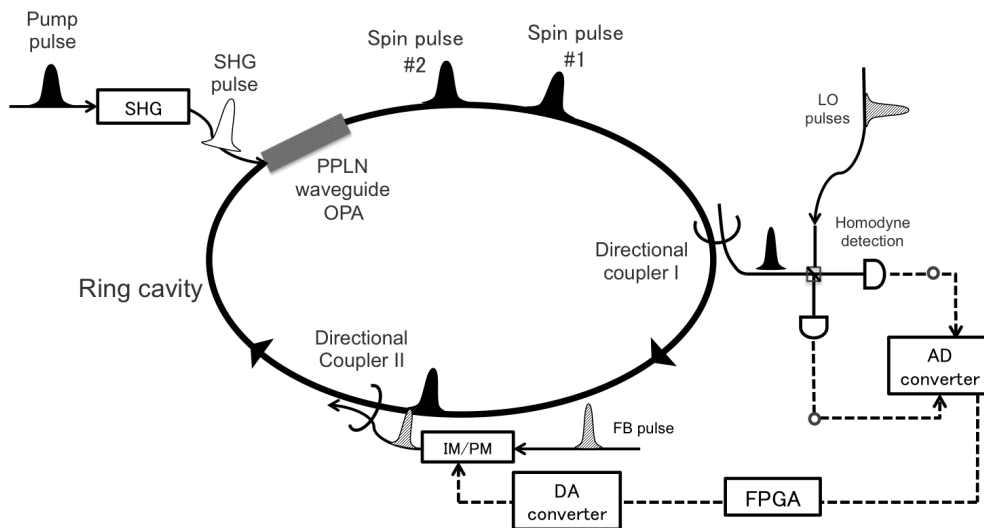
## 2. Coherent Ising Machine (CIM)

### 2.1. Computation Principle of CIM

CIM has been proposed to solve combinatorial optimization problems by exploiting the phase transition of a network of quantum oscillators. This machine searches the spin configuration  $\{\sigma_i\}$  which corresponds to the ground state of the Ising Hamiltonian  $H$  representing cost function of an optimization problem. The Ising Hamiltonian  $H$  is defined as

$$H = - \sum_{i < j} J_{ij} \sigma_i \sigma_j, \quad (1)$$

where  $\sigma_i \in \{\pm 1\}$  is a binary spin state of site  $i$  and interactions between spins are described by the coupling matrix  $\{J_{ij}\}$ . A time-division-multiplexed (TDM) pulsed DOPO with an optical fiber ring cavity can be a leading candidate for a scalable CIM [20].



**Figure 1.** Coherent Ising machine (CIM) based on time-division-multiplexed (TDM) pulsed degenerate optical parametric oscillator (DOPO) with measurement and feedback control. Both local oscillator (LO) pulses and feedback (FB) pulses are taken from the pump laser. A parametric gain is provided by a periodically-poled LiNbO<sub>3</sub> (PPLN) waveguide device and an optical ring cavity is formed by a fiber with  $\sim 1$  km length.

Figure 1 shows the basic configuration of the CIM based on TDM pulsed DOPO [21]. The Ising spin  $\sigma_i$  is represented by the in-phase amplitude  $c_i$  of the  $i$ -th DOPO pulse. The positive  $c_i$  value corresponds to up-spin ( $\sigma_i = +1$ ) while the negative  $c_i$  value corresponds to down-spin ( $\sigma = -1$ ). Each DOPO pulse is in a squeezed vacuum state at a pump rate below oscillation threshold, so that

the positive  $c_i$  (up-spin) and negative  $c_i$  (down-spin) co-exist as a superposition state. This allows the quantum parallel search for the ground state of the Ising Hamiltonian. The Ising coupling  $J_{ij}$  is implemented by optical homodyne detection and electric feedback control. The inferred value  $\tilde{c}_i$  by the optical homodyne measurement is used to compute an appropriate coupling field to other spin  $\sigma_j$  together with the initially memorized  $J_{ij}$  values in the digital circuit. An actual feedback pulse for the  $j$ -th pulse is produced by the optical amplitude/phase modulation based on the feedback circuit output for a part of the pump laser pulse which is phase-coherent with the DOPO pulse due to the phase sensitive amplification of DOPO.

The dynamics of the in-phase component  $c$  and quadrature-phase component  $s$  of the field of the  $i$ -th DOPO pulse is described as [22,23]

$$dc_i = (-1 + p - c_i^2 - s_i^2)c_i dt + \sum_j \xi_{ij} c_j dt + \frac{1}{A_s} \sqrt{c_i^2 + s_i^2 + \frac{1}{2}} dW_{c_i}, \quad (2)$$

$$ds_i = (-1 - p - c_i^2 - s_i^2)c_i dt + \sum_j \xi_{ij} s_j dt + \frac{1}{A_s} \sqrt{c_i^2 + s_i^2 + \frac{1}{2}} dW_{s_i}, \quad (3)$$

where  $\xi_{ij}$  is the coupling constant from the  $j$ -th to the  $i$ -th DOPO pulse and  $p$  is the pump rate. The last terms on the right hand side of the equations describe pump quantum noise and vacuum fluctuation incident from the open port of the output coupler.  $A_s$  is the amplitude of DOPO at a normalized pump rate  $p = 2$ , where  $p = 1$  corresponds to the threshold pump rate.  $dW_{c_i}$  and  $dW_{s_i}$  are independent zero-mean Gaussian noise processes. It is clear from Equation (3) that the quadrature-phase component of an each DOPO pulse decays exponentially and only the in-phase component builds up. At the steady state, the total saturated gain of the network  $\sum_i (p - c_i^2 - s_i^2)$  equals the overall photon decay rate  $\Gamma$ . If we ignore noise terms and quadrature-phase components, and assume  $|\xi_{ij}| \ll |c_i|$ , it can be shown analytically that  $\Gamma$  is described as [18]

$$\Gamma = N - \sum_{i,j} \xi_{ij} \sigma_i \sigma_j + O\left(\frac{\epsilon^3 N^4}{(p-1)^2}\right), \quad (4)$$

where  $\sigma_i = \text{sgn}(c_i)$ ,  $N$  is the number of DOPO pulses and  $\epsilon$  is the order of  $\xi_{ij}$ . If we ignore the higher order correction given by the third term of Equation (4) and exclude the constant term  $N$ , the overall photon decay rate  $\Gamma$  is identical to the Ising Hamiltonian. Therefore, an oscillation mode  $\{\sigma_i\}$  which achieves the minimum overall photon decay rate  $\Gamma$  provides the ground state of the Ising Hamiltonian.

## 2.2. Implementation of Zeeman Terms

The Ising Hamiltonian with a Zeeman term induced by an external magnetic field is represented as

$$H = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i, \quad (5)$$

where  $h_i$  is the local external field to spin  $i$ . In a wide variety of NP problems, they can be mapped to the Ising model with Zeeman terms [12]. Therefore we need to expand the argument of the previous subsection to the Ising model with Zeeman terms. This can be done by adding an extra term to the Equation (2) as

$$dc_i = (-1 + p - c_i^2)c_i dt + \sum_j \xi_{ij} c_j dt + h h_i \langle |c| \rangle dt + \frac{1}{A_s} \sqrt{c_i^2 + s_i^2 + \frac{1}{2}} dW_{c_i}, \quad (6)$$

where  $\langle |c| \rangle = (1/N) \sum_i |c_i|$  is the average absolute amplitude overall DOPO pulses at time  $t$  and  $h$  is a constant that determines the strength of the Zeeman term. Here, we ignore the dynamics of the

quadrature-phase component due to the reason mentioned above. From this equation, the overall photon decay rate is computed, which is equivalent to Equation (5) when  $h$  equals 1.

### 3. The Problem Hamiltonians

#### 3.1. Lead Optimization Procedure

The lead optimization procedure used in this work is based on the paper by Ogata et al. [8]. It is assumed that the coordinates of the compound structures, which are called geometry here, are extracted from a high resolution structure imaging of the protein-ligand complex. Then all the coordinates in the geometry are classified according to their bond order types ( $sp^3$ ,  $sp^2$ , etc.) and atomic species ( $CH_3$ ,  $CH_2$ ,  $CH$ ,  $NH_2$ ,  $NH$ , etc.). For instance, the geometry determined by a target protein is given as follows:  $X \cdots Y \cdots Z$ , where  $X$ ,  $Y$ , and  $Z$  represent the coordinates in the geometry and " $\cdots$ " is a generic representation of the bonds which connects the atomic species. Replacing  $Y$  with an atomic species " $=CH-$ " may generate a chemically incomplete compound  $X=CH-Z$  for which " $=$ " and " $-$ " indicate a double and a single bonds, respectively. Then,  $X$  must be assigned to an atomic species connected through a double bond to  $Y$  (e.g.,  $O=$  or  $CH_2=$ ). Similarly,  $Z$  must be assigned to an atomic species capable of linking to  $Y$  through a single bond (e.g.,  $-CH_3$  or  $-NH_2$ ). By assembling all possible combinations of these atomic species, four compounds are obtained:  $O=CH-CH_3$ ,  $O=CH-NH_2$ ,  $CH_2=CH-CH_3$  and  $CH_2=CH-NH_2$ .

After assigning atomic species to the geometry, the interaction energy between the obtained compounds and the target protein must be evaluated. The cost function, including an interaction energy and solvation energy, is calculated for each atomic species separately as

$$E_{in} = E_{in}^v + E_{in}^e + E_{in}^h + E_{in}^s, \quad (7)$$

where  $i$  is the position to which an atomic species is assigned and  $n$  is the atomic species.  $E_{in}^v$  is the van der Waals interaction energy between the atomic species and the target protein.  $E_{in}^e$  is the electrostatic interaction energy between the two,  $E_{in}^h$  is the hydrogenic bond energy and  $E_{in}^s$  is the solvation energy. The total cost function is represented as  $E_{total} = \sum_i E_{in_i}$ , where  $n_i$  is the atomic species which is assigned to the position  $i$ . The details of the interaction energy are described in the report by Ogata et al. [8]. In the original method, this cost function is calculated for all possible combinations of atomic species and all such ligand compounds are ranked in the ascending order according to the cost function value. Since this approach experiences combinatorial explosion as the problem size becomes large, we take the approach to sample only ligand compounds with low cost function values from all possible combinations. Note that there are  $\sim 3 \times 10^{12}$  ligand compounds even for such a small problem size with  $N_{position} = 16$  and  $N_{atom} = 6$  atomic species where  $N_{position}$  and  $N_{atom}$  are the numbers of sites and atomic species, respectively.

#### 3.2. Mapping to the Ising Hamiltonian

The problem considered above is summarized in the following constrained optimization problems: From given geometry and atomic species, sample the low energy assignments  $\{i, n\}$  of atomic species to geometry satisfying the constraints of the bond order. We can map the problem to the Ising Hamiltonian with Zeeman terms as follows.

Let us consider a binary variable  $x_{in} \in \{0, 1\}$  with position  $i$  on the geometry and atomic species  $n$ :

$$x_{in} = \begin{cases} 1 & \text{if atomic species } n \text{ is assigned to position } i, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

When the total number of atomic species and positions are  $N_{\text{atom}}$  and  $N_{\text{position}}$ , respectively, the number of variables  $x_{in}$  required to compute the cost function is  $N = N_{\text{atom}}N_{\text{position}}$ . The penalty between  $x_{in}$  and  $x_{jm}$  is now defined as

$$J_{in;jm} = \begin{cases} 1 & \text{if position } i \text{ is adjacent to } j \text{ and the bond order between } n \text{ and } m \text{ are inconsistent,} \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

The total cost function with  $x_{in}$  and  $J_{in;jm}$  for the considered problem is described by

$$H = A \sum_i \left( \sum_n x_{in} - 1 \right)^2 + B \sum_{i \neq j} \sum_{n,m} J_{in;jm} x_{in} x_{jm} + C \sum_{i,n} E_{in} x_{in}, \quad (10)$$

where  $A$ ,  $B$  and  $C$  are adjustable constants which determine the contribution weights of each term in the cost function and  $E_{in}$  is defined by Equation (7). The first term in the equation comes from the constraint that exactly one atomic species must be assigned to each position. The second term denotes the constraint of the bond order. The first and second terms must be zero for a legitimate compound satisfying the two constraints. Finally, the third term is the interaction energy which we want to minimize under the two constraints. We appropriately choose the parameters  $A$ ,  $B$  and  $C$ , to efficiently find the ground states on low-energy states of the Hamiltonian under the condition of satisfying the constraints.

Since the above formulation (Equation (10)) uses the binary variable  $x_{in} \in \{0, 1\}$ , we need to transform the variable  $x_{in}$  to  $\sigma_{in} \in \{-1, 1\}$  by the relation  $x_{in} = (\sigma_{in} + 1)/2$  for mapping the cost function on the CIM. Equation (10) can be then represented by

$$H = \sum_{i,j,n,m} \tilde{J}_{in;jm} \sigma_{in} \sigma_{jm} + \sum_{i,n} h_{in} \sigma_{in} + \text{const.}, \quad (11)$$

where

$$\tilde{J}_{in;jm} = \frac{A}{4} \delta_{ij} + \frac{B}{4} (1 - \delta_{ij}) J_{in;jm}, \quad (12)$$

$$h_{in} = \frac{A}{2} (N_{\text{atom}} - 2) + \frac{B}{2} \sum_{j,m} J_{in;jm} + C E_{in}. \quad (13)$$

Since the constant term doesn't change the energy landscape, we can neglect the third term of Equation (11). The resulting cost function is the Ising Hamiltonian with Zeeman terms, so that we can implement it directly on the CIM.

### 3.3. Several Heuristic Modifications

A CIM based on the DOPO network simulates an Ising model and reports not only the ground states but also low-energy excited states. DOPO pulses, which represent Ising spins  $\sigma_i \in \{-1, 1\}$ , can actually take continuous values as the amplitude of the in-phase component  $c$ , whereas the original Ising spins take two discrete spin states. It has been found that CIM solves MAX-CUT problems without a Zeeman term with high performance [18,21]. However, our problem has a Zeeman term and the goal is not only finding the ground states with an exact minimum energy, but rather sampling many low-energy states which can stably bind to the target protein and thus become drug candidates.

Here, we introduce three heuristic methods to improve the performance of the CIM to solve the given problems. A first heuristic method is the gradual pumping where the pump parameter  $p$  is gradually increased from below to above the oscillation threshold. The pump parameter  $p$  acts as a gain, and when there are no mutual couplings, the amplitude of each DOPO pulse at the steady state is proportional to  $\sqrt{p-1}$  for  $p > 1$ . The gradual pumping makes quantum noise effects gradually

weaker as the amplitude grows. In this paper, pump  $p$  is increased linearly with time until some limit value.

As a second heuristic method, we replace the amplitude of the coupling field in Equation (6) as

$$\sum_j \xi_{ij} c_j \rightarrow \langle |c| \rangle \sum_j \xi_{ij} \frac{c_j}{|c_j|}. \quad (14)$$

By this replacement, each DOPO pulse experiences the effect of interaction as if all other DOPO pulses are oscillating with the same amplitude. This modification allows us to avoid the breakdown of the mapping protocol due to the inhomogeneous DOPO amplitudes.

Finally, we consider the relaxation of  $\{J_{in;jm}\}$  (Equation (9)) to a continuous value. When the CIM solves an Ising problem with a Zeeman term, we add it with a constant  $h$  as described in Equation (6). The parameter  $h$  empirically tunes the balance between the Ising interaction term and the Zeeman term. If there are large variation in Zeeman coefficients  $h_{in}$ , the Zeeman Hamiltonian dominates over the Ising Hamiltonian and CIM picks up biased solutions on the same Zeeman energy surface. In our problem, the large variation in  $h_{in}$  is caused by the term  $\sum_{j,m} J_{in;jm}$ . To avoid this situation, we introduce the relaxation of  $\{J_{in;jm}\}$  to a continuous value while keeping symmetry, zero elements and non-zero elements with positive real values so that the second term of Equation (13) satisfies

$$\sum_{j,m} J_{in;jm} = \frac{1}{N} \sum_{i,j,n,m} J_{in;jm}, \text{ for } \forall i, n. \quad (15)$$

Since  $J_{in;jm}$  is only related to the second constraint which should be zero for legitimate states, this relaxation doesn't affect the energy of solutions. If we represent  $\mathbf{J}$  as the vector whose elements are non-zero elements of  $\{J_{in;jm}\}$ , the problem is finding  $\mathbf{J}$  which satisfies

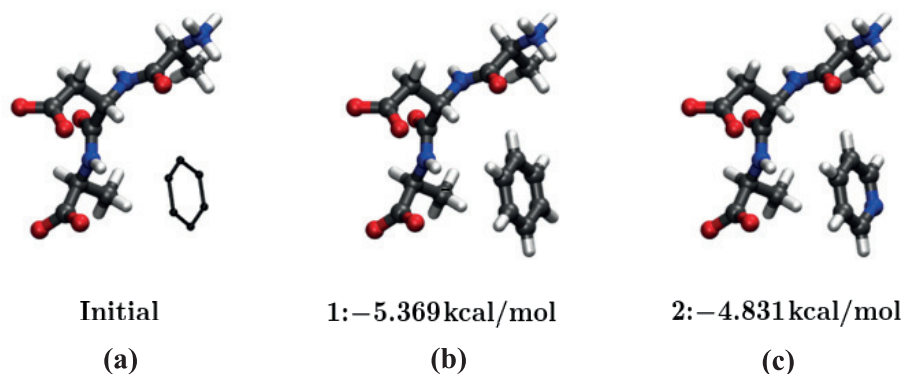
$$\mathbf{W}\mathbf{J} = \mathbf{1} \cdot \left( \frac{1}{N} \sum_{i,j,n,m} J_{in;jm} \right), \quad (16)$$

where  $\mathbf{1}$  is a constant vector whose elements are one,  $\mathbf{W}$  is an  $N_{\text{nonzero}} \times N$  matrix which specifies the locations of non-zero elements of  $\{J_{in;jm}\}$  and  $N_{\text{nonzero}}$  is the number of non-zero elements of  $\{J_{in;jm}\}$ . There are infinite ways of this relaxation because  $N_{\text{nonzero}}$  is larger than  $N$  in general. In this paper, we use  $\mathbf{J}$  which is found by calculating the Moore-Penrose pseudoinverse of  $\mathbf{W}$  [24].

#### 4. Numerical Simulation Results

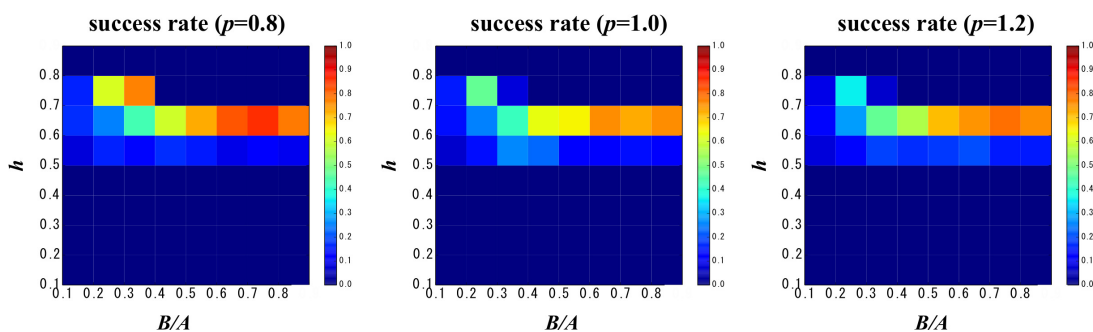
We tested our proposed protocol including heuristic modifications by the numerical simulation for the simple benchmark problem. The benchmark we used here is a 6-membered ring placing near the artificially generated Ala-Asp-Ala tripeptide (Figure 2a). To generate compounds, six types of atomic species,  $=\text{CH}-$ ,  $-\text{CH}=$ ,  $-\text{NH}-$ ,  $=\text{N}-$ ,  $-\text{N}=$  and  $-\text{O}-$ , are assigned to the ring. Therefore, the number of spins  $N$  which CIM needs to solve the problem is  $N = N_{\text{atom}}N_{\text{position}} = 36$ , where  $N_{\text{atom}} = N_{\text{position}} = 6$ . The interaction energies for all possible compounds are calculated in advance. Benzene and pyridine are detected as the configuration with the lowest and 2nd lowest energies, respectively, in 0.53 kcal/mol score difference (Figure 2b,c). The details of the interaction energy  $E_{in}$  and  $J_{in;jm}$  of this benchmark are available in the Supplementary Materials.





**Figure 2.** 6-membered ring placing near Ala-Asp-Ala tripeptide. (a) initial structure; (b) benzene with the lowest energy and (c) pyridine with the 2nd lowest energy.

At first, we tested our constraint Hamiltonian without the score term, i.e., the first and second terms of Equation (10) without the interaction energy ( $C = 0$ ), to investigate the parameter dependence of the probability of satisfying the constraints. Figure 3 shows the success probability of satisfying the constraints over such parameters as pump limit  $p$ , the weight of Zeeman term  $h$  and the ratio of the bond constraint over the position constraint  $B/A$ . The success probability depends strongly on  $h$  rather than  $B/A$  and  $p$ . The optimal value for the weight of the Zeeman term is  $h \sim 0.6$ .



**Figure 3.** The success probability of satisfying the constraints for 1000 identical trials. The parameter  $p$  is the final pump rate for gradual pumping. The parameters of the Hamiltonian (Equation 10) are set to  $A = 1$  and  $C = 0$  for all the results.

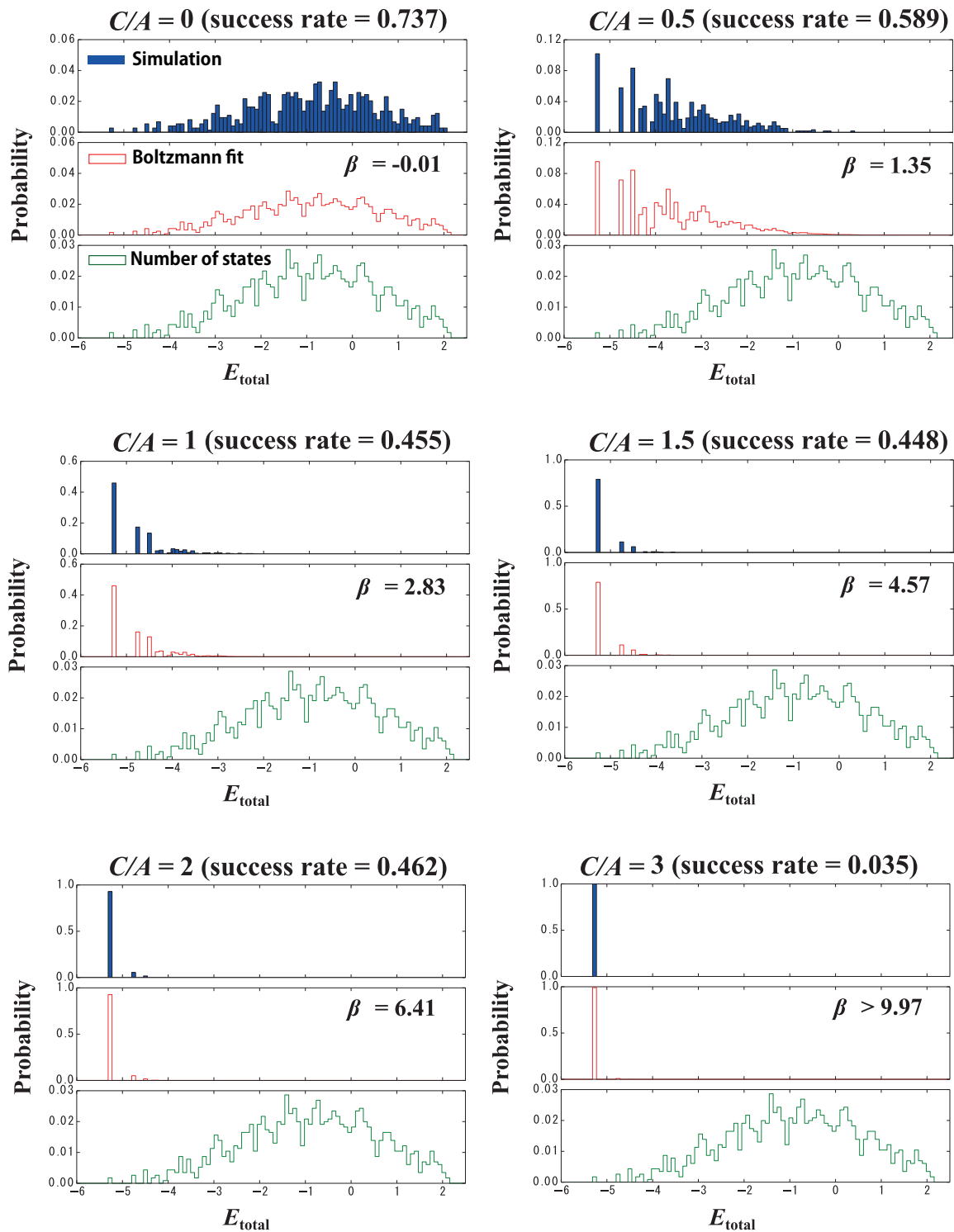
Next, we investigate the statistics for the interaction energy, the third term of Equation (10), with  $p = 1$ ,  $h = 0.6$  and  $B/A = 0.6$ . We plotted the histograms of the interaction energy for various  $C/A$  values only for solutions which satisfy the constraints, given by the first and second terms of Equation (10), in Figure 4. The corresponding Boltzmann distribution at properly chosen temperature is also plotted in red. If we represent the Ising Hamiltonian of Equation (10) as  $H(\mathbf{x}) = H_{\text{cons}}(\mathbf{x}) + CH_{\text{score}}(\mathbf{x})$ , the Boltzmann distribution which satisfy the constraints is represented by

$$P(H_{\text{score}} = E | H_{\text{cons}} = 0) = \frac{D(E) \exp(-\beta_n E)}{Z}, \quad (17)$$

$$Z = \sum_i D(E_i) \exp(-\beta_n E_i), \quad (18)$$

where  $E$  is the interaction energy,  $\beta_n = C/k_B T$  is the normalized temperature parameter,  $k_B$  is the Boltzmann constant,  $T$  is the temperature and  $D(E)$  is the number of states at the energy  $E$ . Changing the parameter  $C$  corresponds to changing the normalized temperature parameter for the Boltzmann distribution. We estimated the normalized temperature parameter for each  $C/A$  by the least squares

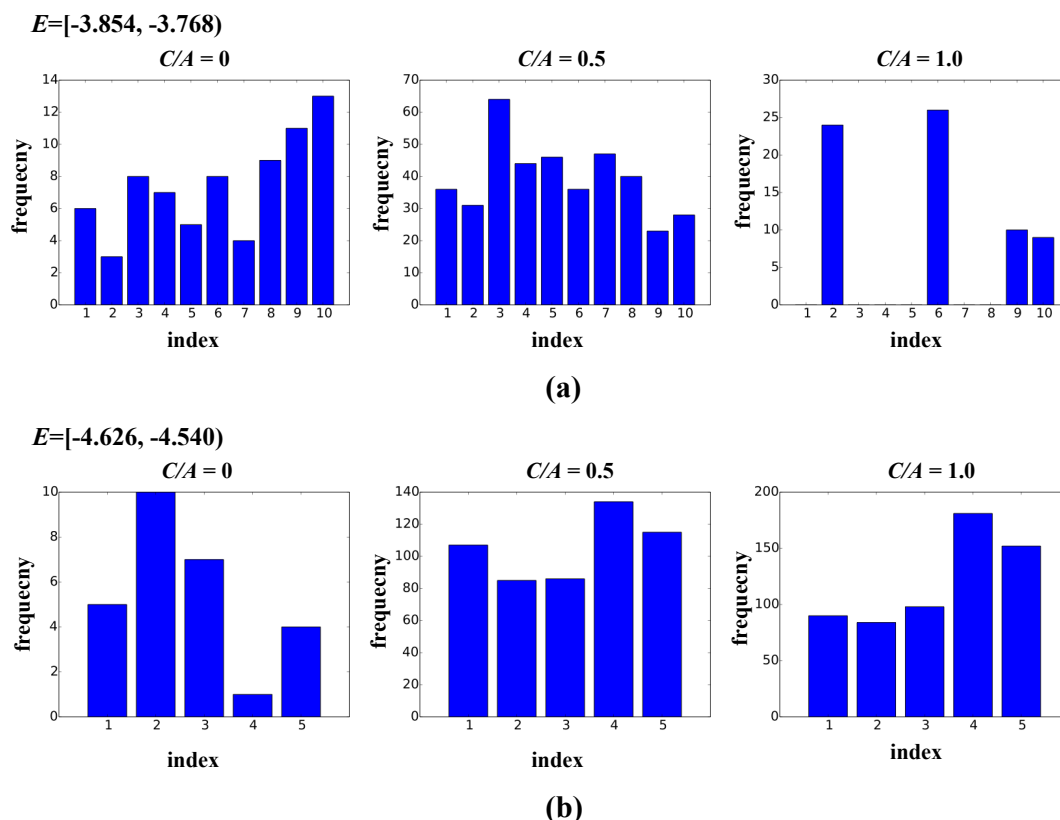
method. The observed distributions are well matched with the Boltzmann distributions although the success probability of satisfying the constraints decreases as  $C$  becomes large.



**Figure 4.** Histograms of the interaction energies of the final states of CIM over 1000 runs. The blue bar is the simulation result and the red dot is the estimated Boltzmann distribution. The green line at the bottom of each figure shows the number of states for the given Hamiltonian. All the histograms are normalized to 1.



Finally, we plotted the histograms for finding degenerate states at the energy interval  $E = [-3.854, -3.768)$  and  $E = [-4.626, -4.540)$  in Figure 5 in order to investigate the random sampling capability for the solutions on the same energy surface. All degenerate states are found with almost equal probability at  $C/A = 0.5$ . Since our purpose is to sample all possible compound configurations with low interaction energies, we confirmed that the optimal parameter of  $C/A$  is  $\sim 0.5$  in terms of both low energy temperature and random sampling capabilities.



**Figure 5.** The histogram of finding the degenerate states on the same energy surface over 1000 runs. (a)  $E = [-3.854, -3.768)$ ; (b)  $E = [-4.626, -4.540)$ .

## 5. Conclusions

In this paper, we proposed the lead optimization protocol for the CIM and tested the method for a simple benchmark problem. This is the first benchmark study of the DOPO-based CIM that the problem Hamiltonian includes Zeeman terms. We confirm that the CIM satisfies the constraints with high success probability and the resulting histograms of obtaining various compound configurations are well matched with the Boltzmann distribution with low energy temperature. The normalized temperature parameter of the Boltzmann distribution can be tuned by changing the weight  $C$  of the interaction Hamiltonian.

We also investigated randomness of the solutions on the same energy surface. All the solutions in the same energy interval are covered with almost identical probability. Since exact randomness and the Boltzmann distribution are not needed for our purpose, the present results were acceptable.

We do not need the same solutions which are already obtained, so that CIM should pick up different solutions in subsequent runs. There are a lot of similar, further tricks that we can employ to improve the performance of our method [25].

**Supplementary Materials:** The data which we used for the benchmark study is available online at [www.mdpi.com/1099-4300/18/10/365/s1](http://www.mdpi.com/1099-4300/18/10/365/s1), Table S1: Interaction Energy  $E_{in}$  of benchmark problem, Table S2:  $J_{in,jm}$  of benchmark problem.

**Acknowledgments:** This work was funded by the Impulsing Paradigm Change through Disruptive Technologies (ImPACT) Program of the Council of Science, Technology and Innovation.

**Author Contributions:** Hiromasa Sakaguchi and Koji Ogata conceived the mapping of the lead optimization procedure. Hiromasa Sakaguchi, Koji Ogata, Tetsu Isomura, Shoko Utsunomiya, Yoshihisa Yamamoto and Kazuyuki Aihara designed the study. Hiromasa Sakaguchi and Yoshihisa Yamamoto conceived the heuristic procedures of CIM. Hiromasa Sakaguchi performed the numerical experiments. Koji Ogata created the benchmark problem. All authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bohacek, R.S.; McMartin, C.; Guida, W.C. The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.* **1996**, *16*, 3–50.
2. Anderson, A.C. The process of structure-based drug design. *Chem. Biol.* **2003**, *10*, 787–797.
3. Lounnas, V.; Ritschel, T.; Kelder, J.; McGuire, R.; Bywater, R.P.; Foloppe, N. Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Comput. Struct. Biotechnol. J.* **2013**, *5*, 1–14.
4. Yamazaki, K.; Kusunose, N.; Fujita, K.; Sato, H.; Asano, S.; Dan, A.; Kanaoka, M. Identification of phosphodiesterase-1 and 5 dual inhibitors by a ligand-based virtual screening optimized for lead evolution. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1371–1379.
5. Yan, S.; Appleby, T.; Larson, G.; Wu, J.Z.; Hamatake, R.K.; Hong, Z.; Yao, N. Thiazolone-acylsulfonamides as novel HCV NS5B polymerase allosteric inhibitors: Convergence of structure-based drug design and X-ray crystallographic study. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1991–1995.
6. Carosati, E.; Mannhold, R.; Wahl, P.; Hansen, J.B.; Fremming, T.; Zamora, I.; Cianchetta, G.; Baroni, M. Virtual screening for novel openers of pancreatic K(ATP) channels. *J. Med. Chem.* **2007**, *50*, 2117–2126.
7. Ogata, K.; Isomura, T.; Yamashita, H.; Kubodera, H. A quantitative approach to the estimation of chemical space from a given geometry by the combination of atomic species. *Mol. Inform.* **2007**, *26*, 596–607.
8. Ogata, K.; Isomura, T.; Kawata, S.; Yamashita, H.; Kubodera, H.; Wodak, S.J. Lead generation and optimization based on protein-ligand complementarity. *Molecules* **2010**, *15*, 4382–4400.
9. Johnson, M.W.; Amin, M.H.S.; Gildert, S.; Lanting, T.; Hamze, F.; Dickson, N.; Harris, R.; Berkley, A.J.; Johansson, J.; Bunyk, P.; et al. Quantum annealing with manufactured spins. *Nature* **2011**, *473*, 194–198.
10. Yamaoka, M.; Yoshimura, C.; Hayashi, M.; Okuyama, T.; Aoki, H.; Mizuno, H. 20k-spin Ising Chip for Combinational Optimization Problem with CMOS Annealing. In Proceedings of the IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 22–26 February 2015.
11. Utsunomiya, S.; Takata, K.; Yamamoto, Y. Mapping of Ising models onto injection-locked laser systems. *Opt. Express* **2011**, *19*, 18091–18108.
12. Lucas, A. Ising formulations of many NP problems. *Front. Phys.* **2014**, *2*, doi:10.3389/fphy.2014.00005.
13. Rieffel, E.G.; Venturelli, D.; O’Gorman, B.; Do, M.B.; Prystay, E.M.; Smelyanskiy, V.N. A case study in programming a quantum annealer for hard operational planning problems. *Quantum Inf. Process.* **2014**, *14*, 1–36.
14. Perdomo-Ortiz, A.; Dickson, N.; Drew-Brook, M.; Rose, G.; Aspuru-Guzik, A. Finding low-energy conformations of lattice protein models by quantum annealing. *Sci. Rep.* **2012**, *2*, 571.
15. Denil, N.; Fretias, N. Toward the Implementation of a Quantum RBM. In Proceedings of the NIPS 2011 Deep Learning and Unsupervised Feature Learning Workshop, Cranada, Spain, 16 December 2011.
16. Dumoulin, V.; Goodfellow, I.J.; Courville, A.; Bengio, Y. On the Challenges of Physical Implementations of RBMs. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec, QC, Canada, 27–31 July 2014; pp. 1199–1205.
17. Adachi, S.H.; Henderson, M.P. Application of Quantum Annealing to Training of Deep Neural Networks. **2015**, arXiv:1510.00635.
18. Wang, Z.; Marandi, A.; Wen, K.; Byer, R.L.; Yamamoto, Y. Coherent Ising machine based on degenerate optical parametric oscillators. *Phys. Rev. A* **2013**, *88*, 063853.
19. Marandi, A.; Wang, Z.; Takata, K.; Byer, R.L.; Yamamoto, Y. Network of Time-Multiplexed Optical Parametric Oscillators as a Coherent Ising Machine. *Nat. Photonics* **2014**, *8*, 937–942.

20. Inagaki, T.; Inaba, K.; Hamerly, R.; Inoue, K.; Yamamoto, Y.; Takesue, H. Large-scale Ising spin network based on degenerate optical parametric oscillators. *Nat. Photonics* **2016**, *10*, 415–419.
21. Haribara, Y.; Utsunomiya, S.; Yamamoto, Y. Computational principle and performance evaluation of coherent Ising machine based on degenerate optical parametric oscillator network. *Entropy* **2016**, *18*, 151.
22. Takata, K.; Marandi, A.; Yamamoto, Y. Quantum correlation in degenerate optical parametric oscillators with mutual injections. *Phys. Rev. A* **2015**, *92*, 043821.
23. Maruo, D.; Utsunomiya, S.; Yamamoto, Y. Truncated Wigner theory of coherent Ising machines based on degenerate optical parametric oscillator. *Phys. Scr.* **2016**, *91*, 083010.
24. Penrose, R. A generalized inverse for matrices. *Math. Proc. Camb. Philos. Soc.* **1955**, *51*, 406–413.
25. Molnár, B.; Ercsey-Ravasz, M. Asymmetric Continuous-Time Neural Networks without Local Traps for Solving Constraint Satisfaction Problems. *PLoS ONE* **2013**, *8*, e73400.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).