

# Irreversible simulated tempering algorithm with skew detailed balance conditions: a learning method of weight factors in simulated tempering

Yuji Sakai<sup>1</sup> and Koji Hukushima<sup>1,2</sup>

<sup>1</sup> Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan

<sup>2</sup> Center for Materials Research by Information Integration, National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

E-mail: yuji0920@huku.c.u-tokyo.ac.jp

**Abstract.** Recent numerical studies concerning simulated tempering algorithm without the detailed balance condition are reviewed and an irreversible simulated tempering algorithm based on the skew detailed balance condition is described. A method to estimate weight factors in simulated tempering by sequentially implementing the irreversible simulated tempering algorithm is studied in comparison with the conventional simulated tempering algorithm satisfying the detailed balance condition. It is found that the total amount of Monte Carlo steps for estimating the weight factors is successfully reduced by applying the proposed method to an two-dimensional ferromagnetic Ising model.

## 1. Introduction

Since the invention of the Markov-chain Monte Carlo (MCMC) method [1], some improvement and development have been made for an efficient sampling, such as extended-ensemble methods [2, 3, 4, 5] and cluster algorithms [6, 7], and they have been developed within the framework of the detailed balance condition (DBC). Recently, the MCMC methods without the DBC have been discussed to improve their performance and several MCMC algorithms violating the DBC have been reported [8, 9, 10, 11, 12]. In addition, the efficiency of MCMC algorithms without the DBC is partly confirmed theoretically, in terms of the relaxation rate [13] and the asymptotic variance [14].

In the present paper, we make a brief review on a simulated tempering algorithm violating the DBC called the irreversible simulated tempering, particularly by imposing a skew detailed balance condition (SDBC), instead of the DBC. The simulated tempering algorithm has a specific weight factor to be fixed in advance, which determines the efficiency of the algorithm. We also discuss the way to estimate weight factors in the simulated tempering by the sequential update with the simulated tempering algorithm with the SDBC.

## 2. Irreversible simulated tempering algorithm

In this section, the outline of the irreversible simulated tempering (IST) algorithm based on the lifting technique with the skew detailed balance condition [15] is briefly explained.



Let  $\mathbf{X}$  be a configuration and  $E(\mathbf{X})$  denotes an energy function of a system to be considered. In statistical physics, the target distribution of the system is often given by the Gibbs-Boltzmann distribution with an inverse temperature  $\beta$ . In the simulated tempering, the inverse temperature is treated as a random variable as well as the configuration  $\mathbf{X}$ . More specifically,  $\beta$  takes  $R$  different values  $\{\beta_r\}_{r=1}^R$  that should be determined before simulation. In addition, a lifting variable  $\varepsilon \in \{+, -\}$  is introduced to the system in the irreversible simulated tempering. Thus, a state in the irreversible simulated tempering is specified by  $(\mathbf{X}, \beta_r, \varepsilon)$ . Accordingly, the target distribution is given as

$$P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon) \propto \exp[-\beta_r E(\mathbf{X}) + g_r], \quad (1)$$

where  $g_r$  denotes a weight factor depending only on the inverse temperature.

An explicit update scheme of the irreversible simulated tempering algorithm consists of two steps. one is the update scheme of an original configuration  $\mathbf{X}$  for a fixed inverse temperature and lifting variable with a conventional MCMC algorithm such as the Metropolis-Hastings algorithm and cluster algorithms. The other is the update scheme of the inverse temperature and the lifting variable for fixed  $\mathbf{X}$ , described as follows:

- (a) Let the current state be  $(\mathbf{X}, \beta_r, \varepsilon)$  and the candidate of the next inverse temperature  $\beta_l$  is determined with the probability  $q_{r,l}^{(\varepsilon)}$  given as follows:

$$q_{1,2}^{(\varepsilon)} = q_{R,R-1}^{(\varepsilon)} = 1, \quad (2)$$

$$q_{r,r\pm 1}^{(\varepsilon)} = \frac{1 \pm \delta\varepsilon}{2}, \quad (3)$$

for  $1 < r < R$ , and  $q_{r,l}^{(\varepsilon)} = 0$  otherwise.

- (b) Accept the next state  $(\mathbf{X}, \beta_l, \varepsilon)$  with the probability  $W_{r,l}^{(\varepsilon)}$  given by

$$W_{r,l}^{(\varepsilon)} = \min \left[ 1, \frac{q_{l,r}^{(-\varepsilon)} P_{\text{IST}}(\mathbf{X}, \beta_l, -\varepsilon)}{q_{r,l}^{(\varepsilon)} P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon)} \right]. \quad (4)$$

- (c) If the trial (b) is rejected, flip the lifting variable  $\varepsilon$  with the probability  $\Lambda_r^{(\varepsilon)}$  given by

$$\Lambda_r^{(\varepsilon)} = \max \left[ 0, \varepsilon \sum_{\varepsilon'=\pm} \sum_{l \neq r} \varepsilon' q_{r,l}^{(-\varepsilon')} W_{r,l}^{(-\varepsilon')} \right] / \left( 1 - \sum_{l \neq r} q_{r,l}^{(\varepsilon)} W_{r,l}^{(\varepsilon)} \right) \quad (5)$$

and set  $(\mathbf{X}, \beta_r, -\varepsilon)$  as the next state.

- (d) If the trial (c) is also rejected, set the current state as the next state.

Note that the acceptance probability satisfies the SDBC with respect to the target distribution  $P_{\text{IST}}(\mathbf{X}, \beta_r, \varepsilon)$  and the global balance condition is fulfilled in the above procedure. The parameter  $\delta$  in the proposal probability  $q_{r,l}^{(\varepsilon)}$  controls the violation of the DBC. When  $\delta$  is set to zero, the DBC is restored.

In Ref. [15], the irreversible simulated tempering algorithm has been applied to the two-dimensional ferromagnetic Ising model as a benchmark. It is numerically shown that the relaxation dynamics of the inverse temperature qualitatively changes from diffusive to ballistic by violating the DBC and consequently the autocorrelation time of the magnetization is reduced several times compared to the conventional simulated tempering for the case with an ideal choice of the weight factors. Thus, it is confirmed that the violation of the DBC can improve the efficiency of simulated tempering algorithm.

### 3. On-the-fly estimation of weight factors

The choice of the weight factors essentially affects the efficiency of the simulated tempering. When  $g_r = -\ln Z(\beta_r)$  with  $Z(\beta)$  being the partition function at  $\beta$ , the marginal probability of inverse temperature of the target distribution is uniform and all the inverse temperature is expected to be sampled uniformly. However, it is difficult to know the exact value of the partition function a priori in general. Then, in practice the weight factors are prepared approximately by some algorithms [16, 17, 18, 19, 20]. In this section, we study an on-the-fly estimation method of the optimal weight factors with the irreversible simulated tempering algorithm. The basic idea of the on-the-fly scheme has already proposed in Ref. [19].

Let  $(\mathbf{X}^{(n)}, \beta^{(n)}, \varepsilon^{(n)})$  be the state of the irreversible simulated tempering after  $n$  MCS and  $N_r^{(n)}$  denotes the histogram of the inverse temperature  $\beta_r$  defined as

$$N_r^{(n)} \equiv \sum_{k=1}^n \delta(\beta^{(k)}, \beta_r), \quad (6)$$

where  $\delta(\cdot, \cdot)$  is the Kronecker delta. Then, the estimator of expectation of energy at  $\beta = \beta_r$  is given as

$$\bar{E}_r^{(n)} \equiv \frac{1}{N_r^{(n)}} \sum_{k=1}^n E(\mathbf{X}^{(k)}) \delta(\beta^{(k)}, \beta_r), \quad (7)$$

and the weight factors are updated iteratively according to the following equation,

$$g_{r+1} = g_r + \frac{\bar{E}_r^{(n)} + \bar{E}_{r+1}^{(n)}}{2} (\beta_{r+1} - \beta_r). \quad (8)$$

The on-the-fly weight determination scheme is given as follows.

- (a) Prepare a set of inverse temperatures  $\{\beta_r\}_{r=1}^R$  such that  $\beta_1 > \beta_2 > \dots > \beta_R$  without loss of generality. Set  $g_r^{(0)} = 0$  for  $r = 1, 2, \dots, R$ .
- (b) Start the IST simulation from the largest inverse temperature  $\beta_1$  and a random initial condition  $\mathbf{X}^{(0)}$  and  $\varepsilon^{(0)}$ . Calculate  $\bar{E}_1^{(n)}$  and update  $g_2$  as  $g_2 = (\beta_2 - \beta_1) \bar{E}_1^{(n)} / 2$  after every MCS.
- (c) If the transition from  $\beta_1$  to  $\beta_2$  is realized for the first time, start to calculate  $\bar{E}_2^{(n)}$  and update  $g_2$  and  $g_3$  according to Eq. (8).
- (d) Continue the IST simulation, calculate  $\bar{E}_1^{(n)}$  and  $\bar{E}_2^{(n)}$ , and update  $g_2$  and  $g_3$  after every MCS. If the transition to  $\beta_3$  is realized for the first time, start to calculate  $\bar{E}_3^{(n)}$  and update  $g_3$  and  $g_4$  according to Eq. (8).
- (e) Continue the above procedure until  $g_R$  is updated. Note that during the simulation, we update weight factors  $g_r$  and  $g_{r+1}$  after each MCS when the current inverse temperature is  $\beta_r$ .

We apply the above algorithm to the two-dimensional ferromagnetic Ising model without external magnetic field. In this work, The number of spins is fixed as  $N = 32^2$  and the set of inverse temperature is prepared as  $\beta_1 = 0.5$ ,  $\beta_R = 0.2$ , and the others are aligned with equal interval. A Monte Carlo step (MCS) is defined as the time unit where  $N$  spin-flip trials with Metropolis-Hastings algorithm and a update trial of  $(\beta, \varepsilon)$  are performed.

First, we study the total number of MCSs required to accomplish the estimation. Figure 1 is showing  $R$ -dependence of the total MCSs. While no difference between  $\delta = 0$  and  $\delta \neq 0$  is

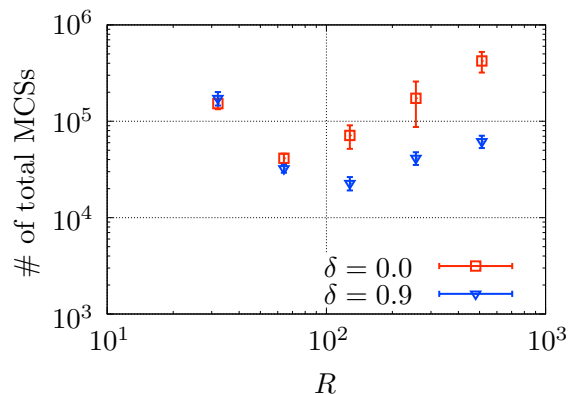


**Figure 1.**  $R$ -dependence of total number of MCSs for the on-the-fly estimation of weight factors. The parameter  $\delta$  is chosen as  $\delta = 0.0$  (red square) and  $\delta = 0.9$  (blue inverted triangle), respectively.

observed in small  $R$  region in Fig. 1, the total MCSs in the case of  $\delta \neq 0$  is about 10 times less than that in the case of  $\delta = 0$  when the number of inverse temperature is sufficiently large. It indicates that the computational cost of the estimation of weight factors can be reduced by violating the DBC.

We next focus on the estimation accuracy of weight factors. Figure 2 shows that the weight factors estimated by our algorithm with the parameter  $\delta = 0$  and  $\delta = 0.9$  are comparable. Thus, we conclude that the violation of the DBC can improve the estimation algorithm proposed in Ref. [19].

#### 4. Summary

In this paper, the irreversible simulated tempering algorithm violating the DBC was briefly reviewed. The update scheme in the simulated tempering algorithm is modified by introducing the lifting technique and the skew detailed balance condition is imposed to the acceptance probability of inverse temperature. It turned out [15] that the violation of the DBC accelerates the relaxation of inverse temperature and the autocorrelation of magnetization in comparison with the conventional simulated tempering. Here we propose an iterative algorithm to estimate the optimal weight factor by implementing the irreversible simulated tempering. By applying the method to the Ising model in two dimensions, it is revealed that the total Monte Carlo steps required in the estimation is successfully reduced without deteriorating the precision of the estimation by violating the DBC. Our results indicate that the violation of the DBC by imposing the SDBC can also improve the efficiency of the estimation of weight factors in simulated tempering. As a future work, our algorithm would be applied to more complicated systems such as the  $q$ -state Potts model with large value of  $q$  and spin glasses.

#### acknowledgments

The authors are grateful to S. Todo for useful comments and for letting us notice the method of Ref. [21]. YS is supported by a Grant-in-Aid from the Japan Society for Promotion of Science (JSPS) Fellows (Grant No. 26-7868). KH is supported by Grants-in-Aid for Scientific Research from MEXT, Japan (Grant Nos. 25610102 and 25120010).

#### References

- [1] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 *J. Chem. Phys.* **21** 1087–92
- [2] Berg B A and Neuhaus T 1992 *Phys. Rev. Lett.* **68** 9
- [3] Hukushima K and Nemoto K 1996 *J. Phys. Soc. Jpn.* **65** 1604
- [4] Lyubartsev A P, Martinovski A A, Shevkunov S V and Vorontsov-Velyaminov P N 1992 *J. Chem. Phys.* **96** 1776
- [5] Marinari E and Parisi G 1992 *Europhys. Lett.* **19** 451

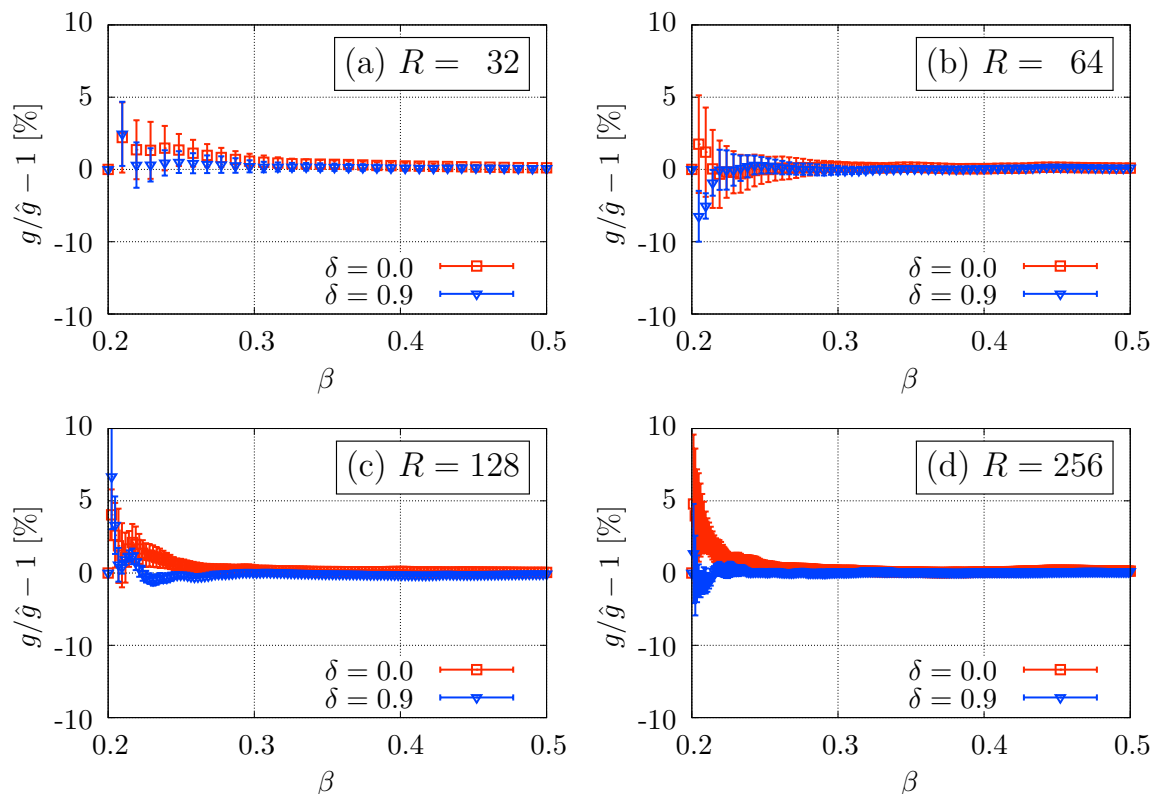


**Figure 2.** (Color online) Estimation accuracy of weight factors as a function of inverse temperature  $\beta$ . The number of inverse temperature is (a) $R = 32$ , (b) $R = 64$ , (c) $R = 128$ , and (d) $R = 256$ . The parameter  $\delta$  is chosen as  $\delta = 0.0$  (red square) and  $\delta = 0.9$  (blue inverted triangle). Horizontal axis in each panel represents the ratio between the weight factor  $g$  obtained by on-the-fly estimation and the weight factor  $\hat{g}$  evaluated by the exact numerical method in Ref. [21].

- [6] Swendsen R H and Wang J S 1987 *Phys. Rev. Lett.* **58** 86
- [7] Wolff U 1989 *Phys. Rev. Lett.* **62** 361
- [8] Suwa H and Todo S 2010 *Phys. Rev. Lett.* **105** 120603
- [9] Turitsyn K S, Chertkov M and Vucelja M 2011 *Physica D* **240** 410
- [10] Fernandes H C and Weigel M 2011 *Comput. Phys. Commun.* **182** 1856
- [11] Schram R D and Barkema G T 2014 *Physica A* **418** 88
- [12] Bierkens J 2015 *Stat. Comput.*
- [13] Ichiki A and Ohzeki M 2013 *Phys. Rev. E* **88** 020101(R)
- [14] Sun Y, Gomez F and Schmidhuber J 2010 NIPS2010 proceedings 2235
- [15] Sakai Y and Hukushima K 2016 *Preprint* cond-mat.stat-mech/1601.04286
- [16] Hansmann U H E and Okamoto Y 1997 *J. Comput. Chem.* **18** 920
- [17] Mitsutake A and Okamoto Y 2004 *J. Chem. Phys.* **121** 2491
- [18] Park S and Pande V S 2007 *Phys. Rev. E* **76** 016703
- [19] Nguyen P H, Okamoto Y and Derreumaux P 2013 *J. Chem. Phys.* **138** 061102
- [20] Valentim A, da uz M G E and Fiore C E 2014 *Comput. Phys. Commun.* **185** 2046
- [21] Kastening B 2001 *Phys. Rev. E* **64** 066106