Derivaton of QUBO formulations for sparse estimation

Tomohiro Yokota[1], Makiko Konoshima[2], Hirotaka Tamura[2], Jun Ohkubo[1,3]

[1]*Graduate School of Science and Enginnering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, 338-8570, Japan*
[2]*Fujitsu Laboratories Ltd., 4-1-1 Kawasaki, Kanagawa 211-8558, Japan*
[3]*JST, PREST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

We propose a quadratic unconstrained binary optimization (QUBO) formulation of the ℓ_1-norm, which enables us to perform sparse estimation in the Ising-type annealing methods including quantum annealing. The QUBO formulation is derived via the Legendre transformation and the Wolfe theorem, which have recently been employed in order to derive the QUBO formulations of ReLU-type functions. Furthermore, it is clarified that a simple application of the derivation method to the ℓ_l-norm case gives a redundant variable; finally a simplified QUBO formulation is obtained by removing the redundant variable.

arXiv:2001.03715v1 [quant-ph] 11 Jan 2020

1. Introduction

In recent years, some novel computing hardwares have been developed and actually provided; there are some Ising-type annealing machines such as "D-Wave 2000" by the Canadian company D-Wave[1,2] and "Fujitsu's Digital Annealer" by the Japanese company Fujitsu, which is a physically inspired annealing accelerator based on classical (i.e., non-quantum) logic circuits.[3] The annealing machines are used to obtain approximate solutions for optimization problems; the optimization problems play important roles in various research areas including data mining and machine learning. Especially, the quantum annealing method has been originally proposed by Kadowaki and Nishimori,[4] and a similar idea called adiabatic quantum computing[5] has attracted many attentions; recently, researches for practical applications have been performed (for example, see the paper by Tanahashi et al.[6]) Discussions for machine learning and quantum Boltzmann machines were given by Biamonte et al.,[7] and there are many challenging tasks from the viewpoints of hardware and software. Although one of the restrictions is that the system size is small, the number of available qubits (or classical bits) has been increasing year by year, which enables us to tackle practical and large optimization problems.

The above annealing hardwares need quadratic unconstrained binary optimization (QUBO) formulations; the hardwares are based on the Ising-type Hamiltonian, and hence it is necessary to convert original cost functions in optimization problems into the QUBO formulations. (The QUBO formulation is equivalent to the Ising model.) Although continuous variables can be expressed as the Ising-type variables via adequate binary-expansions, it is not straightforward in general to reformulate the original cost functions as the QUBO formulations. Some reformulations were given in the review paper by Lucas,[8] and it has been shown that logic gates are expressed in the form of the QUBO formulations.[9] However, a systematic way to derive the QUBO formulations has not been found yet. Recently, the Legendre transformation was employed to derive the QUBO formulation of the q-loss function;[10] the q-loss function was proposed as a cost function with robust characteristics against the label noise in machine learning. The derivation technique based on the Legendre transformation revealed that some mathematical transformation would be needed to transform some types of cost functions into the QUBO formulations. Actually, it has been clarified that the Legendre transformation is not enough to deal with the Rectified Linear-Unit (ReLU) type functions;[11] the Wolfe duality theorem[12] was employed to derive the QUBO formulations for the ReLU-type functions. These works also indicate the fact that the derivation of the QUBO formulations is not straightforward, and we sometimes need further and careful considerations which depend on the original cost functions.

As shown above, there are some works to derive the QUBO formulations for machine learning problems. Of course, there are many other research fields related to optimization problems, and one of them is data analysis (data mining). It has been known that a concept of "regularization" plays an important role in data analysis and, of course, machine learning. For example, a ℓ_2 norm is widely used; a linear regression with the ℓ_2 regularization is called the ridge regression, and it is widely used for various practical issues. In addition, a ℓ_1 norm is also an useful example of regularization, which is used in order to introduce a kind of "sparseness" for the solution; the sparse estimation is one of the hot topics in the research field of data analysis. The least absolute shrinkage and selection operator (LASSO)[13] is a famous practical method to achieve sparse estimations, in which the ℓ_1 norm is added to a least-squares cost function. Recently, the idea of the sparse estimation has been applied to the black hole analysis.[14] The data size of the black hole is so small, and hence it is difficult to observe the image of the black hole directly because of the low-resolution of images. Therefore, simultaneous measurements from radio telescopes all over the world were performed, and the method based on the sparse estimation was applied to the observed big data, by which only essential information is extracted; finally the imaging of the black holes was achieved. Note that the ℓ_2 norm is simply connected to the QUBO formulation because of the quadratic form; in contrast, the ℓ_1 norm has a non-differentiable point, and the QUBO formulation has not been derived yet.

In this paper, the QUBO formulation of the ℓ_1 norm is derived. In order to obtain the QUBO formulation, both the Legendre transformation and the Wolfe duality theorem are employed. Furthermore, it is clarified that only the simple applications of the previous derivation techniques are not enough; through numerical checks and reconsideration for the derived formulation, a simplified QUBO formulation is finally derived. The simplified QUBO formulation has a smaller number of variables than the naive derived formulation; such reduction of the number of variables is important for the hardware implementation because the current Ising-type hardwares have only restricted number of qubits (or classical bits).

The construction of this paper is as follows. Section 2 explains the QUBO formulation and related previous works. The important techniques for the derivations are also given for later use. In Sect. 3, the QUBO formulation of the ℓ_1 norm is derived, and the numerical checks are given. Section 4 gives the main result of this paper; a simplified version of the QUBO form is given, in which a variable is removed from the naive QUBO formulation derived in Sect. 3. Section 5 gives concluding remarks and future works.

2. Backgrounds and preliminaries

The aim of this paper is to derive the QUBO formulation of the ℓ_1 norm-type function. As denoted in the Introduction, the QUBO formulation of a little complicated function, q-loss function, has already been derived.[10] One may think that a kind of combination of the q-loss functions could be used as regularization functions, but the q-loss function is not enough to make the ℓ_1 norm. Based on the derivation techniques used in the q-loss case, the QUBO formulation of the ReLU-type function was discussed.[11] In Ref. 11, two techniques, i.e., the Legendre transformation and the Wolfe duality theorem, were employed, which also play important roles in our discussions. Hence, in this section, we briefly denote some background knowledges and previous works, starting from a brief explanation for QUBO formulations and Ising model. Especially, the Legendre transformation and the Wolfe duality theorem will be concisely denoted.

2.1 QUBO and Ising model

As denoted in the Introduction, the Ising-type annealing machines need the Ising Hamiltonian or the QUBO formulation in order to solve combinatorial optimization problems. The QUBO form has binary variables, which take only 1 or 0, and the 0-1 binary variables are sometimes suitable to consider the combinatorial optimizations; the binary expansions of continuous variables naturally introduce their binary expressions. Of course, since the QUBO formulation and the Ising model are equivalent, it is possible to convert the QUBO form into the Ising model, and vice versa. The Ising model is represented as follows:

$$H = -\sum_{i,j} J_{ij}\sigma_i\sigma_j - \sum_i h_i\sigma_i, \tag{1}$$

where $\sigma_i \in \{-1, +1\}$ is a spin variable for spin i, $J_{ij} \in \mathbb{R}$ a coefficient related to the quadratic term between spins i and j, and $h_i \in \mathbb{R}$ a coefficient for the linear term with spin i. Let $q_i \in \{0, 1\}$ be a binary variable corresponding to the i-th spin, and then by applying the variable transformation $q_i = (\sigma_i + 1)/2$, we have

$$H = -\sum_{i,j} \widetilde{J}_{i,j}q_iq_j - \sum_i \widetilde{h}_iq_i, \tag{2}$$

where $\widetilde{J}_{i,j}$ and $\widetilde{h}_i$ should be transformed from $\{J_{ij}\}$ and $\{h_i\}$ adequately. As for the relations between the QUBO formulation and the Ising Hamiltonian, please see, for example, Ref. 6; in Ref. 6, some examples of the QUBO formulations for typical optimization problems are also given.

2.2 Legendre transformation

For reader's convenience, we here give a brief notation for the Legendre transformation.

If a function f_L is convex, the Legendre transformation of f_L, the so-called conjugate function of f_L, is given as follows:

$$f_L^*(t) = \sup_x\{tx - f_L(x)\}. \tag{3}$$

That is, the variable t is introduced, and the function for x is transformed to the function for t. In addition, (3) is equivalent to following equation:

$$f_L^*(t) = -\inf_x\{f_L(x) - tx\}. \tag{4}$$

2.3 Previous work 1: q-loss function

Here, a brief review of the previous work by Denchev et al.[10] is given. The following q-loss function was proposed in Ref. 10:

$$L_q(m) = \min\left[(1-q)^2, (\max[0, 1-m])^2\right], \tag{5}$$

where $q \in (\infty, 0]$ is a parameter and m is a continuous variable. In Ref. 10, there is a discussion for the application of the q-loss function in machine learning problems, and it was clarified that the q-loss function has a robust feature against label noise. Since (5) has a max function, it would not be easy to see the QUBO formulation of the q-loss function. Denchev et al. employed the Legendre transformation, and finally the following function was derived:[10]

$$L_q(m) = \min_t\left\{(m-t)^2 + (1-q)^2\frac{(1-\mathrm{sign}(t-1))}{2}\right\}, \tag{6}$$

where t is an additional variable which is introduced via the Legendre transformation. Although the variables m and t in (6) are continuous, the usage of the binary expansions gives the QUBO formulation for the q-loss function. As for details of the binary expansions, please see Ref. 10. Note that the sign function in (6) is also expressed as an one-body term when we employ the binary expansion.

2.4 Wolfe-duality

In nonlinear programming and mathematical optimization, the Wolfe duality theorem[12] is used to convert a main problem with inequality constraints to a dual problem. For a differentiable objective function and differentiable constraints, the main problem is written as follows:

$$\begin{cases} \text{minimize}_{\boldsymbol{x}} & f_{\mathrm{W}}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \mathbb{R}^n), \\ \text{subject to} & h_i(\boldsymbol{x}) \le 0 \quad (i = 1, 2, \ldots, l), \end{cases} \tag{7}$$

where $f_{\mathrm{W}}(\boldsymbol{x})$ is a certain convex function to be optimized and $\{h_i(\boldsymbol{x})\}$ are convex and inequality constraints. The Lagrangian function for this optimization problem is

$$L(\boldsymbol{x}, \boldsymbol{z}) = f_{\mathrm{W}}(\boldsymbol{x}) + \boldsymbol{z}^T h(\boldsymbol{x}), \tag{8}$$

where $\boldsymbol{z}$ is a vector of the Legendre coefficients. Then, the Wolfe dual theorem means that the minimization problem in

Eq. (7) is equivalent to the following maximization problem:

$$\begin{cases} \text{maximize}_{x,z} & L(\boldsymbol{x}, \boldsymbol{z}) \quad ((\boldsymbol{x}, \boldsymbol{z}) \in \mathbb{R}^n \times \mathbb{R}^l), \\ \text{subject to} & \nabla L(\boldsymbol{x}, \boldsymbol{z}) = 0 \quad (\boldsymbol{z} \geq 0). \end{cases} \tag{9}$$

As shown above, the Wolfe dual theorem transforms the minimization problem to the maximization problem.

2.5 *Previous work 2: ReLU-type function*

In Ref. 11, the QUBO form of the following ReLU-type function was discussed:

$$f_{\mathrm{ReLU}}(m) = -\min(0, m). \tag{10}$$

Note that the function $f_{\mathrm{ReLU}}(m)$ becomes the conventional ReLU function when the variable transformation $m \to -m$ is employed. As shown in Ref. 11, a naive application of the Legendre transformation to the function $f_{\mathrm{ReLU}}(m)$ in (10) gives the following expression:

$$f_{\mathrm{ReLU}}(m) = -\min_t \{-mt\} \quad \text{subject to} \quad -1 \leq t \leq 0, \tag{11}$$

where t is a new variable which stems from the Legendre transformation.

Although (11) has the QUBO formulation after adequate binary expansions, it is not suitable for optimization problems. This is because the minus sign before the min function; when the ReLU-type function is used as a kind of constraints or penalty terms for an optimization problem with a cost function $C(m)$, the whole minimization problem is, for example, given as follows:

$$\begin{aligned} \min_m \{C(m) + f_{\mathrm{ReLU}}(m)\} &= \min_m \left\{ C(m) - \min_t \{-mt\} \right\} \\ &\neq \min_{m,t} \{C(m) - (-mt)\}, \end{aligned} \tag{12}$$

and hence the cost function $C(m)$ and the ReLU-type function $f_{\mathrm{ReLU}}(m)$ cannot be minimized simultaneously. Therefore, the Wolfe duality theorem was employed in Ref. 11, and finally the following formulation was derived:

$$f_{\mathrm{ReLU}}(m) = \min_{t, z_1, z_2} \left\{ mt + z_1(t+1) - z_2 t - M(-m - z_1 + z_2)^2 \right\}, \tag{13}$$

where M is a large positive constant. It is easy to see that (13) can be used with the combination of the cost function $C(m)$.

Note that it is possible to make the ℓ_1 norm from a combination of two ReLU-type functions. However, this construction causes redundant variables. In the next section, we will directly derive the QUBO formulation starting from the ℓ_1 norm. In the derivation, two techniques explained in Sect. 2.2 and Sect. 2.4 play important roles.

3. Naive derivation of QUBO formulation for ℓ_1-norm

In this section, the Legendre transformation and the Wolfe dual theorem are applied to the ℓ_1 norm-type function naively.

3.1 *QUBO formulation*

Although the ℓ_1-norm is usually denoted as an absolute value of a variable, i.e., $|m|$, here we employ the following function $f(m)$:

$$f(m) = -\min\{-m, m\}. \tag{14}$$

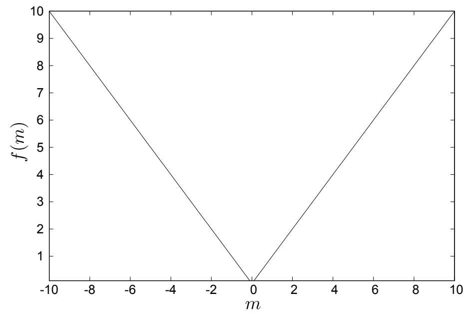

Fig. 1. The function form of ℓ_1-norm in (14).

Note that $f(m)$ can be expressed as follows:

$$\begin{aligned} f(m) &= -\min\{0, m\} - \min\{-m, 0\} \\ &= f_{\mathrm{ReLU}}(m) + f_{\mathrm{ReLU}}(-m), \end{aligned} \tag{15}$$

where $f_{\mathrm{ReLU}}(m)$ is the ReLU-type function in (10). Hence, it is easy to obtain the QUBO formulation for $f(m)$ by using the discussion based on Sect. 2.5. However, the QUBO formulation needs six additional variables. The derivation below enables us to obtain the QUBO formulation for $f(m)$ with only three additional variables.

Here, we try naive application of the Legendre transformation to the function $f(m)$ in (14). In order to employ the Legendre transformation, we give the following form of (14):

$$f(m) = \begin{cases} -m & (m < 0), \\ m & (m \geq 0). \end{cases}$$

Then, the Legendre transformation in (4) is performed for each domain as follows:

(a) $m < 0$:
The gradient in the domain is always -1. Hence, conjugate function is

$$f^*(t) = -\inf_m\{-m - mt\} = -\inf_m\{-m(1+t)\} = 0.$$

In addition, the possible value of t is only $t = -1$.

(b) $m = 0$:
Since the left derivative at this point is $f'_-(m) = -1$ and the right derivative is $f'_+(m) = 1$, the gradient value takes an arbitrary value within -1 to 1. Hence, the conjugate function is $f^*(t) = 0$ with the domain $t \in [-1, 1]$.

(c) $m > 0$:
The gradient in the domain is always 1. Hence, conjugate function is

$$f^*(t) = -\inf_m\{m - mt\} = -\inf_m\{-m(-1+t)\} = 0.$$

In addition, the possible value of t is only $t = 1$.

From the above discussion, the conjugate function of $f(m)$ is $f^*(t) = 0$ $(-1 \leq t \leq 1)$. When we apply the Legendre transformation to $f^*(m)$ again, $f(m)$ is adequately recovered since the function $f(m)$ is convex. Therefore, we find the quadratic form of $f(m)$ as follows:

$$F(m) = -\min_t \{-mt\} \quad \text{subject to} \quad 1 \leq t \leq 1. \tag{16}$$

In order to emphasize the fact that it is the quadratic form of $f(m)$, we newly introduced $F(m)$ instead of $f(m)$.

As shown in Sect. 2.5, although the obtained expression via the Legendre transformation has a quadratic form, it cannot be combined with another cost function. Hence, the Wolfe dual theorem is employed; the following expression is immediately obtained by applying the Wolfe dual theorem to $F(m)$:

$$\widetilde{F}(m) = \max_{t,z_1,z_2}\{-mt - z_1(t+1) + z_2(t-1)\} \tag{17}$$

$$\text{subject to} \quad \begin{cases} -m - z_1 + z_2 = 0, \\ -1 \leq t \leq 1,\ 0 \leq z_1,\ 0 \leq z_2. \end{cases}$$

This reformulation has the equality constraint, $-m - z_1 + z_2 = 0$; in order to embed this constraint into the QUBO formulation, it is enough to add the squared term as a penalty. Therefore, the optimization problem (17) can be represented as follows:

$$\begin{aligned} \widetilde{F}(m) = \min_{t,z_1,z_2} &\{mt + z_1(t+1) - z_2(t-1) \\ &+ M(-m - z_1 + z_2)^2\} \\ \text{subject to} \quad & -1 \leq t \leq 1,\ 0 \leq z_1, 0 \leq z_2. \end{aligned} \tag{18}$$

where M is a constant and takes a large value to ensure the equality constraint, $-m - z_1 + z_2 = 0$, to be satisfied. Note that there are remaining inequality constraints, $-1 \leq t \leq 1, 0 \leq z_1$, and $0 \leq z_2$; these inequality constraints can be easily realized by expanding these variables t, z_1, and z_2, in the binary expressions which satisfy the corresponding domain constraints respectively.

As a result, the QUBO formulation for the ℓ_1-norm is expressed by using additional three variables.

3.2 Numerical validation

Here, numerical checks are given in order to see the validity of the obtained QUBO formulation in (18). We expect that the QUBO formulation will be used within quantum annealing methods or simulated annealing methods, and here simulated annealing algorithms are employed. Since the purpose here is to verify the obtained QUBO formulation, the function in (18) is experimented with continuous variables without binary expansions.

The aim here is to check whether $\widetilde{F}(m)$ in (18) gives the ℓ_1-norm in (14) or not. We randomly generate m and check the value of $\widetilde{F}(m)$ by using a simulated annealing method for continuous variables. m is chosen from an uniform distribution in the range of $[-10, 10]$. A simulated annealing is performed for each chosen m. In each numerical experiment, the initial condition for the additional variables, t, z_1, and z_2 is chosen as follows:

- t is generated from an uniform distribution with the range of $[-1, 1]$.
- z_1 and z_2 are generated from an uniform distribution with the range of $[0, 10]$.

As the simulated annealing method for the continuous variables, we employ a conventional Metropolis-Hastings-type method. In order to generate next candidates of the state, each variable moves by the amount of $+0.001$ or -0.001 with the same probability for each iteration. At each iteration, the temperature is changed with the annealing schedule $T_{n+1} = 0.9999T_n$, where n is the iteration step. The initial temperature is set as $T_1 = 1000$, and the annealing is finished when the temperature is lower than 10^{-3}.

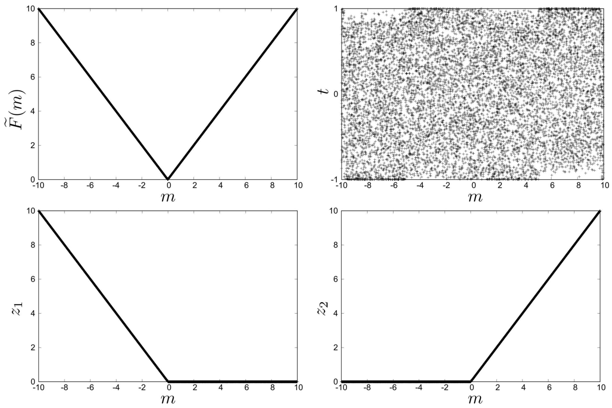

Fig. 2. Numerical results obtained from the annealing method. m is chosen randomly, and the annealing is performed; each m gives a value of $\widetilde{F}(m)$. The values of t, z_1, z_2 at the optimum states are also shown.

Figure 2 shows the results of the annealing. We confirm that the ℓ_1-norm is adequately recovered by the optimization of (18). Not only the value of the cost function $\widetilde{F}(m)$, but also the values of the three additional variables are also shown in Fig. 2; it is clear that z_1 and z_2 converge to specific values, but t takes various values randomly. This means that t would not be necessary for the optimization, which may give us a further simplified QUBO formulation for the ℓ_1-norm.

4. Reduced QUBO formulation

4.1 Reduction of the variable in the Legendre transformation

As discussed above, the naive application of the Legendre transformation and the Wolfe dual theorem gives the QUBO formulation with three additional variables. However, from the numerical experiments in the previous section, it is revealed that the variable t, which stems from the Legendre transformation, may not be necessary for the optimization problem. Because of the restriction of the number of spin variables, it is preferable to have smaller number of variables in general. Hence, here we try further reduction of variable from the QUBO formulation in (18).

In order to achieve the elimination of t from (18), we focus on the equality constraint $-m - z_1 + z_2 = 0$. By employing the equality $z_2 = m + z_1$, we have

$$\begin{aligned} \widetilde{F}(m) &= \min_{t,z_1,z_2}\{mt + z_1(t+1) - z_2(t-1) \\ &\qquad + M(-m - z_1 + z_2)^2\} \\ &= \min_{t,z_1,z_2}\{mt + z_1(t+1) - (m+z_1)(t-1) \\ &\qquad + M(-m - z_1 + z_2)^2\} \\ &= \min_{z_1,z_2}\{z_1 + (m+z_1) + M(-m - z_1 + z_2)^2\} \\ &= \min_{z_1,z_2}\{z_1 + z_2 + M(-m - z_1 + z_2)^2\}. \end{aligned} \tag{19}$$

Then, we finally obtain the following simplified QUBO formulation for the ℓ_1-norm:

$$\widehat{F}(m) = \min_{z_1,z_2}\{z_1 + z_2 + M(-m - z_1 + z_2)^2\}$$

$$\text{subject to} \quad 0 \le z_1, 0 \le z_2. \tag{20}$$

where the new expression $\widehat{F}(m)$ is introduced in order to clarify the difference from (18). Therefore, this conversion from (18) to (20) is possible because the penalty term, $M(-m-z_1+z_2)^2$, forces the equality constraint to be satisfied.

4.2 Numerical validation

In order to check the validity of the simplified QUBO formulation in (20), we again perform numerical experiments. The same settings and procedures as the previous section is employed for the annealing, except for the absence of the variable t. As a consequence, we obtained the same figures in Fig. 2, except for the figure about t (the top-right figure in Fig. 2); not using the redundant variable t, we completely recover the ℓ_1 function. This clarifies that the simplified QUBO formulation works well.

5. Concluding remarks

In this paper, the QUBO formulation for the ℓ_1-norm is derived. Using the Legendre transformation and the Wolfe theorem, a QUBO formulation was derived. Furthermore, it was clarified that a variable, which was introduced by the derivation, is redundant; finally, we numerically confirmed that the final simplified QUBO formulation works well. At first glance, it would be difficult to see the connection between the ℓ_1-norm and the final expression in (20); this nontrivial result is the main contribution of the present work.

There are some remarks for the derivation of the QUBO formulations. As discussed in Sect. 1, there is no systematic way to derive the QUBO formulation. While the Legendre transformation and the Wolfe theorem could be available for various cases, the derivation gives additional variables in general; the reduction of additional variables is important for the Ising-type hardware because of the current limitation of spin variables. As shown in the derivation, the variable t, which is introduced via the Legendre transformation, was reduced finally. It is still not clear whether the usage of the Legendre transformation is necessary or not; at this stage, the procedure (the Legendre transformation $\rightarrow$ the Wolfe dual theorem $\rightarrow$ reduction of variables) is straightforward and understandable. Of course, there could be more suitable derivation methods. In addition, when the ℓ_1-norm is combined to another cost function in order to make sparse estimation with the Ising-type hardware, it is necessary to add two variables for each estimated value, which means that the number of additional variables is still large. It will be important future works to find further reduction methods for practical problems with large size.

1) M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson and G. Rose, Nature **473**, 194 (2011).
2) P. I. Bunyk, E. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, J. Whittaker, IEEE Trans. Appl. Supercond. **24**, 1700110 (2014).
3) M. Aramon, G. Rosenberg, E. Valiante, T. Miyazawa, H. Tamura, and H. G. Katzgraber, Front. Phys. **7**, 48 (2019).
4) T. Kadowaki and H. Nishimori, Phys. Rev. E **58**, 5355 (1998).
5) E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).
6) K. Tanahashi, S. Takayanagi, T. Motohashi, and S. Tanaka, J. Phys. Soc. Jpn. **88**, 061010 (2019).
7) J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe and S. Lloyd, Nature **549**, 195 (2017).
8) A. Lucas, Front. Phys. **2**, 5 (2014).
9) J. D. Whitfield, M. Faccin, and J. D. Biamonte, Europhys. Lett. **99**, 57004 (2012).
10) V. Denchev, N.Ding, S. V. N Vishwanathan, and H. Neven, in Proceedings of the 29th International Conference on Machine Learning, p.863 (2012).
11) G. Sato, M. Konoshima, T. Ohwa, H. Tamura, and J. Ohkubo, Phys. Rev. E **99**, 042106 (2019).
12) P. Wolfe, Quart. Appl. Math. **19**, 239 (1961).
13) R. Tibshirani, J. R. Statist. Soc, B **58**, 267 (1996).
14) M. Honma, K. Akiyama, F. Tazaki, K. Kuramochi, S. Ikeda, K. Hada, and M. Uemura, J.Phys.: Conf. Ser. **699**, 012006 (2016).