Statistical Mechanical Analysis of Phase Unwrapping — One-Dimensional Model —

Tatsuya UEZU[1*], Mayu HACHISUKA[2], and Yohei SAIKA[3]

[1] *Nara Women's University, Nara 630-8506, Japan*

[2] *Graduate School of Humanities and Sciences, Nara Women's University, Nara 630-8506, Japan*

[3] *Department of Information and Computer Engineering, National Institute of Technology, Gunma College, Toriba 580, Maebashi 371-8530, Japan*

(Received January 21, 2023)

We study a phase unwrapping model in the one-dimensional space on the basis of Bayesian inference using the maximizer of posterior marginals (MPM) estimate by the statistical mechanical methods. We propose a model in which the recursion relations to obtain statistical quantities such as MPM estimates are derived. We introduce the three state Potts model to handle the discontinuities in observed data, and propose two methods, the step and direct methods. We derive the recursion relations for MPM estimates of hyperparameters and phase differences in both methods, and investigate the random and regular phase differences, and previously studied other type of random phase differences. We find that the phase differences are inferred fairly well in rather wide ranges of noise amplitudes. The ranges depend on samples and the system sizes. Furthermore, we find that the step method has performance in phase unwrapping comparable to the direct method, and that it is much faster in numerical computation and applicable to much larger system sizes than the direct method.

*tuezu0204@gmail.com

1. Introduction

Phase unwrapping[1] is a technique to infer true phases from corrupted and reduced phases. We give one example of one-dimensional phase unwrapping. That is the vestibulo-ocular reflex (VOR) of vertebrate animals, which is the eye movement system that stabilizes the field of view. When the head rotates to one direction, the eyes rotate to the opposite direction to retain the field of view. From experimentally observed rotation velocity of the head $f_i(t)$ and the eye movement velocity $f_o(t)$, their Fourier transformations $\hat{f}_i(\omega)$ and $\hat{f}_o(\omega)$ are calculated. From $\frac{\hat{f}_o(\omega)}{\hat{f}_i(\omega)} = a(\omega) + ib(\omega)$, the gain $\sqrt{a(\omega)^2 + b(\omega)^2}$ and the phase $\phi(\omega) = \tan^{-1}(\frac{b(\omega)}{a(\omega)})$ are obtained. $\phi(\omega)$ represents the phase shift between both velocities and has important information on eye movement response to head movement at each frequency. Since $\phi(\omega)$ is reduced to the value in the interval $[-\pi, \pi)$, it is necessary to unwrap the phase. It is well known that if there exists external noise, the phase unwrapping becomes difficult.

Many techniques of phase unwrapping have been proposed from various viewpoints.[2–11] However, in general two-dimensional models, theoretical treatments are so difficult that we may not able to confirm simulation results theoretically. It is quite desirable to introduce a model in which statistical quantities such as the maximizer of posterior marginals (MPM) estimates are analytically derived.

One-dimensional phase unwrapping models have been investigated from various viewpoints, such as the non-Gaussian filtering method,[12] the maximum a posterior probability (MAP) or MPM estimations based on the Bayesian inference,[8,13] and so on. In this paper, we propose a model in which the phase differences are generated from the correlated Gaussian distribution and external noises are generated from the Gaussian distribution. In our formulation, phase differences are corrupted by external noise and reduced to $[-\pi, \pi)$ during the observation process. We assume that the discontinuities in the reduced data, which we call jumps, are not very big, and introduce the three state Potts model.[11,14] We propose two methods, the step and direct methods. In the step method, the corrupted data are inferred from the reduced data in step I, and the phase differences are inferred from the corrupted data in step II. Hyperparameters are introduced and their priors and likelihood functions are assumed by imposing continuity of phases and/or phase differences. On the other hand, in the direct method, we infer the phase differences from the reduced data directly. We derive the recursion relations for the variables by which the marginal distributions and the MPM estimates are calculated by using the so-called transfer matrix (TM) method[14] which is also used in Information science[15,16] etc. and has different names such as the belief propagation (BP).

We perform numerical calculations in not only the random phase difference but also in the regular phase differences for small system sizes. Next, we study the large system sizes by the step method because the direct method is not available. Finally, we study the case that the phases are generated from correlated Gaussian distributions and suffer from the external noises generated from the von-Mises distribution studied in Ref. 13. By regarding the phases in Ref. 13 as the phase differences, we apply the step method to these phase differences and study the regions where the jump inference succeeds by comparing with results of the previous study.

The construction of the paper is as follows. In §2, we formulate the direct and step methods. In §3, we show numerical results for the random and regular phase differences. §4 contains a summary and discussion of the results. In the Appendices, we derive some mathematical relations, and describe the outline of the derivation of the marginal distributions and the MPM estimates for the step and direct methods.

2. Formulation

Let θ_i be the coordinate at which i-th phase ξ_i is observed ($i = 1, \cdots, L$). The true phase difference is $m_i = \xi_{i+1} - \xi_i$ ($i = 1, \cdots, L-1$). We define $\boldsymbol{m} = \{m_1, m_2, \cdots, m_{L-1}\}$. We assume that the white Gaussian noise z_i with mean 0 and variance σ_0^2 is added to the phase difference m_i during observation. Let y_i be a corrupted phase difference

$$y_i = m_i + z_i, \ i = 1, \cdots, L-1. \tag{1}$$

Furthermore, we assume that the corrupted phase differences are observed at the middle point $\hat{\theta}_i$ between θ_i and θ_{i+1} and are reduced to values in $[-\pi, \pi)$ ($i = 1, \cdots, L-1$). Let τ_i be the reduced phase difference which is defined by

$$\tau_i = y_i \mod 2\pi, \tau_i \in [-\pi, \pi), \ i = 1, \cdots, L-1. \tag{2}$$

Then,

$$y_i = \tau_i + 2\pi n_i. \tag{3}$$

We call n_i a jump. Thus, the observable is $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_{L-1})$. In order to apply the Bayesian inference, we assume that phase differences obey some probability distribution. In this paper, we assume that m_i is generated from the following correlated Gaussian distribution,

$$P(\boldsymbol{m}) = C_0 \exp(-\frac{1}{2\eta_0^2}\sum_{i=1}^{L}(m_i - m_{i-1})^2), \tag{4}$$

$$C_0 = \prod_{l=1}^{L-1}\sqrt{\frac{\lambda_l}{2\pi}}\frac{1}{\eta_0}, \tag{5}$$

$$\lambda_l = 2(1-\cos(\frac{l\pi}{L})), \quad l = 1, 2, \ldots, L-1, \tag{6}$$

where $m_0 = m_L = 0$ and $\eta_0 > 0$. See Appendix A. By the direct method, the phase differences $\boldsymbol{m}$ are inferred from $\boldsymbol{\tau}$ directly, and by the step method, the jumps $\boldsymbol{n} = (n_1, n_2, \cdots, n_{L-1})$ are inferred from $\boldsymbol{\tau}$ firstly, and then $\boldsymbol{m}$ are inferred from $\boldsymbol{y} = \boldsymbol{\tau} + 2\pi\boldsymbol{n}$. We introduce the random variables $\boldsymbol{s} = (s_1, s_2, \cdots, s_{L-1})$ and $\boldsymbol{x} = (x_1, x_2, \cdots, x_{L-1})$ which are used to estimate $\boldsymbol{n}$ and $\boldsymbol{m}$, respectively. $P(\boldsymbol{x})$ is given by

$$P(\boldsymbol{x}) = C \exp(-\frac{1}{2\eta^2}\sum_{i=1}^{L}(x_i - x_{i-1})^2), \tag{7}$$

$$C = \prod_{l=1}^{L-1}\sqrt{\frac{\lambda_l}{2\pi}}\frac{1}{\eta}, \tag{8}$$

where $x_0 = x_L = 0$ and $\eta > 0$. Furthermore, we assume that the number of jumps are not very many, and s_i takes three values of $0, \pm 1$, that is, we study the three state Potts model.

2.1 Formulation for direct method

Firstly, we explain the direct method. From the prior $P(\boldsymbol{x})$ and the conditional probability $P(\tau_l|x_l)$, we obtain the posterior distribution $P(x_l|\boldsymbol{\tau})$. See Appendix B.

$$P(\tau_l|x_l) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{n_l=-\infty}^{\infty}\exp[-\frac{1}{2\sigma^2}(\tau_l + 2\pi n_l - x_l)^2], \tag{9}$$

$$P(\boldsymbol{\tau}|\boldsymbol{x}) = \prod_{l=1}^{L-1} P(\tau_l|x_l), \tag{10}$$

$$\begin{aligned} P(x_l|\boldsymbol{\tau}) &= \frac{1}{Z_{\mathrm{d}}}\int d\boldsymbol{x}^{(l)} P(\boldsymbol{\tau}|\boldsymbol{x})P(\boldsymbol{x}) \\ &= \frac{1}{Z_{\mathrm{d}}} C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\left[\prod_{j<l}\sqrt{\frac{2\pi}{a_j^{\mathrm{u}}}}\right]\left[\prod_{j>l}\sqrt{\frac{2\pi}{a_j^{\mathrm{d}}}}\right]\sum_{\boldsymbol{n}}\exp[-\frac{1}{2}a_l^{(\boldsymbol{n})}x_l^2 + b_l^{(\boldsymbol{n})}x_l + c_l^{(\boldsymbol{n})}], \end{aligned} \tag{11}$$

$$Z_{\mathrm{d}} = \int d\boldsymbol{x} P(\boldsymbol{\tau}|\boldsymbol{x})P(\boldsymbol{x}), \tag{12}$$

$$d\boldsymbol{x} \equiv dx_1 dx_2 \ldots dx_{L-1},$$

$$d\boldsymbol{x}^{(l)} \equiv dx_1 \ldots\ldots dx_{L-1} \quad \text{(excluding } dx_l), \tag{13}$$

$$\sum_{\boldsymbol{n}} \equiv \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \cdots \sum_{n_{L-1}=-\infty}^{\infty} . \tag{14}$$

The derivation of these relations and definitions of $a_j^{\mathrm{d}}, a_j^{\mathrm{u}}$ and their recursion relations to obtain $a_l^{(\boldsymbol{n})}, b_l^{(\boldsymbol{n})}, c_l^{(\boldsymbol{n})}$ are shown in Appendix C. In the MPM estimates, since the expression of $P(x_l|\boldsymbol{\tau})$, eq. (11), contains infinitely many terms in the summation, we restrict the value of n_i to three values $0, \pm 1$ assuming that jumps are not very big. We replace $\sum_{\boldsymbol{n}}$ by the following $\sum'_{\boldsymbol{n}}$,

$$\sum_{\boldsymbol{n}}{}' = \sum_{n_1=0,\pm1} \sum_{n_2=0,\pm1} \cdots \sum_{n_{L-1}=0,\pm1} . \tag{15}$$

We denote the MPM estimates for m_l and n_l as $\hat{m}_l$ and $\hat{n}_l$, respectively. $\hat{m}_l$ is x_l which maximizes $\sum'_{\boldsymbol{n}} \exp[-\frac{1}{2}a_l^{(\boldsymbol{n})}x_l^2 + b_l^{(\boldsymbol{n})}x_l + c_l^{(\boldsymbol{n})}]$. $\hat{n}_l$ is determined as the value s_l which minimizes $|\hat{m}_l - (\tau_l + 2\pi s_l)|$.

<u>Hyperparameter inference</u>

The posterior probability for η and σ is given by

$$P(\eta, \sigma|\boldsymbol{\tau}) \quad \propto \quad P(\boldsymbol{\tau}|\eta, \sigma)P(\eta, \sigma). \tag{16}$$

We assume that we do not have any information on the prior $P(\eta, \sigma)$, and put $P(\eta, \sigma)$=constant. We obtain

$$P(\eta, \sigma|\boldsymbol{\tau}) \quad \propto \quad P(\boldsymbol{\tau}|\eta, \sigma) = \int d\boldsymbol{x} P(\boldsymbol{\tau}|\boldsymbol{x}, \sigma)P(\boldsymbol{x}|\eta) = Z_{\mathrm{d}}. \tag{17}$$

Z_{d} is calculated as

$$Z_{\mathrm{d}} \quad = \quad C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1} \left[\prod_{j<l}\sqrt{\frac{2\pi}{a_j^{\mathrm{u}}}}\right] \left[\prod_{j>l}\sqrt{\frac{2\pi}{a_j^{\mathrm{d}}}}\right] \sum_{\boldsymbol{n}}{}' \sqrt{\frac{2\pi}{a_l^{(\boldsymbol{n})}}} \exp[\frac{\left(b_l^{(\boldsymbol{n})}\right)^2}{2a_l^{(\boldsymbol{n})}} + c_l^{(\boldsymbol{n})}], \tag{18}$$

where l is any integer value from 1 to $L-1$. Note that $P(\boldsymbol{\tau}|\boldsymbol{x}, \sigma)$ is normalized since $[\prod_{i=1}^{L-1} \int_{-\pi}^{\pi} d\tau_i] P(\boldsymbol{\tau}|\boldsymbol{x}, \sigma) = 1$. Thus, we only have to maximize Z_{d}. Instead of Z_{d}, we maximize $\ln Z_{\mathrm{d}}$ with respect to σ and η.

$$\ln Z_{\mathrm{d}} \quad = \quad -(L-1)\ln(\eta\sigma) - \frac{1}{2}\sum_{j<l} \ln a_j^{\mathrm{u}} - \frac{1}{2}\sum_{j>l} \ln a_j^{\mathrm{d}}$$

$$+\ln\Big(\sum_{\boldsymbol{n}}{}' \frac{1}{\sqrt{a_i^{(\boldsymbol{n})}}}\exp[\frac{\Big(b_l^{(\boldsymbol{n})}\Big)^2}{2a_l^{(\boldsymbol{n})}} + c_l^{(\boldsymbol{n})}]\Big) + \text{constant}. \quad (19)$$

2.2 Formulation for step method

2.2.1 Step I

We infer the jumps $\boldsymbol{n}$. We set the prior of s_i as

$$P(\boldsymbol{s}) \quad \propto \quad e^{-h\sum_{i=1}^{L-1}|s_i|^p}. \quad (20)$$

For any positive integer p, $|s_i|^p = |s_i|$ since $s_i = 0, \pm 1$. We put $p = 2$ in this paper. The jump is more difficult to take place for $h > 0$, and is easier to do for $h < 0$. The likelihood function $P(\boldsymbol{\tau}|\boldsymbol{s})$ is assumed to be

$$P(\boldsymbol{\tau}|\boldsymbol{s}) \quad \propto \quad e^{-H(\boldsymbol{\tau},\boldsymbol{s})}, \quad (21)$$

$$H(\boldsymbol{\tau},\boldsymbol{s}) \quad \equiv \quad \frac{1}{2\tilde{\eta}^2}\sum_{i=1}^{L}\Big(\tau_i + 2\pi s_i - (\tau_{i-1} + 2\pi s_{i-1})\Big)^2, \quad (22)$$

where $\tau_0 = \tau_L = 0,\ s_0 = s_L = 0$. That is, we assume that phase differences change continuously. By the Bayesian formula, we obtain

$$P(\boldsymbol{s}|\boldsymbol{\tau}) \quad \propto \quad P(\boldsymbol{\tau}|\boldsymbol{s})P(\boldsymbol{s}) \propto e^{-\mathcal{H}_{\mathrm{I}}(\boldsymbol{s},\boldsymbol{\tau})}, \quad (23)$$

$$\begin{aligned}\mathcal{H}_{\mathrm{I}}(\boldsymbol{s},\boldsymbol{\tau}) \quad &\equiv \quad H(\boldsymbol{\tau},\boldsymbol{s}) + h\sum_{i=1}^{L-1}s_i^2 \\ &= \quad \frac{1}{2\tilde{\eta}^2}\sum_{i=1}^{L}\Big(\tau_i + 2\pi s_i - (\tau_{i-1} + 2\pi s_{i-1})\Big)^2 + h\sum_{i=1}^{L-1}s_i^2. \end{aligned} \quad (24)$$

In order to estimate n_i and y_i, we adopt the MPM inference. Let $P_i(s_i)$ be the marginal distribution of s_i. Since s_i takes values 0, and ± 1, we obtain

$$P_i(0) = 1 - \langle s_i^2\rangle,\ P_i(1) = (\langle s_i^2\rangle + \langle s_i\rangle)/2,\ P_i(-1) = (\langle s_i^2\rangle - \langle s_i\rangle)/2. \quad (25)$$

Thus, the MPM estimate $\hat{n}_i$ for n_i is the argument which gives the maximum among $P_i(0), P_i(1)$ and $P_i(-1)$. We derive analytic formulae for $\langle s_i\rangle$ and $\langle s_i^2\rangle$ by using the TM method. See Appendix D.

Hyperparameter inference

Let us denote the summation of $s_l = 0, \pm 1$ as Tr_l. We obtain

$$P(h,\tilde{\eta}|\boldsymbol{\tau}) \quad \propto \quad P(\boldsymbol{\tau}|h,\tilde{\eta}) = [\prod_{l=1}^{L-1}\mathrm{Tr}_l]P(\boldsymbol{\tau},\boldsymbol{s}|h,\tilde{\eta}) \propto [\prod_{l=1}^{L-1}\mathrm{Tr}_l]e^{-\mathcal{H}_{\mathrm{I}}} = Z_{\mathrm{I}}. \quad (26)$$

Since $P(\boldsymbol{\tau}|h,\tilde{\eta})$ should be normalized, it is necessary to calculate $[\prod_{l=1}^{L-1}\int_{-\pi}^{\pi}d\tau_l]Z_{\rm I}$.

$$[\prod_{l=1}^{L-1}\int_{-\pi}^{\pi}d\tau_l]Z_{\rm I}$$
$$=[\prod_{l=1}^{L-1}\int_{-\pi}^{\pi}d\tau_l][\prod_{i=1}^{L-1}{\rm Tr}_i]\exp[-h\sum_{i=1}^{L-1}|s_i|^p-\frac{1}{2\tilde{\eta}^2}\sum_{i=1}^{L}\Big(\tau_i+2\pi s_i-(\tau_{i-1}+2\pi s_{i-1})\Big)^2].$$

This is easily calculated numerically by the TM method. We maximize $\ln Z$ with respect to h and $\tilde{\eta}$ to obtain the hyperparameters.

2.2.2 Step II

We infer the estimate $\hat{y}_i$ for the corrupted phase difference y_i by using the MPM estimate of the jump $\hat{n}_i$, that is, $\hat{y}_i=\tau_i+2\pi\hat{n}_i$. Below, we denote $\hat{\boldsymbol{y}}$ by $\boldsymbol{y}$ for simplicity. We proceed to estimate $\boldsymbol{m}$ from $\boldsymbol{y}$. The conditional probability $P(\boldsymbol{y}|\boldsymbol{x})$ is assumed to be the uncorrelated Gaussian distribution the same as $P(\boldsymbol{y}|\boldsymbol{m})$ with the standard deviation σ instead of σ_0.

$$P(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{L-1}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y_i-x_i)^2}{2\sigma^2}}. \tag{27}$$

From the Bayesian formula, we have

$$P(\boldsymbol{x}|\boldsymbol{y}) \propto P(\boldsymbol{y}|\boldsymbol{x})P(\boldsymbol{x})=C_{\rm II}\frac{1}{(\sqrt{2\pi}\sigma)^{L-1}}e^{-\mathcal{H}_{\rm II}}, \tag{28}$$

$$Z_{\rm II} = \int d\boldsymbol{x}e^{-\mathcal{H}_{\rm II}}, \tag{29}$$

$$\mathcal{H}_{\rm II} = \frac{1}{2\sigma^2}\sum_{i=1}^{L-1}(x_i-y_i)^2+\frac{1}{2\eta^2}\sum_{i=1}^{L}(x_i-x_{i-1})^2. \tag{30}$$

Let us consider the method to estimate m_i from $\boldsymbol{y}$ by the MPM inference. Since the marginal distribution $P_i(x_i)$ for x_i is the Gaussian distribution, the maximum of $P_i(x_i)$ is attained at the average value $\langle x_i\rangle$. Thus, the MPM estimate $\hat{m}_i$ is $\langle x_i\rangle$. We use the TM method to derive the formula for $\langle x_i\rangle$. See Appendix E.

Hyperparameter inference

$$P(\eta,\sigma|\boldsymbol{y}) \propto P(\boldsymbol{y}|\eta,\sigma)=C_{\rm II}\frac{1}{(\sqrt{2\pi}\sigma)^{L-1}}\int d\boldsymbol{x}e^{-\mathcal{H}_{\rm II}}=C_{\rm II}\frac{1}{(\sqrt{2\pi}\sigma)^{L-1}}Z_{\rm II}. \tag{31}$$

Since $P(\boldsymbol{y}|\eta,\sigma)$ is normalized, we maximize $I \equiv C_{\rm II}\frac{1}{(\sqrt{2\pi}\sigma)^{L-1}}Z_{\rm II}$ with respect to η and σ. Instead, we maximize $\ln I$.

$$\ln I \quad = \quad -(L-1)\ln(\eta\sigma) + \ln Z_{\rm II} + \text{constant}. \tag{32}$$

3. Numerical results

3.1 Random phase differences

We generate phase differences and corrupted phase differences for $\eta_0, \sigma_0 = 0.1, 0.2, \cdots, 2.0$. For the hyperparameter inference, we scan $h = -10.0, -9.9, \cdots, 10.0$ with $\Delta h = 0.1$, or $h = -0.1, -0.099, \cdots, 0.1$ with $\Delta h = 0.001$, and $\tilde{\eta} = 0.1, 0.2, \cdots, 2.0$ for step I, and $\eta, \sigma = 0.1, 0.2, \cdots, 2.0$ for step II and the direct method. We decide the optimal parameters which maximize the likelihood functions. We also use the gradient method, and find that both methods give almost the same optimal parameters and the MPM estimates. We generate two types A and B of phase differences and external noises. For type A, we generate $\{m_l^1\}$ from the correlated Gaussian distribution (eq. (4)) with $\eta_0 = 1$ and $\{z_l^1\}$ with the Gaussian distribution with mean 0 and standard deviation 1. Then, we set $m_l = \eta_0 m_l^1$ and $z_l = \sigma_0 z_l^1$. We generate only one sample. For type B, we generate $\{m_l\}$ for each η_0 from the correlated Gaussian distribution (eq. (4)) with η_0. For $\{z_l\}$, we generate $\{z_l^1\}$ with the Gaussian distribution with mean 0 and standard deviation 1 for each sample, and set $z_l = \sigma_0 z_l^1$. We generate 10 samples. That is, the latter case is more random than the former case.

Jump inference

Firstly, we show the numerical results for the jump inference. Let n_l be the true value of the jump at the location $\hat{\theta}_l$, and $\sigma_{0,J}$ be the maximum value of σ_0 below which the jump inference succeeds. That is, for $\sigma_0 \le \sigma_{0,J}$, the jump inference succeeds. In Fig. 1, we show the performance of jump inferences for the step and direct methods by heat map in the (η_0, σ_0) plane for types A and B and $L = 18$. Furthermore, we investigate the ratio of the number of locations l where the jump inferences succeed to the total number of locations $L-1$.

$$r \quad = \quad \frac{l}{L-1}. \tag{33}$$

In Fig. 2, we show the ratios for the step and direct methods, $r^{\rm s}$ and $r^{\rm d}$, and $R = r^{\rm s}/r^{\rm d}$. As is seen from Figs. 1 and 2, both methods give comparable jump inferences although the jump inference depends on samples. In Fig. 3, we show the η_0 dependences of $\langle\sigma_{0,J}\rangle$ averaged over 10 samples for type B. We find that for both methods $\langle\sigma_{0,J}\rangle \gtrsim 1$ for

$\eta_0 \lesssim 1.2$.

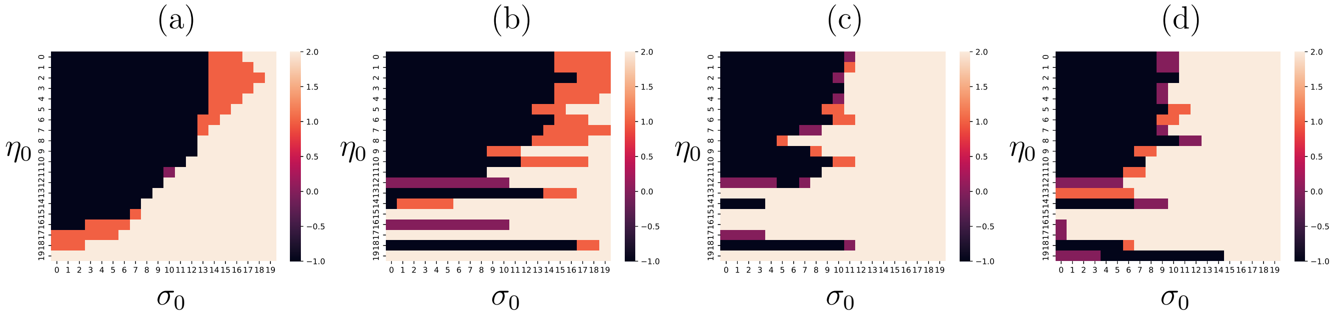

Fig. 1. (Color) Heat map of jump inference in (η_0, σ_0) plane. $L = 18$. Step and direct methods. The abscissa is σ_0 which ranges from 0.1 to 2.0 with increment 0.1 from left to right, and the ordinate on the left is η_0 which ranges from 0.1 to 2.0 with increment 0.1 from top to bottom. $(0, 1, \cdots, 19)$ in the heat map corresponds to $(0.1, 0.2, \cdots, 2.0)$ for the value of η_0 and σ_0. The ordinate on the right corresponds to the color. Black(-1): jump is successfully inferred in both methods, magenta(0): jump is successfully inferred only in the step method, orange(1): jump is successfully inferred only in the direct method, white(2): jump is not inferred successfully in both methods. (a) Type A with $\Delta h = 0.1$, (b), (c), (d) type B with $\Delta h = 0.001$. (b) sample 0, (c) sample 1, (d) sample 2.

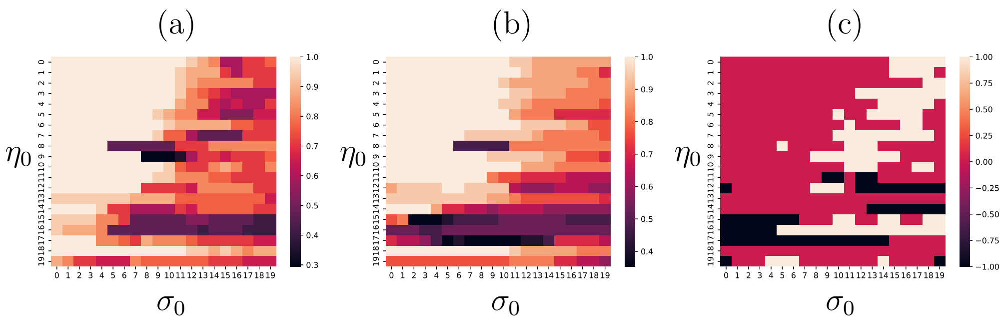

Fig. 2. (Color) Heat map of the ratio $r = l/(L-1)$ in (η_0, σ_0) plane. $L = 18$. Step and direct methods. Type B, sample 1. (a) r^{s}, (b) r^{d}, (c) $R = r^{\mathrm{s}}/r^{\mathrm{d}}$. -1 : $R > 1.1$, 1: $R < 0.9$, 0: $0.9 \leq R \leq 1.1$.

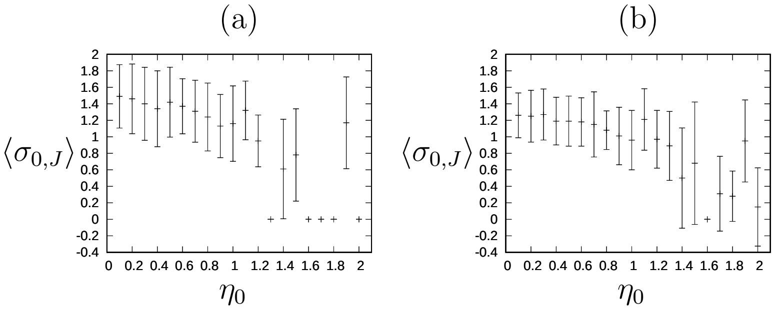

Fig. 3. η_0 dependences of sample averaged $\langle \sigma_{0,J} \rangle$ over 10 samples. $L = 18$. Type B. $\Delta h = 0.001$. (a) direct method, (b) step method.

Error comparison between step and direct methods

We define errors E_1 and E_2 as $E_1 = \sum_{i=1}^{L-1}(m_i - \tau_i)^2 / \sum_{i=1}^{L-1}(m_i)^2$ and $E_2 = \sum_{i=1}^{L-1}(m_i - \hat{m}_i)^2 / \sum_{i=1}^{L-1}(m_i)^2$, where $\hat{m}_i$ is the MPM estimate for m_i. For the step and direct methods, we denote E_2 as E_2^{s} and E_2^{d}, respectively. We show the heat map of $E_2^{\mathrm{s}}/E_2^{\mathrm{d}}$ for types A and B in Fig. 4. As is seen from the figures, there is a tendency that in region where the jump inference succeeds in both methods, the errors for the step method are smaller than or comparable to those for the direct method.

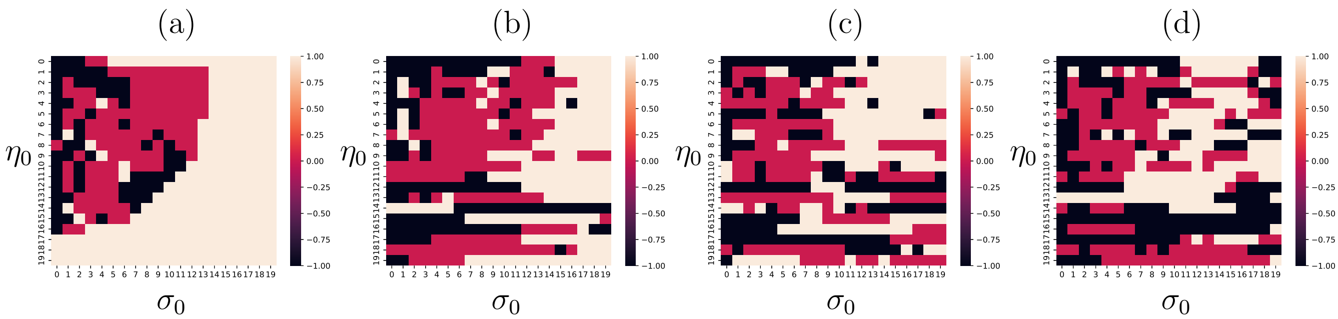

Fig. 4. (Color) Error comparison between the step and direct methods. $L = 18$. Black(-1): $E_2^{\mathrm{s}}/E_2^{\mathrm{d}} < 0.9$, white(1): $E_2^{\mathrm{s}}/E_2^{\mathrm{d}} > 1.1$, magenta(0): $0.9 \leq E_2^{\mathrm{s}}/E_2^{\mathrm{d}} \leq 1.1$. (a): type A, (b), (c), (d): type B. (b) sample 0, (c) sample 1, (d) sample 2.

Optimal hyperparameters

As for h, in almost all regions where the number of jumps is 0, h_{opt} takes the value 10 for type A and 0.1 for type B which are the maximum of the scan range. It is reasonable because as h increases the jump is difficult to take place. Outside the region, it takes small values. As for $\tilde{\eta}$, it takes values from the minimum 0.1 to the maximum 2.0 we scan for both types. From numerical results, as expected, we find that for both methods $\eta_{\rm opt}$ and $\sigma_{\rm opt}$ are almost linearly dependent on η_0 and σ_0, respectively.

Location dependence of MPM inference

We show two examples of the location dependence of the MPM inference for the step and direct methods for type A. The case that the result of the step method is better than that of the direct method is shown in Fig. 5(a), and the opposite case is shown in Fig. 5(b).

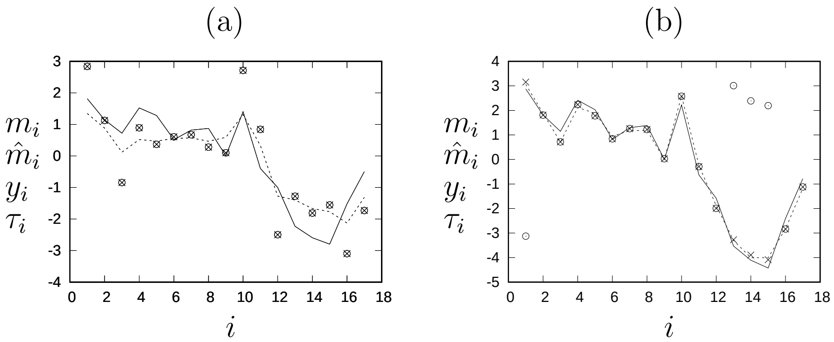

Fig. 5. $L = 18$. Solid line: true phase difference m_i. $\circ$: reduced phase difference τ_i. Dashed line: MPM estimate $\hat{m}_i$ for m_i. $\times$: phase difference with noise y_i. (a) Step method. $\eta_0 = 1.2, \sigma_0 = 1.1, E_1 = 0.49$. $E_2^{\rm s} = 0.20$. For direct method, $E_2^{\rm d} = 1.1$. (b) Direct method. $\eta_0 = 1.9, \sigma_0 = 0.3, E_1 = 1.8$. $E_2^{\rm d} = 0.015$. For step method, $E_2^{\rm s} = 4.4$.

3.2 Regular phase differences obtained from phases

Next, we show the numerical results for phase differences which are generated from regular phases. Even if the phases are regular, the phase differences are rather complicated. Therefore, we apply the theory for random phases developed in this paper to phase differences generated from regular phases.

We generate phases from the following polynomials.

$$\phi_i = \begin{cases} b\frac{L}{12}[1 - \left(\frac{i-x_0}{a}\right)^n] & |i - x_0| \le a, \\ 0 & |i - x_0| > a, \end{cases} \quad i = 1, \cdots, L,$$

$$m_i = \phi_{i+1} - \phi_i, \ i = 1, \cdots, L-1.$$

where $x_0 = \frac{L+1}{2}$, $a = a_0(L-1)$. We set $a_0 = 0.4, 0.6, b = 3, 5$ and $n = 2, 4$. We generate the corrupted phase differences for $\sigma_0 = 0.1, 0.2, \cdots, 2.0$. For the hyperparameter inference, we scan $h = -10, -9.9, \cdots, 10$, $\tilde{\eta}, \eta$ and $\sigma = 0.1, 0.2, \cdots, 2$.

<u>Jump inference</u>

We show the numerical results of the jump inference for $L = 14, 18$ and 20 and $b = 5$. We set $a_0 = 0.4$, and then the phase difference abruptly changes at $|i - \frac{L+1}{2}| \simeq a (= 0.4(L-1))$. In Fig. 6, we show the σ_0 dependence of the ratio r (eq. (33)) for the step and direct methods, r^{s} and r^{d}. The average of phase differences is almost 0. The region where the jump inference succeeds in the step method is the same as that in the direct method for all L.

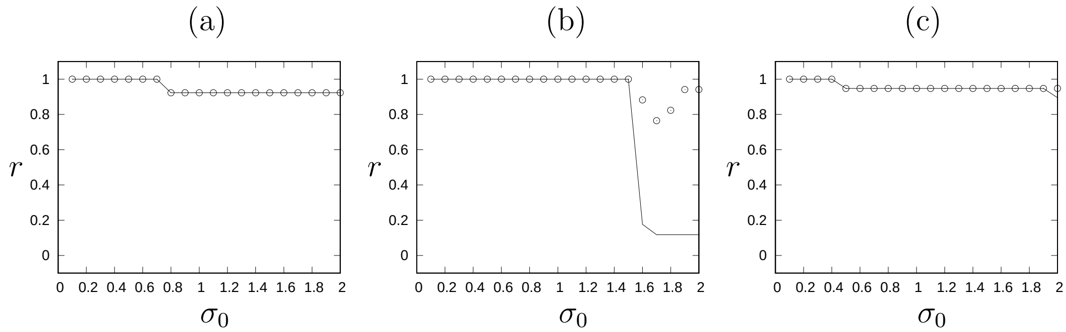

Fig. 6. σ_0 dependence of r. Solid line: r^{s} (step1), ○: r^{d} (direct method). (a) $L = 14$, (b) $L = 18$, (c) $L = 20$.

Error comparison between step and direct methods

We compare errors $E_2^{\rm s}$ and $E_2^{\rm d}$. In Fig. 7, we show the ratio $E_2^{\rm s}/E_2^{\rm d}$. From this figure, we find that the step method is better than or comparable to the direct method for σ_0 where jumps are correctly inferred. This result is similar to that in the random phase difference case.

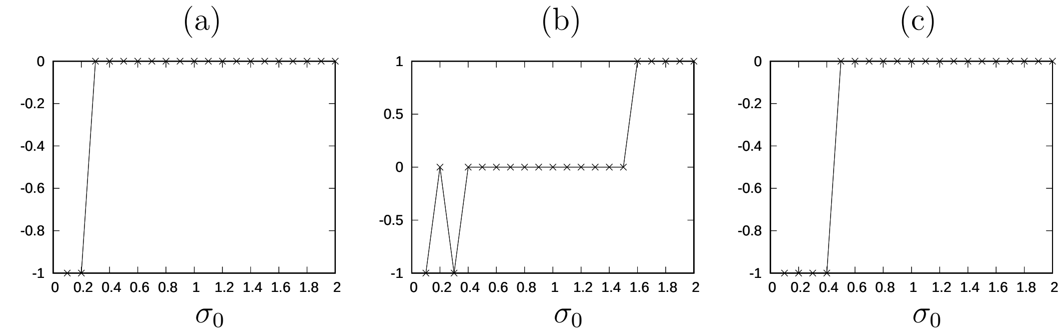

Fig. 7. σ_0 dependences of the ratio of errors $E_2^{\rm s}/E_2^{\rm d}$. -1: $E_2^{\rm s}/E_2^{\rm d} < 0.9$, 0: $0.9 \leq E_2^{\rm s}/E_2^{\rm d} \leq 1.1$, 1: $E_2^{\rm s}/E_2^{\rm d} > 1.1$. (a) $L = 14$, (b) : $L = 18$, (c) : $L = 20$.

Optimal hyperparameters

From numerical results, we find that $h_{\rm opt}$ takes the small positive values for σ_0 where non-zero jumps are inferred correctly, and for σ_0 where zero jump is inferred correctly it takes the value 10 which is the maximum of the scan range. This is the same as in the random phase difference case. Let σ_m be the standard deviation of phase differences. $\sigma_m \simeq 1.4$ for $L = 14, 18, 20$. As for $\tilde{\eta}$, it takes comparable value with the value of σ_m when jump inference succeeds. For the region of σ_0 where jumps are inferred correctly in both methods, $\eta_{\rm opt}$ is nearly equal to σ_m and $\sigma_{\rm opt}$ takes values $0.2 \sim 0.8$. From these results of hyperparameter inference, $\tilde{\eta}_{\rm opt}$ in the step method and $\eta_{\rm opt}$ in the step and direct methods take values similar to σ_m in the region where the jump inference succeeds. Therefore, without performing the hyperparameter inference, we also performed numerical calculations replacing $\tilde{\eta}$ and η by $\langle\tau\rangle$, the average of the reduced phase differences, and obtained similar results. Here, we use $\langle\tau\rangle$ instead of σ_m since observers do not know σ_m. We omit the details.

Location dependence of MPM inference

We show the two cases of the location dependence of the MPM inference for $L = 20$.

One case is that the MPM inference by the step method is better than that by the direct method (Fig. 8(a)) and the other case is that the MPM inferences of both methods are comparable (Fig. 8(b)).

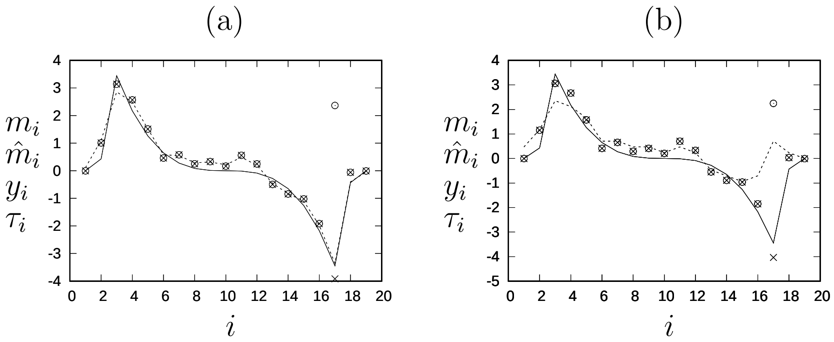

Fig. 8. $L = 20, b = 5$. Notations are the same as in Fig. 5. (a) $\sigma_0 = 0.4, E_2^{\mathrm{s}}/E_2^{\mathrm{d}} < 0.9, E_1 = 0.94$. Step method. $E_2^{\mathrm{s}} = 0.045$. For direct method, $E_2^{\mathrm{d}} = 0.064$. (b) $\sigma_0 = 0.5, 0.9 \leq E_2^{\mathrm{s}}/E_2^{\mathrm{d}} \leq 1.1, E_1 = 0.93$. Direct method. $E_2^{\mathrm{d}} = 0.61$. For step method,$E_2^{\mathrm{s}} = 0.62$.

We also performed numerical calculations for $a_0 = 0.6$ where there is no discontinuity of phase differences and for phases generated from the second order polynomial. We obtained similar results to those shown above and omit them.

In this section, for small system sizes such as L up to 20, we found that the step and direct methods give comparable MPM estimates of phase differences when jumps are correctly inferred. In the next section, we perform the MPM estimate by the step method in the large system sizes where the direct method is not available.

4. Large L cases

Firstly, we study random phase differences of type B. For $L = 65$, we generate 100 samples and show the heat map of the ratio of samples in which the jump inference succeeds at all locations and that of the root mean square error (RMSE) in Fig. 9. RMSE is defined as RMSE=$\sqrt{\frac{1}{L-1}\sum_{l=1}^{L-1}(m_l - \hat{m}_l^{\mathrm{s}})^2}$. In Fig. 10, we show the σ_0 dependence of the sample average of the RMSE. From these figures, we note that in the region $\sqrt{\eta_0^2 + \sigma_0^2} \lesssim 1$ in (η_0, σ_0) plane, the jump inference succeeds.

Next, we study the regular phase differences. For $L = 50, 500$ and $b = 3$ and 5, we show the results of jump inferences in Fig. 11. σ_m is 0.86, 1.43, 0.85, and 1.41, $\sigma_{0,J}$ is 0.5, 0.3, 0.6 and 0.5 for $(L, b) = (50, 3), (50, 5), (500, 3), (500, 5)$, respectively. From these results we find that our theory is applicable to the regular phases and large L

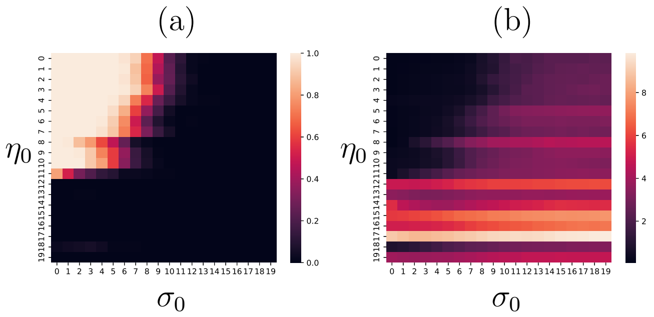

Fig. 9. (Color) Heat map in (η_0, σ_0) plane for type B. $L = 65$. $\Delta h = 0.001$. (a) Ratio of samples in which jump inference succeeds at all locations, (b) sample average of RMSE.

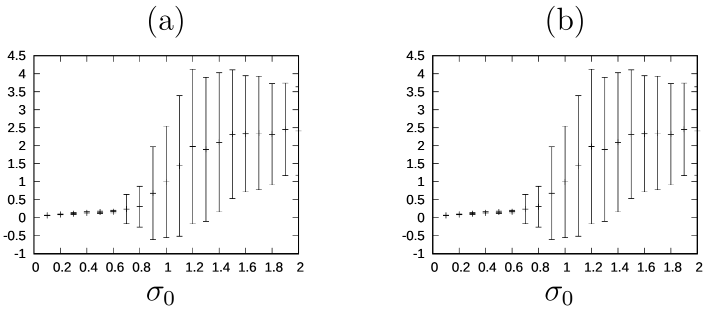

Fig. 10. σ_0 dependence of RMSE. $L = 65$. $\Delta h = 0.001$. (a) $\eta_0 = 0.1$, (b) $\eta_0 = 1.0$.

cases. As for the optimal hyperparameters, there are several results which are different from the small L cases. We find that $h_{\rm opt}$ becomes negative when $r^{\rm s}$ becomes of order of 0.5, and this is quite reasonable because jumps take place easier for negative h as seen from eq. (20). $\tilde{\eta}_{\rm opt}$ increases from the small value to 2.0 almost linearly and then takes the value 2.0 because 2.0 is the maximum value of the parameter scan. We show the location dependences of the phase differences in Fig. 12 for $L = 50$ and $L = 500$ in the cases that the jump inferences succeed.

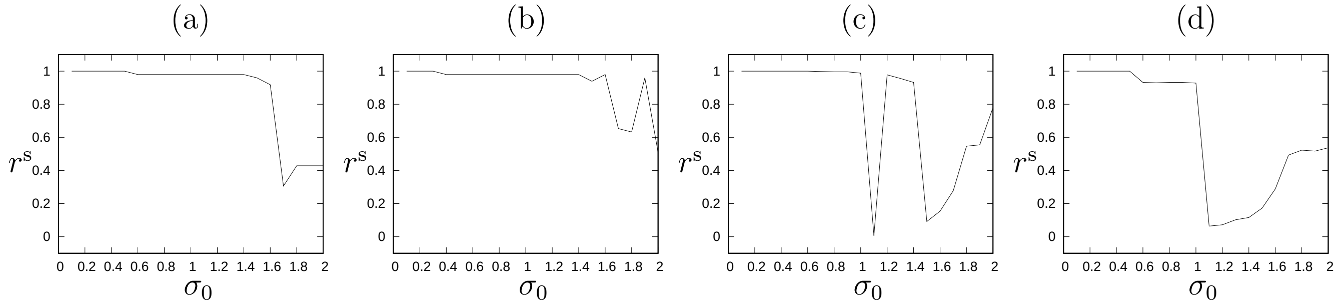

Fig. 11. σ_0 dependences of r^{s}, the ratio of number of locations where the jump inference succeeds to number of total locations. Step method. (a), (b) $L = 50$, (c), (d) $L = 500$. (a), (c) $b = 3$, (b), (d) $b = 5$.

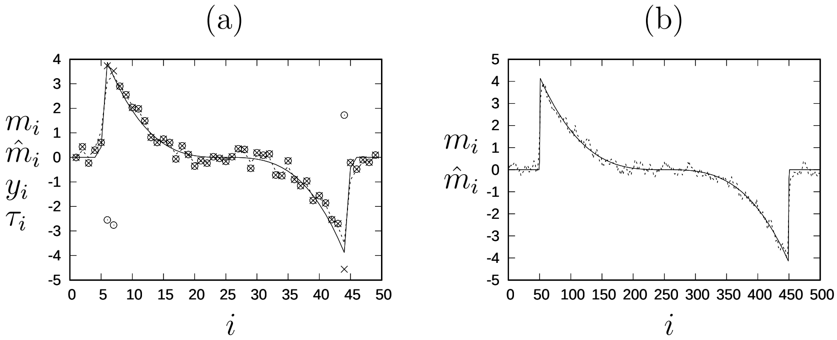

Fig. 12. $a_0 = 0.4, n = 4, b = 5$. Step II. (a) $L = 50$. Notations are the same as in Fig. 5. $\sigma_0 = 0.3, E_2^{\mathrm{s}} = 0.035, E_2^{\mathrm{s}}/E_1 < 0.9$ and jump inference succeeds with non-zero jump. (b) $L = 500$. Solid curve: m_i, dashed curve: MPM estimate of m_i. Corrupted and reduced phase differences are not drawn for clear figures. $\sigma_0 = 0.5, E_2^{\mathrm{s}} = 0.025, E_2^{\mathrm{s}}/E_1 < 0.9$ and jump inference succeeds with non-zero jump.

Now, we show the numerical results when the phase differences and the external noises are generated from the correlated Gaussian distribution and the von-Mises distribution, respectively. We generate 10 samples of external noises for type B. In Ref. 13, the location dependence of phase differences are shown for $N = 64, \alpha = 1, \beta = 5$ and $h = 10^{-3}$. These values of N, α and β correspond to $L = 65, \eta_0 = 1$ and $\sigma_0 = 1/\sqrt{5} \simeq 0.44$ in our notations, respectively. As for h, we scan $h = -0.1, -0.099, \cdots, 0.1$ and the scan range of other hyperparameters are the same as in the previous calculations. In Fig. 13, we show η_0 dependences of $\langle \sigma_{0,J} \rangle$with error bars, where bracket implies the average over 10 samples. As L increases, $\langle \sigma_{0,J} \rangle$ decreases approximately, as expected.

Finally, let us compare our results with previous ones. In Ref. 13, the posterior mean (PM) method is used and the phases are inferred by taking the mean values of phases with respect to the posterior distribution, and the jumps are successfully inferred except for the uniform shift. In our method, from Fig. 13(b), we note that $\langle \sigma_{0,J} \rangle \sim 0.5$ at $\eta_0 = 1.0$. In more details, we found that the jump inference succeeds in 7 of 10 samples for $\eta_0 = 1$ and $\sigma_0 = 0.5$. Since the performance of the jump inference depends on samples and system sizes, it is difficult to judge whether the present method or the method of the previous study is better. In Fig. 14, the location dependence of phase differences is shown when the jump inference succeeds. In this case, the hyperparameters are $\tilde{\eta}_{\rm opt} = 1.1, \eta_{\rm opt} = 1.0$ and $\sigma_{\rm opt} = 0.4$, and similar to the values $\eta = 1.0$ and $\sigma = 0.5$, respectively. $h_{\rm opt} = 0.024$ is small as in the cases of the Gaussian noise when the MPM inferences succeed with non-zero jumps.

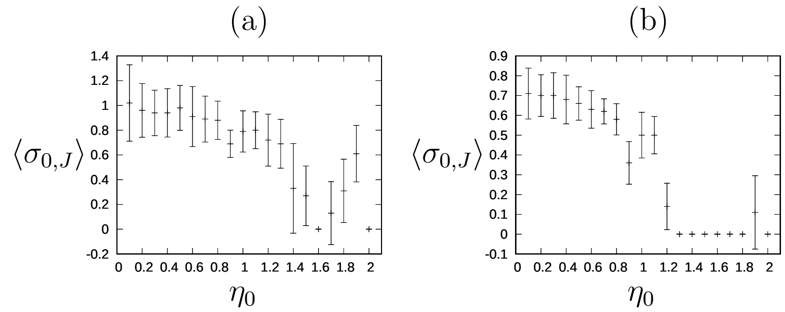

Fig. 13. η_0 dependence of $\langle \sigma_{0,J} \rangle$ with errors. (a) $L = 18$, (b) $L = 65$.

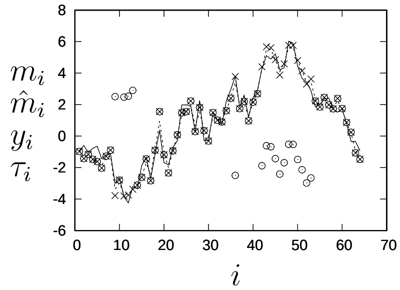

Fig. 14. Location dependence of phase differences. Notations are the same as in Fig. 5. $L = 65, \eta_0 = 1, \sigma_0 = 0.5$. Step II. sample 1. $E_1 = 1.42, E_2^{\rm s} = 0.024$.

5. Summary and discussion

In this paper, we proposed a one-dimensional phase unwrapping model and analyzed it by the transfer matrix (TM) method. We assumed that during the observation process of phase differences, these are corrupted by Gaussian noises and are reduced to values in $[-\pi, \pi)$. Thus, the reduced data may have discontinuities, which we call jumps, as a function of location. In order to deal with the discontinuities, we introduced the three state Potts model assuming that the jumps are not very big. We formulated the problem on the basis of Bayesian inference using the maximizer of the posterior marginal (MPM) estimate.

Firstly, we studied the case that the phase differences are generated from the correlated Gaussian distribution. We proposed two methods, the step and direct methods. In the step method, we inferred the corrupted data from the reduced data in step I, and inferred the phase differences from the inferred corrupted data in step II. In the direct method, we inferred the phase differences from the reduced data directly. In both methods, we derived the expressions of the MPM estimates of the phase differences and the posterior probabilities of the hyperparameters. We performed numerical calculations for several samples of phase differences and external noises. We compared the performance of the step and direct methods for small values of L up to 20, because numerical calculations of the direct method are difficult for values of L larger than 20 as explained below. Here, L is the system size, the number of locations where phases are observed. We found that the MPM estimates of the phase differences for both methods are comparable although the jump inference depends on samples. Next, we applied the present theory to the regular phase differences obtained from phases generated from polynomials. Similar to the case of random phase differences, we found that both methods give comparable results. Then, we studied the large L cases by the step method. We studied the random phase differences for $L = 65$, and the regular phase differences for $L = 50$ and $L = 500$. We found that we can infer the corrupted data and the phase differences fairly well by using inferred hyperparameters in rather wide ranges of noise amplitudes. The ranges depend on samples and the system sizes as in the cases of small system sizes. Finally, we applied the step method to the case that the phases are generated from the correlated Gaussian distributions as in our case, but the external noises are generated from the von-Mises distribution studied in Ref. 13. We studied the regions where the jump inference succeeds and found that in 7 of 10 samples we can infer the jumps successfully at almost the same parameters studied in Ref. 13. Since the performance of the jump inference depends on samples and the system sizes, it is difficult to judge whether the present method or the method of the previous study is better.

Now, let us discuss applicability of the direct and step methods to large system sizes and their computational times. The direct method is applicable to system sizes of at most 20. This is because for the direct method the numbers of summations of exponential functions in the formulae of the MPM estimate (eq. (11)) and the likelihood function (eq. (18)) are extremely large. For example, the numbers are $3^{L-1} = 3^{17} = 129,140,163$ for $L = 18$ and $3^{499} \simeq 10^{238}$ for $L = 500$. No such summations exist in the step method. For $L = 18$, the computational time to calculate the optimal hyperparameters by the raster scan and the MPM estimates for one set of (η_0, σ_0) par sample is about 503 minutes for the direct method and about 1.13 seconds for the step method. That is, the step method is about 26,700 times faster than the direct method. The main reason for this is that summations of exponential functions exist in the direct method and the number of them is of order 10^8. Therefore, the step method has significant advantages of applicability to large system sizes and extremely short computational times.

We restrict ourselves to the one-dimensional model in this paper. By a similar method to the present one, the analyses of the model which is composed of two strings, we call it the ladder model, and the two-dimensional model are under investigation. We will report the results of the investigation in the future.

Acknowledgment

One of the authors, T. U. is grateful to M. Kawamura for valuable discussions, and D. Kuwahara and Y. Hirata for the explanations of the experiments on the one-dimensional phase unwrapping.

Appendix A: Derivation of Eigenvalues

In this Appendix, we derive the eigenvalues and the normalization factor C_0 of the correlated Gaussian distribution (eq. (4)).

$$P(\boldsymbol{m}) = C_0 \exp(-\frac{1}{2\eta_0^2}\sum_{i=1}^{L}(m_i - m_{i-1})^2), \tag{A·1}$$

$$C_0 = \prod_{l=1}^{L-1} \sqrt{\frac{\lambda_l}{2\pi}\frac{1}{\eta_0}}, \tag{A·2}$$

$$\lambda_l = 2(1-\cos(\frac{l\pi}{L})) \quad l = 1,2,\ldots,L-1. \tag{A·3}$$

Introducing the matrix A, $P(\boldsymbol{m})$ is written as

$$P(\boldsymbol{m}) = C_0 \exp(-\frac{1}{2\eta_0^2}\boldsymbol{m}^{\mathrm{T}}A\boldsymbol{m}), \tag{A·4}$$

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -1 & 2 \end{pmatrix}. \tag{A·5}$$

We derive the eigenvalues and eigenvectors of the following $n \times n$ matrix B_n for $n \geq 1$.

$$B_n = \begin{pmatrix} \alpha & \beta & 0 & 0 & \cdots & 0 \\ \beta & \alpha & \beta & 0 & \cdots & 0 \\ 0 & \beta & \alpha & \beta & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta & \alpha & \beta \\ 0 & \cdots & 0 & 0 & \beta & \alpha \end{pmatrix}. \tag{A·6}$$

Defining $b_n = \det B_n$, it is calculated as

$$b_n = \alpha b_{n-1} - \beta^2 b_{n-2},\ n \geq 3,$$
$$b_1 \equiv \alpha.$$

Defining $c_n = b_n - ub_{n-1}$ for $n \geq 2$, we obtain $c_n = vc_{n-1}$, where u and v satisfy $u+v=\alpha, uv=\beta^2$. That is, these are the solutions of the equation $z^2-\alpha z+\beta^2=0$, $u,v=\frac{1}{2}(\alpha \pm \sqrt{\alpha^2-4\beta^2})$. Then, we obtain $c_n = v^{n-2}c_2$, $c_2 = b_2 - ub_1 = v^2$. Therefore, $c_n = v^n$ for $n \geq 2$. Thus, $b_n = v^n + ub_{n-1}$ for $n \geq 2$. From this relation, we obtain $b_n = \sum_{k=0}^{n} v^k u^{n-k}$ for $n \geq 1$. Now, let us derive the eigenvalue λ of the matrix A. The characteristic equation of A is $\det(\lambda E - A) = 0$, where E is the unit matrix. This is equal to $b_{L-1} = 0$ with $\alpha = \lambda - 2$ and $\beta = 1$. Now, we show $\lambda \neq 0$ and $\lambda \neq 4$. If $\lambda = 0$ or $\lambda = 4$, we find $\alpha^2 = 4$ and $u = v = \alpha/2 = \pm 1$. Therefore, $b_{L-1} = Lu^{L-1} \neq 0$. Thus, $\lambda \neq 0$ and $\lambda \neq 4$. Then, u and v are different and are not equal to 0. Let us put $u = re^{i\theta}$

where r and θ are real numbers. It follows $v = e^{-i\theta}/r$. Thus, we obtain

$$b_{L-1} = \sum_{k=0}^{L-1} v^k u^{L-1-k} = u^{L-1}\sum_{k=0}^{L-1}(\frac{v}{u})^k = u^{L-1}\frac{1-(\frac{v}{u})^L}{1-\frac{v}{u}} = u^{L-1}\frac{1-(e^{-2i\theta}/r^2)^L}{1-e^{-2i\theta}/r^2}.$$

Note that $\frac{v}{u} = e^{-2i\theta}/r^2 \neq 1$. From $b_{L-1} = 0$, it follows $(e^{-2i\theta}/r^2)^L = 1$. Thus, we obtain $r = 1, e^{-2i\theta} = \exp(\frac{2\pi l}{L}i), l = 1, 2, \cdots, L-1$. Therefore, $u_l = e^{i\theta_l} = \exp(-\frac{\pi l}{L}i), l = 1, 2, \cdots, L-1$. From the relation $u + v = \lambda - 2$, we obtain $\lambda_l = 2 + e^{i\theta_l} + e^{-i\theta_l} = 2 + 2\cos(\frac{\pi l}{L})$. By changing the label l to $L-l$, we obtain $\lambda_l = 2(1+\cos(\frac{\pi(L-l)}{L})) = 2(1-\cos(\frac{\pi l}{L}))$. As easily checked, the normalized eigenvector $\boldsymbol{t}_k$ belonging to $\lambda_l = 2(1-\cos(\frac{\pi l}{L}))$ is given by

$$\boldsymbol{t}_k = \sqrt{\frac{2}{L}}\begin{pmatrix} \sin\left(\frac{k\pi}{L}\right) \\ \vdots \\ \sin\left(\frac{lk\pi}{L}\right) \ \ (l \text{ th row}) \\ \vdots \\ \sin\left(\frac{(L-1)k\pi}{L}\right) \end{pmatrix}.$$

Defining $U = (\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_{L-1})$ and $\boldsymbol{v} = U^{\mathrm{T}}\boldsymbol{m}$, we obtain $\boldsymbol{m}^{\mathrm{T}}A\boldsymbol{m} = \sum_{l=1}^{L-1}\lambda_l v_l^2$. Thus, we obtain

$$C_0^{-1} = \int d\boldsymbol{m}\exp[-\frac{1}{2\eta_0^2}\boldsymbol{m}^{\mathrm{T}}A\boldsymbol{m}] = \int d\boldsymbol{v}\exp[-\frac{1}{2\eta_0^2}\sum_{l=1}^{L-1}\lambda_l v_l^2] = \prod_{l=1}^{L-1}\sqrt{\frac{2\pi}{\lambda_l}}\eta_0.$$

Appendix B: Derivation of conditional probability

Here, we derive the conditional probability $P(\tau_l|x_l)$, eq. (9). Since $y_l = x_l + z_l = \tau_l + 2\pi n_l$, the joint probability distribution of τ_l and n_l is given by

$$P(\tau_l, n_l|x_l) = P(\tau_l + 2\pi n_l|x_l, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{(\tau_l + 2\pi n_l - x_l)^2}{2\sigma^2}]. \quad \text{(B·1)}$$

Thus, we obtain

$$P(\tau_l|x_l) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{n_l=-\infty}^{\infty}\exp[-\frac{1}{2\sigma^2}(\tau_l + 2\pi n_l - x_l)^2]. \quad \text{(B·2)}$$

The normalization of $P(\tau_l|x_l)$ is shown as follows.

$$\begin{aligned}\int_{-\pi}^{\pi}P(\tau_l|x_l)d\tau_l &= \frac{1}{\sqrt{2\pi}\sigma}\sum_{n_l=-\infty}^{\infty}\int_{-\pi}^{\pi}\exp[-\frac{1}{2\sigma^2}(\tau_l + 2\pi n_l - x_l)^2]d\tau_l \\ &= \frac{1}{\sqrt{2\pi}\sigma}\sum_{n_l=-\infty}^{\infty}\int_{(2n_l-1)\pi}^{(2n_l+1)\pi}\exp[-\frac{1}{2\sigma^2}(t_l - x_l)^2]dt_l \\ &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp[-\frac{1}{2\sigma^2}(t_l - x_l)^2]dt_l = 1,\end{aligned}$$

where $t_l = \tau_l + 2\pi n_l$.

Appendix C: Derivation of marginal distributions for direct method

Firstly, we give several expressions of concerned variables in the following.

$$\begin{aligned}
P(\boldsymbol{\tau}|\boldsymbol{x})P(\boldsymbol{x}) &= C(\frac{1}{\sqrt{2\pi}\sigma})^{L-1}\sum_{\boldsymbol{n}}{}' e^{-\mathcal{H}_{\boldsymbol{n}}}, \\
\mathcal{H}_{\boldsymbol{n}} &= \frac{1}{2\sigma^2}\sum_{l=1}^{L-1}(\tau_l + 2\pi n_l - x_l)^2 + \frac{1}{2\eta^2}\sum_{l=0}^{L-1}(x_l - x_{l+1})^2, \\
Z_{\mathrm{d}} &= \int d\boldsymbol{x}P(\boldsymbol{\tau}|\boldsymbol{x})P(\boldsymbol{x}) = C(\frac{1}{\sqrt{2\pi}\sigma})^{L-1}\sum_{\boldsymbol{n}}{}'\int d\boldsymbol{x}e^{-\mathcal{H}_{\boldsymbol{n}}}, \\
e^{-\mathcal{H}_{\boldsymbol{n}}} &\equiv \prod_{l=0}^{L-1}\Psi_l(x_l, x_{l+1}, n_l), \\
\Psi_l(x_l, x_{l+1}, n_l) &= e^{-\frac{1}{2\sigma^2}(\tau_l+2\pi n_l - x_l)^2 - \frac{1}{2\eta^2}(x_l - x_{l+1})^2},\ l = 1, \cdots, L-2, \\
\Psi_0(x_1) &= e^{-\frac{1}{2\eta^2}x_1^2}, \\
\Psi_{L-1}(x_{L-1}, n_{L-1}) &= \exp[-\frac{1}{2\sigma^2}(\tau_{L-1} + 2\pi n_{L-1} - x_{L-1})^2 - \frac{1}{2\eta^2}x_{L-1}^2],
\end{aligned}$$

where $x_0 = x_L = 0$. For later use, we define $u_l(x_l), d_l(x_l)$, $a_l^{\mathrm{u}}, a_l^{\mathrm{d}}$ etc. as follows.

$$u_1(x_1) = \Psi_0(x_1), \tag{C·1}$$

$$u_{l+1}(x_{l+1}) = \int dx_l \Psi_l(x_l, x_{l+1}, n_l)u_l(x_l) \tag{C·2}$$

$$= \int dx_l \exp[-\frac{a_l^{\mathrm{u}}}{2}x_l^2 + (\frac{x_{l+1}}{\eta^2} + f_l^{\mathrm{u}})x_l - \frac{x_{l+1}^2}{2\eta^2} + h_l^{\mathrm{u}}],\ l = 1, 2, \cdots, L-1, \tag{C·3}$$

$$d_{L-1}(x_{L-1}) = \Psi_{L-1}(x_{L-1}, n_{L-1}), \tag{C·4}$$

$$d_{l-1}(x_{l-1}) = \int dx_l \Psi_{l-1}(x_{l-1}, x_l, n_{l-1})d_l(x_l) \tag{C·5}$$

$$= \int dx_l \exp[-\frac{a_l^{\mathrm{d}}}{2}x_l^2 + (\frac{x_{l-1}}{\eta^2} + f_l^{\mathrm{d}})x_l - \frac{x_{l-1}^2}{2\eta^2} + h_l^{\mathrm{d}}],\ l = L-1, L-2, \cdots, 1. \tag{C·6}$$

Now, let us calculate the marginal distributions.

$$P(x_l|\boldsymbol{\tau}) = \frac{1}{Z_{\mathrm{d}}}C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}'\int d\boldsymbol{x}^{(l)}\exp(-\mathcal{H}_{\boldsymbol{n}})$$

$$= C(\frac{1}{\sqrt{2\pi}\sigma})^{L-1}\sum_{\boldsymbol{n}}{}' \int dx_{i-1}dx_{i+1}u_{i-1}(x_{i-1})d_{i+1}(x_{i+1}) \times\Psi_{i-1}(x_{i-1},x_i,n_{i-1})\Psi_i(x_i,x_{i+1},n_i)$$

$$\equiv C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}' \mu_l(x_l,\boldsymbol{n}),\ \ l=2,3,\cdots,L-2, \quad \text{(C·7)}$$

$$P(x_1|\boldsymbol{\tau}) = \frac{1}{Z_{\rm d}}C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}' d_1(x_1,\boldsymbol{n}) \equiv \frac{1}{Z_{\rm d}}C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}' \mu_1(x_1,\boldsymbol{n}), \quad \text{(C·8)}$$

$$P(x_{L-1}|\boldsymbol{\tau}) = \frac{1}{Z_{\rm d}}C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}' u_{L-1}(x_{L-1},\boldsymbol{n}) \equiv \frac{1}{Z_{\rm d}}C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\sum_{\boldsymbol{n}}{}' \mu_{L-1}(x_{L-1},\boldsymbol{n}). \quad \text{(C·9)}$$

Since $\mu_l(x_l,\boldsymbol{n})$s are all Gaussian distributions, we denote them as

$$\mu_l(x_l,\boldsymbol{n}) \propto e^{-\frac{1}{2}a_l^{(\boldsymbol{n})}x_l^2+b_l^{(\boldsymbol{n})}x_l+c_l^{(\boldsymbol{n})}}, \quad \text{(C·10)}$$

where $a_l^{(\boldsymbol{n})}, b_l^{(\boldsymbol{n})}$ and $c_l^{(\boldsymbol{n})}$ depend on $\boldsymbol{n}$. For any l among $1,2,3,\cdots,L-1$, $Z_{\rm d}$ is expressed as

$$Z_{\rm d} = \int dx_l P(x_l|\boldsymbol{\tau}) = C(\frac{1}{\sqrt{2\pi}\sigma})^{L-1}\sum_{\boldsymbol{n}}{}' \int dx_l\mu_l(x_l,\boldsymbol{n}) \quad \text{(C·11)}$$

$$= C\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{L-1}\left[\prod_{j<l}\sqrt{\frac{2\pi}{a_j^{\rm u}}}\right]\left[\prod_{j>l}\sqrt{\frac{2\pi}{a_j^{\rm d}}}\right]\sum_{\boldsymbol{n}}{}' \sqrt{\frac{2\pi}{a_l^{(\boldsymbol{n})}}}e^{\frac{\left(b_l^{(\boldsymbol{n})}\right)^2}{2a_l^{(\boldsymbol{n})}}+c_l^{(\boldsymbol{n})}}. \quad \text{(C·12)}$$

$a_l^{(\boldsymbol{n})}, b_l^{(\boldsymbol{n})}$ and $c_l^{(\boldsymbol{n})}$ are defined as follows.

$$a_1^{(\boldsymbol{n})} = -\frac{1}{\eta^4 a_2^{\rm d}}+\frac{2}{\eta^2}+\frac{1}{\sigma^2},$$

$$a_l^{(\boldsymbol{n})} = -\frac{1}{\eta^4 a_{l-1}^{\rm u}}-\frac{1}{\eta^4 a_{l+1}^{\rm d}}+\frac{2}{\eta^2}+\frac{1}{\sigma^2} \quad (l=2,\cdots,L-2),$$

$$a_{L-1}^{(\boldsymbol{n})} = -\frac{1}{\eta^4 a_{L-2}^{\rm u}}+\frac{2}{\eta^2}+\frac{1}{\sigma^2},$$

$$b_1^{(\boldsymbol{n})} = \frac{f_2^{\rm d}}{\eta^2 a_2^{\rm d}}+\frac{\tau_1+2\pi n_1}{\sigma^2},$$

$$b_l^{(\boldsymbol{n})} = \frac{f_{l-1}^{\rm u}}{\eta^2 a_{l-1}^{\rm u}}+\frac{f_{l+1}^{\rm d}}{\eta^2 a_{l+1}^{\rm d}}+\frac{\tau_l+2\pi n_l}{\sigma^2} \quad (l=2,\cdots,L-2),$$

$$b_{L-1}^{(\boldsymbol{n})} = \frac{f_{L-2}^{\rm u}}{\eta^2 a_{L-2}^{\rm u}}+\frac{\tau_{L-1}+2\pi n_{L-1}}{\sigma^2},$$

$$c_1^{(\boldsymbol{n})} = \frac{(f_2^{\rm d})^2}{2a_2^{\rm d}}+h_2^{\rm d}-\frac{(\tau_1+2\pi n_1)^2}{2\sigma^2},$$

$$
\begin{aligned}
c_l^{(\boldsymbol{n})} &= \frac{(f_{l-1}^{\rm u})^2}{2a_{l-1}^{\rm u}} + \frac{(f_{l+1}^{\rm d})^2}{2a_{l+1}^{\rm d}} + h_{l-1}^{\rm u} + h_{l+1}^{\rm d} - \frac{(\tau_l + 2\pi n_l)^2}{2\sigma^2} \quad (l = 2, \cdots, L-2), \\
c_{L-1}^{(\boldsymbol{n})} &= \frac{(f_{L-2}^{\rm u})^2}{2a_{L-2}^{\rm u}} + h_{L-2}^{\rm u} - \frac{(\tau_{L-1} + 2\pi n_{L-1})^2}{2\sigma^2}.
\end{aligned}
$$

From eqs. (C·1) - (C·6), the following recursion relations for $a_i^{\rm u}, a_i^{\rm d}$ etc. are obtained.

$$
\begin{aligned}
a_1^{\rm u} &= a_{L-1}^{\rm d} = \frac{1}{\sigma^2} + \frac{2}{\eta^2}, \\
f_1^{\rm u} &= \frac{\tau_1 + 2\pi n_1}{\sigma^2}, \\
f_{L-1}^{\rm d} &= \frac{\tau_{L-1} + 2\pi n_{L-1}}{\sigma^2}, \\
h_1^{\rm u} &= -\frac{(\tau_1 + 2\pi n_1)^2}{2\sigma^2}, \\
h_{L-1}^{\rm d} &= -\frac{(\tau_{L-1} + 2\pi n_{L-1})^2}{2\sigma^2}, \\
a_{i+1}^{\rm u} &= \frac{1}{\sigma^2} + \frac{2}{\eta^2} - \frac{1}{\eta^4 a_i^{\rm u}} \quad (i = 1, \cdots, l-2), \\
a_{i-1}^{\rm d} &= \frac{1}{\sigma^2} + \frac{2}{\eta^2} - \frac{1}{\eta^4 a_i^{\rm d}} \quad (i = l+2, \cdots, L-1), \\
f_{i+1}^{\rm u} &= \frac{\tau_{i+1} + 2\pi n_{i+1}}{\sigma^2} + \frac{f_i^{\rm u}}{\eta^2 a_i^{\rm u}} \quad (i = 1, \cdots, l-2), \\
f_{i-1}^{\rm u} &= \frac{\tau_{i-1} + 2\pi n_{i-1}}{\sigma^2} + \frac{f_i^{\rm d}}{\eta^2 a_i^{\rm d}} \quad (i = l+2, \cdots, L-1), \\
h_{i+1}^{\rm u} &= -\frac{(\tau_{i+1} + 2\pi n_{i+1})^2}{\sigma^2} + \frac{(f_i^{\rm u})^2}{2a_i^{\rm u}} + h_i^{\rm u} \quad (i = 1, \cdots, l-2), \\
h_{i-1}^{\rm d} &= -\frac{(\tau_{i-1} + 2\pi n_{i-1})^2}{\sigma^2} + \frac{(f_i^{\rm d})^2}{2a_i^{\rm d}} + h_i^{\rm d} \quad (i = l+2, \cdots, L-1).
\end{aligned}
$$

Appendix D: Derivation of MPM estimate for step method, step I

We rewrite $\mathcal{H}_{\rm I}$ as

$$\mathcal{H}_{\rm I} = -\mathcal{L} - Y, \tag{D·1}$$

where $\mathcal{L}$ and Y are given as follows.

$$\mathcal{L} = \sum_{i=1}^{L-1} (a_i s_i + b_i s_i^2 - c_i s_i s_{i+1}), \tag{D·2}$$

$$Y = -\frac{1}{2\tilde{\eta}^2} \sum_{i=1}^{L} (\tau_i - \tau_{i-1})^2, \tag{D·3}$$

$$a_i = -\frac{2\pi}{\tilde{\eta}^2}(2\tau_i - \tau_{i+1} - \tau_{i-1}), \ (i = 1, \cdots, L-1), \tag{D·4}$$

$$b_i = -\frac{4\pi^2}{\tilde{\eta}^2} - h = b,\ (i = 1, \cdots, L-1), \tag{D·5}$$

$$c_i = -\frac{4\pi^2}{\tilde{\eta}^2},\ (i = 1, \cdots, L-2),\ c_{L-1} = 0. \tag{D·6}$$

Let us denote the summation of $s_l = 0, \pm 1$ as Tr_l. Then, the partition function Z_I is

$$Z_\mathrm{I} = \mathrm{Tr}_{L-1}\mathrm{Tr}_{L-2}\cdots\mathrm{Tr}_1 e^{-\mathcal{H}_\mathrm{I}} = e^Y \mathrm{Tr}_{L-1}\mathrm{Tr}_{L-2}\cdots\mathrm{Tr}_1 e^{\mathcal{L}}. \tag{D·7}$$

Let $\mathcal{G}_1$ be the term which includes s_1 and $\overline{\mathcal{G}}_1$ be the other terms in $\mathcal{L}$.

$$\mathcal{L} = \mathcal{G}_1 + \overline{\mathcal{G}}_1. \tag{D·8}$$

We define $\mathcal{L}_l$ as

$$\mathcal{L}_l \equiv \sum_{i=l+1}^{L-1} (a_i s_i + b_i s_i^2 - c_i s_i s_{i+1}). \tag{D·9}$$

Thus, we have

$$\mathcal{G}_1 = a_1 s_1 + b_1 s_1^2 - c_1 s_1 s_2, \quad \overline{\mathcal{G}}_1 = \mathcal{L}_1. \tag{D·10}$$

Note that $\mathcal{L} = \mathcal{L}_0,\ \mathcal{L}_{L-1} = 0$. $\mathrm{Tr}_1\ e^{\mathcal{G}_1}$ is given by

$$\mathrm{Tr}_1\ e^{\mathcal{G}_1} = 1 + 2e^{b_1}\cosh(a_1 - c_1 s_2). \tag{D·11}$$

Since $s_i = 0, \pm 1$, we obtain

$$e^{Bs_i} = 1 + \sinh(B)s_i + (\cosh(B) - 1)s_i^2, \tag{D·12}$$

$$1 + A\cosh(B + Cs_i) = De^{Es_i + Fs_i^2}, \tag{D·13}$$

where

$$D = 1 + A\cosh(B), \tag{D·14}$$

$$E = \frac{1}{2}\ln\left[\frac{1 + A\cosh(B+C)}{1 + A\cosh(B-C)}\right], \tag{D·15}$$

$$F = \frac{1}{2}\ln\left[\frac{(1 + A\cosh(B+C))(1 + A\cosh(B-C))}{(1 + A\cosh(B))^2}\right]. \tag{D·16}$$

Now, we use the TM method. For $l = 2, \cdots, L$, we successively define $\mathcal{G}_l + \overline{\mathcal{G}}_l$ as

$$\mathrm{Tr}_l e^{\mathcal{G}_l + \overline{\mathcal{G}}_l} = D_l e^{\mathcal{G}_{l+1} + \overline{\mathcal{G}}_{l+1}},\ l = 1, 2, 3, \cdots, L-1. \tag{D·17}$$

Thus, we obtain

$$\mathcal{G}_l = u_l s_l + v_l s_l^2 - w_l s_l s_{l+1},\ l = 1, \cdots, L-1, \tag{D·18}$$

$$\overline{\mathcal{G}}_l = \mathcal{L}_l,\ l = 1, \cdots, L-1, \tag{D·19}$$

$$(u_l, v_l, w_l) = (a_l + E_{l-1}, b_l + F_{l-1}, c_l),\ l = 2, \cdots, L-1, \tag{D·20}$$

where we define

$$D_l = D(u_l, v_l) \equiv 1 + 2e^{v_l}\cosh(u_l),\ \ l = 1, \cdots, L-1, \tag{D·21}$$

$$E_l = E(u_l, v_l, w_l) = \frac{1}{2}\ln\frac{D(u_l - w_l, v_l)}{D(u_l + w_l, v_l)},\ \ l = 1, \cdots, L-1, \tag{D·22}$$

$$F_l = F(u_l, v_l, w_l) = \frac{1}{2}\ln\left[\frac{D(u_l - w_l, v_l)(D(u_l + w_l, v_l)}{D(u_l, v_l)^2}\right],\ \ l = 1, \cdots, L-1. \tag{D·23}$$

From eq. (D·10), we have

$$(u_1, v_1, w_1) = (a_1, b_1, c_1). \tag{D·24}$$

Thus, we obtain

$$Z_{\rm I} \ = \ e^Y D_1 D_2 D_3 \cdots D_{L-2} D_{L-1}.$$

Let us denote the mapping (D·20) as

$$\boldsymbol{u}_{l+1} \ = \ \boldsymbol{a}_{l+1} + \boldsymbol{G}(\boldsymbol{u}_l),\ \ l = 1, 2, \cdots, L-2, \tag{D·25}$$

$$\boldsymbol{u}_l = (u_l, v_l, w_l)^{\rm T},\ \ \boldsymbol{a}_l = (a_l, b, c_l)^{\rm T}, \tag{D·26}$$

$$\boldsymbol{G}(\boldsymbol{u}_l) = (E(\boldsymbol{u}_l), F(\boldsymbol{u}_l), 0)^{\rm T}, \tag{D·27}$$

where $\boldsymbol{u}_l, \boldsymbol{a}_l$ and $\boldsymbol{G}(\boldsymbol{u}_l)$ are column vectors and T denotes the transpose. We obtain

$$w_l \ = \ c_l = -\frac{4\pi^2}{\tilde{\eta}^2},\ \ l = 1, \cdots, L-2, \tag{D·28}$$

$$w_{L-1} \ = \ c_{L-1} = 0. \tag{D·29}$$

The average values of s_r and s_r^2 $(r = 1, \cdots, L-1)$ are calculated as

$$\begin{aligned} \langle s_r \rangle \ &= \ \frac{1}{Z_{\rm I}}{\rm Tr}_{L-1}{\rm Tr}_{L-2}\cdots{\rm Tr}_1(s_r e^{-\mathcal{H}_{\rm I}}) \\ &= \ \frac{\partial}{\partial u_r}\ln Z_{\rm I} = \sum_{k=r}^{L-1}\frac{1}{D_k}\frac{\partial D_k}{\partial u_r}, \end{aligned} \tag{D·30}$$

$$\langle s_r^2 \rangle \ = \ \frac{\partial}{\partial v_r}\ln Z_{\rm I} = \sum_{k=r}^{L-1}\frac{1}{D_k}\frac{\partial D_k}{\partial v_r}. \tag{D·31}$$

In order to estimate $\langle s_r \rangle$ and $\langle s_r^2 \rangle$, the expressions for $\frac{\partial D_k}{\partial u_l}$ and $\frac{\partial D_k}{\partial v_l}$ are necessary. These are given as

$$\begin{aligned} \left(\frac{\partial D_k}{\partial u_l}, \frac{\partial D_k}{\partial v_l}\right) \ &= \ \big(\bar{D}_u(u_k, v_k), \bar{D}_v(u_k, v_k)\big) \\ &\quad \times dG(\boldsymbol{u}_{k-1})dG(\boldsymbol{u}_{k-2})\cdots dG(\boldsymbol{u}_{l+1})dG(\boldsymbol{u}_l), \end{aligned} \tag{D·32}$$

$$(1 \le l \le L-1,\ l < k \le L-1),$$

where

$$\bar{D}_u(u,v) \equiv \frac{\partial D(u,v)}{\partial u} = 2e^v \sinh(u), \tag{D·33}$$

$$\bar{D}_v(u,v) \equiv \frac{\partial D(u,v)}{\partial v} = 2e^v \cosh(u), \tag{D·34}$$

$$dG(\boldsymbol{u}) = \begin{pmatrix} \frac{\partial E(\boldsymbol{u})}{\partial u} & \frac{\partial E(\boldsymbol{u})}{\partial v} \\ \frac{\partial F(\boldsymbol{u})}{\partial u} & \frac{\partial F(\boldsymbol{u})}{\partial v} \end{pmatrix},\ \boldsymbol{u} = (u,v,w). \tag{D·35}$$

Appendix E: Derivation of MPM estimate for step method, step II

Let us calculate $\langle x_i \rangle$. $-\mathcal{H}_{\rm II}$ is rewritten as

$$-\mathcal{H}_{\rm II} = \mathcal{L} - Y, \tag{E·1}$$

where

$$\mathcal{L} = \sum_{i=1}^{L-1} \Bigl(-\frac{a_i}{2} x_i^2 + b_i x_i + c_i x_i x_{i+1} \Bigr), \tag{E·2}$$

$$Y = \frac{1}{2\sigma^2} \sum_{i=1}^{L-1} y_i^2, \tag{E·3}$$

$$a_i = \frac{1}{\sigma^2} + \frac{2}{\eta^2} = a\ (i = 1, \cdots, L-1), \tag{E·4}$$

$$b_i = \frac{1}{\sigma^2} y_i\ (i = 1, \cdots, L-1),\ c_i = \frac{1}{\eta^2}\ (i = 1, \cdots, L-2),\ c_{L-1} = 0. \tag{E·5}$$

Let $\mathcal{F}_l$ be the integration operator with respect to x_l. Then, we have

$$\mathcal{F}_{L-1} \circ \mathcal{F}_{L-2} \circ \cdots \mathcal{F}_1[e^{-\mathcal{H}_{\rm II}}] = e^{-Y} \mathcal{F}_{L-1} \circ \mathcal{F}_{L-2} \circ \cdots \mathcal{F}_1[e^{\mathcal{L}}]. \tag{E·6}$$

In $\mathcal{L}$, let $\mathcal{G}_1$ be terms which include x_1 and $\overline{\mathcal{G}}_1$ be other terms.

$$\mathcal{L} = \mathcal{G}_1 + \overline{\mathcal{G}}_1. \tag{E·7}$$

We successively define $\mathcal{G}_l + \overline{\mathcal{G}}_l$ $(l = 2, \cdots, L)$ as

$$\mathcal{F}_l[e^{\mathcal{G}_l + \overline{\mathcal{G}}_l}] = A_l e^{\mathcal{G}_{l+1} + \overline{\mathcal{G}}_{l+1}},\ l = 1, 2, \cdots, L-1, \tag{E·8}$$

where $\mathcal{G}_l$ consists of terms which have x_l and $\overline{\mathcal{G}}_l$ consists of other terms. Thus, we have

$$\mathcal{G}_l = -\frac{u_l}{2} x_l^2 + P_l x_l + w_l x_l x_{l+1},\ l = 2, \cdots, L-1,\ \mathcal{G}_L = 0, \tag{E·9}$$

$$\overline{\mathcal{G}}_l = \mathcal{L}_l + \sum_{i=1}^{l-1} \frac{P_i^2}{2u_i},\ l = 2, \cdots, L-1,\ \overline{\mathcal{G}}_L = \sum_{i=1}^{L-1} \frac{P_i^2}{2u_i}, \tag{E·10}$$

$$A_l = \sqrt{\frac{2\pi}{u_l}},\ l = 1, \cdots, L-1, \tag{E·11}$$

$$u_1 = a_1,\ P_1 = b_1,\ w_1 = c_1, \tag{E·12}$$

$$(u_l, P_l, w_l) = (a - \frac{w_{l-1}^2}{u_{l-1}}, b_l + \frac{P_{l-1}w_{l-1}}{u_{l-1}}, c_l),\ l = 2, \cdots, L-1. \tag{E·13}$$

We obtain the following relations.

$$w_l = c_l = \frac{1}{\eta^2},\ l = 1, \cdots, L-2,\ w_{L-1} = c_{L-1} = 0, \tag{E·14}$$

$$u_{l+1} = a - \frac{1}{\eta^4}\frac{1}{u_l} \equiv g(u_l),\ l = 1, \cdots, L-2, \tag{E·15}$$

$$P_{l+1} = b_{l+1} + \frac{1}{\eta^2}\frac{1}{u_l}P_l,\ l = 1, \cdots, L-2. \tag{E·16}$$

In order that the integration by x_i converges, $u_l > 0$ should hold for $l = 1, \cdots, L-1$. Let us prove this. Since $a = \frac{1}{\sigma^2} + \frac{2}{\eta^2}$, we have $u_1 = a > 0$. The fixed points of the mapping $u_{l+1} = g(u_l)$ are $u_\pm = \frac{1}{2}\{a \pm \sqrt{a^2 - (\frac{2}{\eta^2})^2}\}$. Both solutions are real and positive, and u_+ is stable and u_- is unstable. Since $u_+ < a = u_1$, u_l is positive and monotonically decreases, and tends to u_+ as $l \to \infty$. This completes the proof. Therefore, we obtain

$$Z_{\mathrm{II}} = e^{-Y} A_1 A_2 \cdots A_{L-2} A_{L-1} e^{\sum_{i=1}^{L-1} \frac{P_i^2}{2u_i}}. \tag{E·17}$$

By a similar procedure, we derive the expression of the average $\langle x_r \rangle$ as

$$\langle x_r \rangle = \eta^2 \sum_{k=r}^{L-1} \left(\prod_{m=r}^{k} \frac{1}{\hat{u}_m} \right) P_k. \tag{E·18}$$

where we define $\hat{u}_i = \eta^2 u_i$.

References

1) D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software* (Wiley, New York, 1998).
2) D. L. Fried, J. Opt. Soc. Am. **67** (1977) 370.
3) R. H. Hudgin, J. Opt. Soc. Am. **67** (1977) 375.
4) R. J. Noll, J. Opt. Soc. Am. **68** (1978) 139.
5) R. M. Goldstein, and H. A. Zebker, Nature **328** (1987) 707.
6) J. L. Marroquin, and M. Rivera, J. Opt. Soc. Am. **12** (1995) 2393.
7) L. Guerriero, G. Nico, G. Pasquariello, and R. Stramaglia, Appl. Opt. **37**(2000) 3053.
8) G. Nico, G. Palubinskas, and M. Datcu, IEEE Trans. Signal-Processing, **48** (2000) 2545.
9) H. Nishimori, *Theory of spin glasses and information; An introduction* (Oxford, London, 2001).
10) Y. Saika, and H. Nishimori, Prog. Theor. Phys. Suppl.,**157** (2005) 292.
11) Y. Saika, and T. Uezu, Interdisciplinary Information Science, **19** (2013) 73.
12) G. Kitagawa and W. Gersch, *Smoothness Priors Analysis of Time Series* (Springer, 1996).
13) Y. Nakashima, Y. Igarashi, Y. Naruse, and M. Okada. J. Phys. Soc. Jpn. 87, 084801 (2018).
14) R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1989).
15) M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, 2009).
16) C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).