Symmetry, Integrability and Geometry: Methods and Applications SIGMA **3** (2007), 018, 7 pages

Lattice Field Theory with the Sign Problem and the Maximum Entropy Method*

Masahiro IMACHI †, *Yasuhiko SHINNO* ‡ *and Hiroshi YONEYAMA* §

† *Kashiidai, Higashi-ku, Fukuoka, 813-0014, Japan*

‡ *Takamatsu National College of Technology, Takamatsu 761-8058, Japan*

§ *Department of Physics, Saga University, Saga, 840-8502, Japan*
E-mail: *yoneyama@cc.saga-u.ac.jp*

Received September 30, 2006, in final form January 19, 2007; Published online February 05, 2007
Original article is available at `http://www.emis.de/journals/SIGMA/2007/018/`

Abstract. Although numerical simulation in lattice field theory is one of the most effective tools to study non-perturbative properties of field theories, it faces serious obstacles coming from the sign problem in some theories such as finite density QCD and lattice field theory with the θ term. We reconsider this problem from the point of view of the maximum entropy method.

Key words: lattice field theory; sign problem; maximum entropy method

2000 Mathematics Subject Classification: 65C05; 65C50; 68U20; 81T25; 81T80

1 Introduction

Lattice field theory is a powerful method to study non-perturbative aspects of quantum field theory. Although numerical simulation is one of the most effective tools to study non-perturbative properties of field theories, it faces serious obstacles in theories such as finite density QCD and lattice field theory with the θ term. This is because the Boltzmann weights are complex and this makes it difficult to perform Monte Carlo (MC) simulations on a Euclidean lattice. This is the complex action problem, or the sign problem. In the present talk, we review the analysis of the sign problem based on the maximum entropy method (MEM) [1, 2, 3]. For details, refer to [4, 5, 6]. The MEM is well known as a powerful tool for so-called ill-posed problems, where the number of parameters to be determined is much larger than the number of data points. It has been applied to a wide range of fields, such as radio astrophysics and condensed matter physics.

In this talk we deal only with lattice field theory with the θ term. It is believed that the θ term could affect the dynamics at low energy and the vacuum structure of QCD, but it is known from experimental evidence that the value of θ is strongly suppressed in Nature. From the theoretical point of view, the reason for this is not clear yet. Hence, it is important to study the properties of QCD with the θ term to clarify the structure of the QCD vacuum [7, 8, 9]. For theories with the θ term, it has been pointed out that rich phase structures could be realized in θ space. For example, the phase structure of the $Z(N)$ gauge model was investigated using free energy arguments, and it was found that oblique confinement phases could occur [8, 9]. In CP^{N-1} models, which have several dynamical properties in common with QCD, it has been shown that a first-order phase transition exists at $\theta = \pi$ [10, 11, 12].

In order to circumvent the sign problem, the following method is conventionally employed [11, 12]. The partition function $\mathcal{Z}(\theta)$ can be obtained by Fourier-transforming the topological charge

*This paper is a contribution to the Proceedings of the O'Raifeartaigh Symposium on Non-Perturbative and Symmetry Methods in Field Theory (June 22–24, 2006, Budapest, Hungary). The full collection is available at http://www.emis.de/journals/SIGMA/LOR2006.html distribution $P(Q)$, which is calculated with a real positive Boltzmann weight:

$$\mathcal{Z}(\theta) = \frac{\int [d\bar{z}dz] e^{-S+i\theta\hat{Q}(\bar{z},z)}}{\int [d\bar{z}dz] e^{-S}} \equiv \sum_Q e^{i\theta Q} P(Q),$$

where

$$P(Q) \equiv \frac{\int [d\bar{z}dz]_Q e^{-S}}{\int d\bar{z}dz e^{-S}}. \tag{1}$$

The measure $[d\bar{z}dz]_Q$ in equation (1) is such that the integral is restricted to configurations of the field z with topological charge Q. Also, S represents the action.

In the study of CP^{N-1} models, it is known that this algorithm works well for a small lattice volume V and in the strong coupling region [10, 11, 13, 14]. As the volume increases or in the weak coupling region, however, this strategy too suffers from the sign problem for $\theta \simeq \pi$. The error in $P(Q)$ masks the true values of $\mathcal{Z}(\theta)$ in the vicinity of $\theta = \pi$, and this results in a fictitious signal of a phase transition [13, 14]. This is called 'flattening', because the free energy becomes almost flat for θ larger than a certain value. This problem could be remedied by reducing the error in $P(Q)$. This, however, is hopeless, because the amount of data needed to reduce the error to a given level increases exponentially with V.

Here, we are interested in whether the MEM can be applied effectively to the study of the θ term and reconsider the flattening phenomenon of the free energy in terms of the MEM. The MEM is based upon Bayes' theorem. It derives the most probable parameters by utilizing data sets and our knowledge about these parameters in terms of the probability. The probability distribution, which is called the posterior probability, is given by the product of the likelihood function and the prior probability. The latter is represented by the Shannon-Jaynes entropy, which plays an important role to guarantee the uniqueness of the solution, and the former is given by χ^2. It should be noted that artificial assumptions are not needed in the calculations, because the determination of a unique solution is carried out according to probability theory. Our task is to determine the image for which the posterior probability is maximized.

We present the results for the analysis by i) using mock data and ii) using the MC data. For the former, we use the Gaussian form of $P(Q)$. The Gaussian form is realized in many cases, such as the strong coupling region of the CP^{N-1} model and the 2-d $U(1)$ gauge model. For using MC data, we simulate the CP^{N-1} model and apply the MEM to the obtained data. This paper is organized as follows. In the following section, we give an overview of the origin of flattening and summarize the procedure for the analysis of the MEM. The results obtained by use of the MEM are presented in Section 3.

2 Flattening and MEM

2.1 Flattening

The free energy density $f(\theta)$ is calculated by Fourier-transforming $P(Q)$ obtained by MC simulation. Let us call this method the FTM. The quantity $f(\theta)$ is defined as

$$f(\theta) = -\frac{1}{V} \ln \sum_Q P(Q) e^{i\theta Q},$$

where $V = L^2$, the square of the lattice size.

The MC data for $P(Q)$ consist of the true value, $\tilde{P}(Q)$, and its error, $\Delta P(Q)$. When the error at $Q = 0$ dominates because of the exponential damping of $P(Q)$, $f(\theta)$ is closely approximated by

$$f(\theta) \simeq -\frac{1}{V} \ln \left[e^{-V\tilde{f}(\theta)} + \Delta P(0) \right],$$

where $\tilde{f}(\theta)$ is the true value of $f(\theta)$. Because $\tilde{f}(\theta)$ is an increasing function of θ, $\Delta P(0)$ dominates for large values of θ. If $|\Delta P(0)| \simeq e^{-V\tilde{f}(\theta)}$ at $\theta = \theta_{\rm f}$, then $f(\theta)$ becomes almost flat for $\theta \gtrsim \theta_{\rm f}$. This is called "flattening of the free energy density", and it has been misleadingly identified as a first order phase transition, because the first derivative of $f(\theta)$ appears to jump at $\theta = \theta_{\rm f}$. To avoid this problem, we must carry out the order of e^V measurements in the FTM.

2.2 MEM

In this subsection, we briefly explain the MEM in terms of the θ term. In a parameter inference, such as the χ^2 fitting, the inverse Fourier transform

$$P(Q) = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \mathcal{Z}(\theta) e^{-i\theta Q} \tag{2}$$

is used. In the numerical calculations, we use the discretized version of equation (2); $P(Q) = \sum_n K_{Q,n} \mathcal{Z}_n$, where $K_{Q,n}$ is the Fourier integral kernel and $\mathcal{Z}_n \equiv \mathcal{Z}(\theta_n)$. In order for the continuous function $\mathcal{Z}(\theta)$ to be reconstructed, a sufficient number of values of θ, which we denote by N_θ, is required so that the relation $N_\theta > N_Q$ holds, where N_Q represents the number of data points in $P(Q)$ ($Q = 0, 1, \ldots, N_Q - 1$). A straightforward application of the χ^2 fitting to the case $N_\theta > N_Q$ leads to degenerate solutions. This is an ill-posed problem.

The maximum entropy method is suitable to solve this ill-posed problem, yielding a unique solution. The MEM is based upon Bayes' theorem, expressed as

$$\mathrm{prob}(\mathcal{Z}(\theta)|P(Q), I) = \frac{\mathrm{prob}(P(Q)|\mathcal{Z}(\theta), I)\,\mathrm{prob}(\mathcal{Z}(\theta)|I)}{\mathrm{prob}(P(Q)|I)},$$

where $\mathrm{prob}(A|B)$ is the conditional probability that A occurs under the condition that B occurs. The posterior probability $\mathrm{prob}(\mathcal{Z}(\theta)|P(Q), I)$ is the probability that the partition function $\mathcal{Z}(\theta)$ is realized when the MC data $\{P(Q)\}$ and prior information I are given. The likelihood function $\mathrm{prob}(P(Q)|\mathcal{Z}(\theta), I)$ is given by

$$\mathrm{prob}(P(Q)|\mathcal{Z}(\theta), I) = \frac{1}{X_L} e^{-\frac{1}{2}\chi^2},$$

where X_L is a normalization constant and χ^2 is a standard χ^2 function.

The probability $\mathrm{prob}(\mathcal{Z}(\theta)|I)$ is given in terms of an entropy S as

$$\mathrm{prob}(\mathcal{Z}(\theta)|I) = \frac{1}{X_s(\alpha)} e^{-\alpha S},$$

where α and $X_S(\alpha)$ are a positive parameter and an α-dependent normalization constant, respectively. As S, the Shannon–Jaynes entropy is conventionally employed:

$$S = \sum_{n=1}^{N_\theta} \left[\mathcal{Z}_n - m_n - \mathcal{Z}_n \ln \frac{\mathcal{Z}_n}{m_n} \right].$$

Here $m_n \equiv m(\theta_n)$ represents a default model.

The posterior probability $\mathrm{prob}(\mathcal{Z}_n|P(Q), I)$, thus, is given by

$$\mathrm{prob}(\mathcal{Z}_n|P(Q), I, \alpha, m) = \frac{1}{X_L X_s(\alpha)} e^{-\frac{1}{2}\chi^2 + \alpha S} \equiv \frac{e^{W[\mathcal{Z}]}}{X_L X_s(\alpha)}.$$

For the prior information I, we impose the criterion

$$\mathcal{Z}_n > 0, \qquad \text{so that} \quad \mathrm{prob}(\mathcal{Z}_n \le 0\,|\,I, \alpha, m) = 0.$$

The most probable image of $\mathcal{Z}_n$, denoted as $\hat{\mathcal{Z}}_n$, is calculated according to the following procedures [3, 4].

1. Maximizing $W[\mathcal{Z}]$ to obtain the most probable image $\mathcal{Z}_n^{(\alpha)}$ for a given α:

$$\frac{\delta}{\delta \mathcal{Z}_n} W\left[\mathcal{Z}\right]\Big|_{\mathcal{Z}=\mathcal{Z}^{(\alpha)}} = \frac{\delta}{\delta \mathcal{Z}_n}\left(-\frac{1}{2}\chi^2 + \alpha S\right)\Big|_{\mathcal{Z}=\mathcal{Z}^{(\alpha)}} = 0.$$

2. Averaging $\mathcal{Z}_n^{(\alpha)}$ to obtain the α-independent most probable image $\mathcal{Z}_n$:

$$\hat{\mathcal{Z}}_n = \int d\alpha \ \mathcal{Z}_n^{(\alpha)} \ \mathrm{prob}(\alpha|P(Q), I, m).$$

   The range of integration is determined so that the relation

$$\mathrm{prob}(\alpha|P(Q), I, m) \geq \mathrm{prob}(\hat{\alpha}|P(Q), I, m)/10$$

   holds, where $\mathrm{prob}(\alpha|P(Q), I, m)$ is maximized at $\alpha = \hat{\alpha}$.

3. Error estimation: The error of the most probable output image $\hat{\mathcal{Z}}_n$ is calculated as the uncertainty of the image, which takes into account the correlations of the images $\hat{\mathcal{Z}}_n$ among various values of θ_n:

$$\langle(\delta\hat{\mathcal{Z}}_n)^2\rangle \equiv \int d\alpha \ \langle(\delta\mathcal{Z}_n^{(\alpha)})^2\rangle \ \mathrm{prob}(\alpha|P(Q), I, m). \tag{3}$$

   Here $\delta\hat{\mathcal{Z}}_n$ and $\delta\mathcal{Z}_n^{(\alpha)}$ represent the error in $\hat{\mathcal{Z}}_n$ and that in $\mathcal{Z}_n^{(\alpha)}$, respectively.

3 Results

3.1 Mock data: Gaussian

Firstly, we present the results of the analysis by using mock data. For this, we use the Gaussian $P(Q)$ as

$$P(Q) = A \exp\left[-\frac{c}{V}Q^2\right],$$

where, in the case of the 2-d $U(1)$ gauge model, c is a constant depending on the inverse coupling constant β, and V is the lattice volume. The constant A is fixed so that $\sum_Q P(Q) = 1$. The distribution $P(Q)$ is analytically transformed by use of the Poisson sum formula into the partition function

$$\mathcal{Z}_{\mathrm{pois}}(\theta) = A\sqrt{\frac{\pi V}{c}} \sum_{n=-\infty}^{\infty} \exp\left[-\frac{V}{4c}(\theta - 2\pi n)^2\right]. \tag{4}$$

To prepare the mock data, we add noise with variance $\delta \times P(Q)$ to the Gaussian $P(Q)$. In the analysis, we consider sets of data with various values of δ and study the effects of δ.

In Fig. 1 the Gaussian topological charge distribution and corresponding $f(\theta)$ obtained by using the FTM are shown for various lattice volumes. For small volumes, the behavior of $f(\theta)$ is smooth. For large volume ($V = 50$), however, clear flattening is observed. For $V = 30$, some data are missing. This is due to the fact that $\mathcal{Z}(\theta)$ could take negative values because of large errors. This is also called flattening, because its origin is the same as that stated above.

For the $P(Q)$ data corresponding to small volumes without flattening, the MEM successfully reproduces $f(\theta)$. Fig. 2 displays $\mathcal{Z}(\theta)$ for $V = 50$ by using the MEM. Here, the Gaussian default $m(\theta) = \exp\left[-\gamma\frac{\ln 10}{\pi^2}\theta^2\right]$ with $\gamma = 5.5$ is used. The result of the MEM does not show flattening but smooth behavior, in agreement with the exact result calculated by equation (4) (Poisson). In contrast, the FTM yields flattening. See [4] for details.

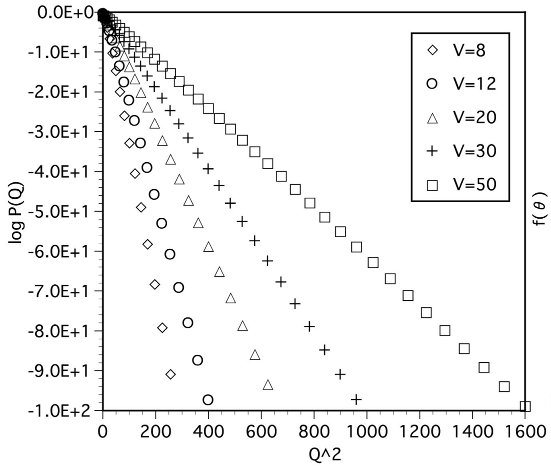

Figure 1. Gaussian topological charge distribution and corresponding $f(\theta)$ obtained by using the Fourier method (FTM) for various lattice volumes [4]. The parameter δ is chosen to be 1/400.

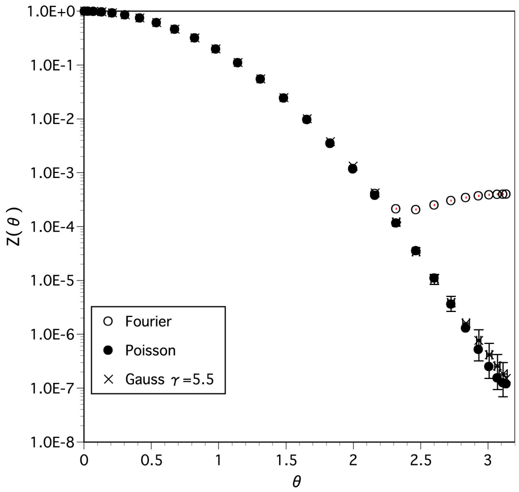

Figure 2. $\hat{\mathcal{Z}}(\theta)$ (crosses) with the error bars for the Gaussian default model with $\gamma = 5.5$ [4]. Here, $V = 50$. Compared to the result of the FTM (circles), a remarkable improvement is clearly seen.

3.2 MC data

In this subsection, we apply the MEM to real Monte Carlo data by simulating the CP^3 model. (For details, see [6].) For this we used a fixed point action [15, 16] and various lattice volumes $L \times L$. Among these, we concentrate on the data for $L = 38$ as the non-flattening case and $L = 50$ as the flattening case. We systematically studied the flattening phenomenon by adopting a variety of default model $m(\theta)$ and prior probability of the parameter α. For the latter, $g(\alpha)$ dependence appearing in

$$\mathrm{prob}(\alpha|P(Q), I, m) \equiv P(\alpha) \propto g(\alpha) e^{W(\alpha)+\Lambda(\alpha)},$$

is investigated. The function $g(\alpha)$ represents the prior probability of α and is chosen according to prior information. In general, two types of $g(\alpha)$ are employed, one according to Laplace's rule, $g_{\mathrm{Lap}}(\alpha) = \mathrm{const}$, and one according to Jeffrey's rule, $g_{\mathrm{Jef}}(\alpha) = 1/\alpha$. The latter rule is determined by requiring that $P(\alpha)$ be invariant with respect to a change in scale, because α is a scale factor. The former rule means that we have no knowledge about the prior information of α. In general, the most probable image $\hat{\mathcal{Z}}(\theta)$ depends on $g(\alpha)$. We investigate the sensitivity

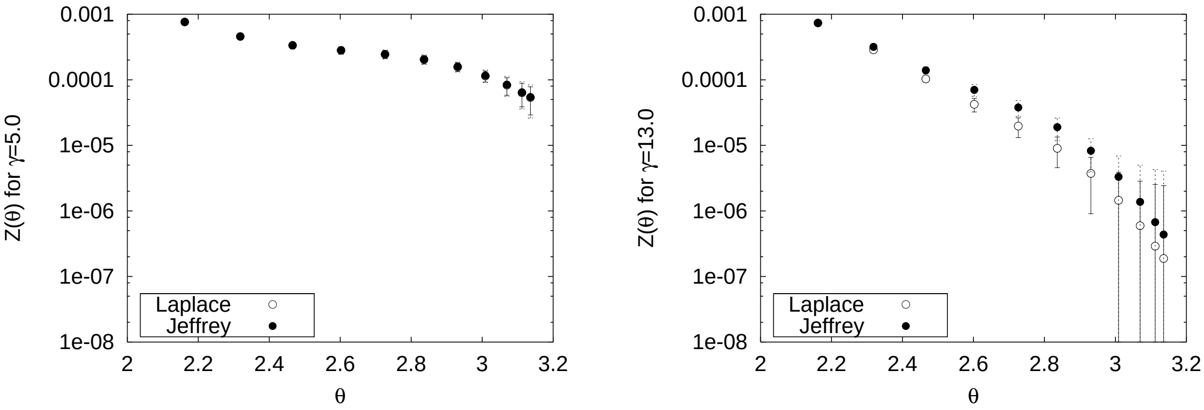

Figure 3. $\hat{\mathcal{Z}}_{\rm Lap}(\theta)$ and $\hat{\mathcal{Z}}_{\rm Jef}(\theta)$ for $\theta \in [2.0, \pi]$ [6].

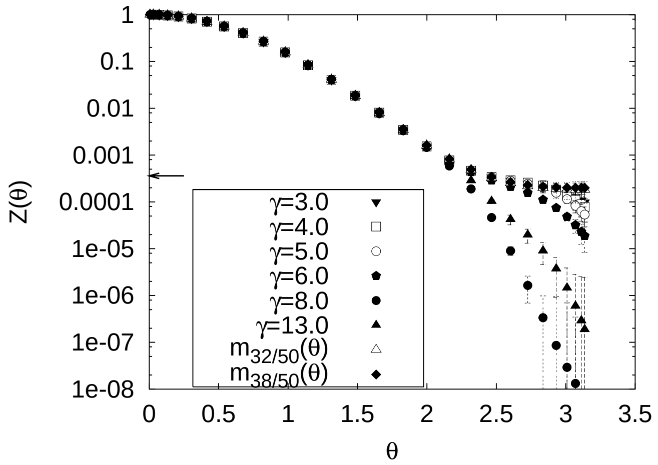

Figure 4. The most probable images $\hat{\mathcal{Z}}(\theta)$ for various $m(\theta)$ [6].

of $\hat{\mathcal{Z}}(\theta)$ to the choice of $g(\alpha)$ by studying a relative difference

$$\Delta(\theta) \equiv \frac{|\hat{\mathcal{Z}}_{\rm Lap}(\theta) - \hat{\mathcal{Z}}_{\rm Jef}(\theta)|}{(\hat{\mathcal{Z}}_{\rm Lap}(\theta) + \hat{\mathcal{Z}}_{\rm Jef}(\theta))/2}, \tag{5}$$

where $\hat{\mathcal{Z}}_{\rm Lap}(\theta)$ and $\hat{\mathcal{Z}}_{\rm Jef}(\theta)$ represent the most probable images according to Laplace's rule and Jeffrey's rule, respectively.

In the case without flattening ($L = 38$), the MEM yielded images $\hat{\mathcal{Z}}(\theta)$ that are almost independent of $m(\theta)$ and $g(\alpha)$. The most probable images $\hat{\mathcal{Z}}(\theta)$ are in agreement with the result of the FTM within the errors.

In the case with flattening ($L = 50$), we found that the statistical fluctuations of $\hat{\mathcal{Z}}(\theta)$ become smaller as the number of measurements increases except near $\theta = \pi$. We also found that $\hat{\mathcal{Z}}(\theta)$ with large errors depends strongly on $g(\alpha)$ in the region of large θ, where the $g(\alpha)$ dependence of $\hat{\mathcal{Z}}(\theta)$ was estimated using the quantity $\Delta(\theta)$. For $\theta \lesssim 2.3$, $\hat{\mathcal{Z}}(\theta)$ agrees with the result of the FTM. For $\theta \gtrsim 2.3$, $\hat{\mathcal{Z}}(\theta)$ behaves smoothly, while the FTM develops flattening. In Fig. 3 , we compare $\hat{\mathcal{Z}}(\theta)$ in larger values of θ for $g_{\rm Lap}(\alpha)$ and $g_{\rm Jef}(\alpha)$ when two different Gaussian default models are used ($\gamma = 5.0$ and 13.0). It is noted that $\gamma = 5.0$ is the case in which the smallest values of $\Delta(\theta)$ in equation (5) are observed in this θ region among various default models.

Our results are summarized in Fig. 4. All the results obtained using the MEM behave smoothly over the entire range of θ. Errors are estimated from uncertainties of the images according to equation (3). For larger values of θ, $\hat{\mathcal{Z}}(\theta)$ depends strongly on $m(\theta)$. Each of these images could be a candidate for the true image. This $m(\theta)$ dependence of $\hat{\mathcal{Z}}(\theta)$ may reflect the flattening phenomenon. If we had proper knowledge about $m(\theta)$ as prior information, we could identify the true image in a probabilistic sense. Such knowledge may also allow us to clarify the relationship between the default model dependence and the systematic error, which is not included in the figure. This will be a task to be pursued at the next stage.

The MEM provides a probabilistic point of view in the study of theories with the sign problem. It may then be worthwhile to study lattice QCD with a finite density in terms of the MEM.

Acknowledgements

One of the authors (H.Y.) is grateful to the organizers of the O'Raifeartaigh Symposium on Non-Perturbative and Symmetry Methods in Field Theory for providing him an opportunity to present a talk. He also thanks them for their warm hospitality. The symposium reminded him well that he had enjoyed intensive discussions with Lochlainn and other colleagues (A. Wipf, especially) at DIAS.

References

[1] Bryan R.K., Maximum entropy analysis of oversampled data problems, *Eur. Biophys. J.* **18** (1990), 165–174.

[2] Jarrell M., Gubernatis J.E., Bayesian inference and the analytic continuation of imaginary-time quantum Monte Carlo data, *Phys. Rep.* **269** (1996), 133–195.

[3] Asakawa M., Hatsuda T., Nakahara Y., Maximum entropy analysis of the spectral functions in lattice QCD, *Prog. Part. Nuclear Phys.* **46** (2001), 459–508, hep-lat/0011040.

[4] Imachi M., Shinno Y., Yoneyama H., Maximum entropy method approach to θ term, *Progr. Theoret. Phys.* **111** (2004), 387–411, hep-lat/0309156.

[5] Imachi M., Shinno Y., Yoneyama H., True or fictitious flattening?: MEM and the θ term, hep-lat/0506032.

[6] Imachi M., Shinno Y., Yoneyama H., Sign problem and MEM in lattice field theory with the θ term, *Progr. Theoret. Phys.* **115** (2006), 931–949, hep-lat/0602009.

[7] 't Hooft G., Topology of the gauge condition and new confinement phases in non-Abelian gauge theories, *Nuclear Phys. B* **190** (1981), 455–478.

[8] Cardy J.L., Rabinovici E., Phase structure of $\mathrm{Z_p}$ models in the presence of a θ parameter, *Nuclear Phys. B* **205** (1982), 1–16.

[9] Cardy J. L., Duality and the parameter in Abelian lattice models, *Nuclear Phys. B* **205** (1982), 17–26.

[10] Seiberg N., Topology in strong coupling, *Phys. Rev. Lett.* **53** (1984), 637–640.

[11] Bhanot G., Rabinovici E., Seiberg N., Woit P., Lattice vacua, *Nuclear Phys. B* **230** (1984), 291–298.

[12] Wiese U.-J., Numerical simulation of lattice θ-vacua: the 2-d $U(1)$ gauge theory as a test case, *Nuclear Phys. B* **318** (1989), 153–175.

[13] Plefka J.C., Samuel S., Monte Carlo studies of two-dimensional systems with a θ term, *Phys. Rev. D* **56** (1997), 44–54, hep-lat/9704016.

[14] Imachi M., Kanou S., Yoneyama H., Two-dimensional CP^2 model with θ term and topological charge distributions, *Progr. Theoret. Phys.* **102** (1999), 653–670, hep-lat/9905035.

[15] Hasenfratz P., Niedermayer F., Perfect lattice action for asymptotically free theories, *Nuclear Phys. B* **414** (1994), 785–814,hep-lat/9308004.

[16] Burkhalter R., Imachi M., Shinno Y., Yoneyama H., CP^{N-1} models with θ term and fixed point action *Progr. Theoret. Phys.* **106** (2001), 613–640, hep-lat/0103016.