An information geometrical approach to the mean-field approximation for quantum Ising spin models

2008 J. Phys. A: Math. Theor. 41 065005

(http://iopscience.iop.org/1751-8121/41/6/065005)

J. Phys. A: Math. Theor. **41** (2008) 065005 (16pp) doi:10.1088/1751-8113/41/6/065005

An information geometrical approach to the mean-field approximation for quantum Ising spin models

Nihal Yapage and Hiroshi Nagaoka

Graduate School of Information Systems, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan

E-mail: nihal@hn.is.uec.ac.jp and nagaoka@is.uec.ac.jp

Received 16 September 2007
Published 29 January 2008
Online at stacks.iop.org/JPhysA/41/065005

Abstract
We study the mean-field approximation for a general class of quantum Ising spin states from an information geometrical point of view. The states we consider are assumed to have at most second-order interactions with arbitrary but deterministic coupling coefficients. We call such a state a quantum Boltzmann machine (QBM) for the reason that it can be regarded as a quantum extension of the equilibrium distribution of a (classical) Boltzmann machine (CBM), which is a well-known stochastic neural network model. The totality of QBMs is then shown to form a quantum exponential family and thus can be considered as a smooth manifold having similar geometrical structures to those of CBMs. We elaborate on the significance and usefulness of information geometrical concepts, in particular the e- and m-projections, in studying the naive mean-field approximation for QBMs. We also discuss the higher-order corrections to the naive mean-field approximation based on the idea of the Plefka expansion in statistical physics. We elucidate the geometrical essence of the corrections and provide the expansion coefficients with expressions in terms of information geometrical quantities. Here, one may note this work as the information geometrical interpretation of (Plefka T 2006 *Phys. Rev.* E **73** 016129) and as the quantum extension of (Tanaka T 2000 *Neural Comput.* **12** 1951–68).

PACS numbers: 02.40.−k, 03.67.−a, 05.30.−d, 02.30.Mv, 05.50.+q, 75.10.Jm

1. Introduction

In many areas of physics, information theory and other related areas, one often encounters the calculation of quantities such as expectations and correlations of a system with respect to a probability distribution or a density operator with a complicated structure. This is, in general, computationally a very time-consuming problem since the required time increases exponentially with the number of elements in the system irrespective of being classical or quantum. Thus, it is inevitable to employ an approximation method to get rid of this difficulty. On the other hand, information geometry as a new field has been found very useful in understanding the deep mathematical structures in the interface of many fields.

Mean-field approximation, originated in statistical physics, has been widely used both in classical and quantum physics as well as in other fields such as information theory, statistics, etc. The basic idea of the mean-field approximation is to use a simple tractable family of probability distributions (or density operators) to calculate characteristic quantities related to a probability distribution (or a density operator) including mutual interactions. In particular, Tanaka [4, 5] has studied the mean-field approximation for a general class of classical Ising spin models which are identified with the equilibrium distributions of stochastic neural networks called Boltzmann machines (or *classical Boltzmann machines* (CBMs) in this paper) from the viewpoint of information geometry; see also [6, 7] for related works.

Motivated by the above works, we study in this paper the mean-field approximation for a class of quantum Ising spin states, which correspond to the equilibrium distributions of CBMs and are called *quantum Boltzmann machines* (QBMs), from a viewpoint of quantum information geometry. Similar to CBMs, each QBM has two kinds of real-valued parameters, namely h, the thresholds in neural network contexts and the external fields in physical contexts, and w, the coupling coefficients (to be defined in section 2). We remark that these parameters are arbitrary but deterministic in our model, while the coupling coefficients are usually considered to be random variables in case of spin glasses. Here, it should be noted that the QBMs, at present, lack any notions corresponding to the stochastic dynamics of CBMs which determine their equilibrium distributions. This means that our approach does not suggest how to quantize the neural aspects of CBMs, and we simply introduce QBMs as a general class of quantum states.

We regard the set of QBMs as a quantum exponential family, a smooth manifold which has a similar form as an exponential family in statistics, on which a Riemannian metric and a couple of affine connections are naturally defined. These differential geometrical structures turn out to have a characteristic property called the dually flatness and are closely related to the quantum relative entropy. Using this setup, we elucidate the geometrical essence of the naive mean-field approximation for the QBMs as well as the higher-order extensions based on the idea of the Plefka expansion.

The structure of this paper is as follows. In the next section, we define QBMs corresponding to the CBMs. Section 3 introduces classical and quantum exponential families. It is shown that the manifolds of QBMs and product states form quantum exponential families. In sections 4 and 5 we describe some relevant concepts from quantum information geometry. We devote section 6 to derive the naive mean-field equation for QBMs using information geometrical notions. Section 7 discusses the higher-order mean-field approximations for QBMs using a Taylor expansion of the quantum relative entropy. Furthermore, we elaborate on the unified view of the naive and higher-order mean-field approximations. Discussion and conclusions in section 8 terminate the paper.

2. Quantum spin states and QBMs

Let us consider an n-element system of quantum Ising spins. Each element is represented as a quantum bit (qubit) or quantum spin-$\frac{1}{2}$ with the local Hilbert space $\mathbb{C}^2$, and the

n-element system corresponds to $\mathcal{H} \equiv (\mathbb{C}^2)^{\otimes n}$. Let $\mathcal{S}$ be the set of faithful states on $\mathcal{H}$; $\mathcal{S} = \{\rho \mid \rho = \rho^* > 0 \text{ and } \mathrm{Tr}\,\rho = 1\}$. An element of $\mathcal{S}$ is said to have at most kth-order interactions if it is written as

$$\begin{aligned}\rho_\theta &= \exp\left\{\sum_{i,s}\theta^{(1)}_{is}X_{is} + \sum_{i<j}\sum_{s,t}\theta^{(2)}_{ijst}X_{is}X_{jt} + \cdots\right.\\ &\qquad\left. + \sum_{i_1<\cdots<i_k}\sum_{s_1\ldots s_k}\theta^{(k)}_{i_1\ldots i_k s_1\ldots s_k}X_{i_1s_1}\cdots X_{i_ks_k} - \psi(\theta)\right\}\\ &= \exp\left\{\sum_{j=1}^{k}\sum_{i_1<\cdots<i_j}\sum_{s_1\ldots s_j}\theta^{(j)}_{i_1\ldots i_j s_1\ldots s_j}X_{i_1s_1}\cdots X_{i_js_j} - \psi(\theta)\right\}\end{aligned} \tag{1}$$

with

$$\psi(\theta) = \log\mathrm{Tr}\exp\left\{\sum_{j=1}^{k}\sum_{i_1<\cdots<i_j}\sum_{s_1\ldots s_j}\theta^{(j)}_{i_1\ldots i_j s_1\ldots s_j}X_{i_1s_1}\cdots X_{i_js_j}\right\}, \tag{2}$$

where $X_{is} = I^{\otimes(i-1)}\otimes X_s\otimes I^{\otimes(n-i)}$, $\theta = \left(\theta^{(j)}_{i_1\ldots i_j s_1\ldots s_j}\right)$. Here, I is the identity matrix on $\mathcal{H}$ and X_s for $s\in\{1,2,3\}$ are the usual Pauli matrices given by

$$X_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad X_2 = \begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}, \qquad X_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Letting $\mathcal{S}_k$ be the totality of states ρ_θ of the above form, we have the hierarchy $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_n = \mathcal{S}$. Note that $\mathcal{S}_1$ is the set of product states $\rho_1\otimes\rho_2\otimes\cdots\otimes\rho_n$.

Our main concern in the present paper lies in approximating an element of $\mathcal{S}_2$ by a product state in $\mathcal{S}_1$, so the set $\mathcal{S}_2$ is of special importance. In the sequel, we let $h_{is} = \theta^{(1)}_{is}$ and $w_{ijst} = \theta^{(2)}_{ijst}$ to rewrite (1) for $k = 2$ as

$$\rho_{h,w} = \exp\left\{\sum_{i,s}h_{is}X_{is} + \sum_{i<j}\sum_{s,t}w_{ijst}X_{is}X_{jt} - \psi(h,w)\right\}, \tag{3}$$

where $h = (h_{is})$ and $w = (w_{ijst})$. The real dimension of $\mathcal{S}_2$ is $3n(3n-1)/2$ which gives the number of parameters to specify a density operator $\rho_{h,w}$.

A classical counterpart of the state (3) is the probability distribution for binary sequences $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \{-1,+1\}^n$ of the form

$$p_{h,w}(\boldsymbol{x}) = \exp\left\{\sum_i h_i x_i + \sum_{i<j} w_{ij}x_i x_j - \psi(h,w)\right\}, \tag{4}$$

where $h = (h_i)$ and $w = (w_{ij})$. This is known to appear as the equilibrium distribution of a stochastic neural network called the Boltzmann machine [1], which is referred to as a CBM in this paper with C standing for 'classical'. A CBM consists of a number, say n, of elements that are connected in some way and is specified by a set of parameters $(h,w) = (h_i, w_{ij})$. Here, $h_i \in \mathbb{R}$ denotes the threshold value of element i and $w_{ij}\in\mathbb{R}$ denotes the coupling coefficient between the two elements i and j. Each element i takes a binary value $x_i \in \{-1,+1\}$ as its state, which fluctuates according to a stochastic rule depending on both the parameters (h,w) and the values of the other elements. This process defines a Markov chain on the product set $\{-1,+1\}^n$, whose equilibrium (stationary) distribution is given by (4). Noting that the correspondence $p_{h,w} \leftrightarrow (h,w)$ is one to one, we can, at least mathematically, identify each

CBM with its equilibrium probability distribution. Actually, Tanaka [4, 5] has studied the mean-field approximation for a distribution of the form (4), calling it a Boltzmann machine. Correspondingly, an element of $\mathcal{S}_2$ of the form (3) is called a *quantum Boltzmann machine* or a QBM in this paper (see also [2, 3]), although we have no quantum dynamics corresponding to the stochastic state change of a CBM at present. Physically, a QBM simply means a general quantum state for n-fold spins with at most second-order interactions which are arbitrary and deterministic, not random as in a spin glass.

The elements of $\mathcal{S}_1$ are represented as $\rho_{h,0}$ by letting $w = 0$ in (3). In the sequel, we write them as

$$\tau_{\bar{h}} = \exp\left\{\sum_{i,s} \bar{h}_{is} X_{is} - \psi(\bar{h})\right\} \tag{5}$$

by using new symbols τ and $\bar{h} = (\bar{h}_{is})$ when we wish to make it clear that we are treating $\mathcal{S}_1$ instead of $\mathcal{S}_2$. We have

$$\tau_{\bar{h}} = \bigotimes_{i=1}^{n} \exp\left\{\sum_{s} \bar{h}_{is} X_s - \psi_i(\bar{h}_i)\right\}, \tag{6}$$

where $\bar{h}_i = (\bar{h}_{is})_s$ and

$$\begin{aligned}\psi_i(\bar{h}_i) &= \log \operatorname{Tr} \exp\left\{\sum_{s} \bar{h}_{is} X_s\right\} \\ &= \log\{\exp(\|\bar{h}_i\|) + \exp(-\|\bar{h}_i\|)\}\end{aligned} \tag{7}$$

with $\|\bar{h}_i\| \overset{\text{def}}{=} \sqrt{\sum_s (\bar{h}_{is})^2}$. Note that

$$\psi(\bar{h}) = \sum_{i} \psi_i(\bar{h}_i). \tag{8}$$

3. Quantum exponential families

Let $\mathcal{X}$ be a finite set or, more generally, a measurable space with an underlying measure $\mathrm{d}\mu$. When a family of probability distributions on $\mathcal{X}$ (probability mass functions for a finite $\mathcal{X}$ and probability density functions for a general $(\mathcal{X}, \mathrm{d}\mu)$), say $\mathcal{M} = \{p_\theta \mid \theta = (\theta^i); i = 1, \ldots, m\}$, is represented in the form

$$p_\theta(x) = \exp\left\{c(x) + \sum_{i=1}^{m} \theta^i f_i(x) - \psi(\theta)\right\}, \qquad x \in \mathcal{X}, \tag{9}$$

$\mathcal{M}$ is called an *exponential family*. Here, $\theta^i; i = 1, \ldots, m$ are real-valued parameters, c and f_i are the functions on $\mathcal{X}$ and $\psi(\theta)$ is a real-valued convex function. For instance, the equilibrium distributions (4) of CBMs form an exponential family. The notion of exponential family is very important in statistics and information geometry, and is also useful in studying properties of CBMs with their mean-field approximations. We introduce a quantum version of this notion in the following.

Let $\mathcal{H}$ be a finite dimensional Hilbert space and denote the totality of faithful states on $\mathcal{H}$ by

$$\mathcal{S} = \{\rho \mid \rho = \rho^* > 0 \text{ and } \operatorname{Tr}\rho = 1\}.$$

Suppose that a parametric family $\mathcal{M} = \{\rho_\theta \mid \theta = (\theta^i); i = 1, \ldots, m\} \subset \mathcal{S}$ is represented in the form

$$\rho_\theta = \exp\left\{C + \sum_{i=1}^{m} \theta^i F_i - \psi(\theta)\right\}, \tag{10}$$

where F_i $(i = 1, \ldots, m)$ and C are the Hermitian operators and $\psi(\theta)$ is a real-valued function. We assume in addition that the operators $\{F_1, \ldots, F_m, I\}$, where I is the identity operator, are linearly independent to ensure that the parametrization $\theta \mapsto \rho_\theta$ is one to one. Then $\mathcal{M}$ forms an m-dimensional smooth manifold with a coordinate system $\theta = (\theta^i)$. In this paper, we call such an $\mathcal{M}$ a *quantum exponential family*[1], or QEF for short, with *natural* (or *canonical*) *coordinates* $\theta = (\theta^i)$. It is easy to see that $\mathcal{S}$ is a QEF of dimension $(\dim \mathcal{H})^2 - 1$. Note also that for any $1 \leqslant k \leqslant n$ the set $\mathcal{S}_k$ of states (1) forms a QEF, including $\mathcal{S}_2$ of QBMs and $\mathcal{S}_1$ of product states.

If we let

$$\eta_i(\theta) \stackrel{\text{def}}{=} \mathrm{Tr}[\rho_\theta F_i], \tag{11}$$

then $\eta = (\eta_i)$ and $\theta = (\theta^i)$ are in one-to-one correspondence. That is, we can also use η instead of θ to specify an element of $\mathcal{M}$. These (η_i) are called the *expectation coordinates* of $\mathcal{M}$.

In particular, the natural coordinates of $\mathcal{S}_2$ are given by $(h, w) = (h_{is}, w_{ijst})$ in (3), while the expectation coordinates are $(m, \mu) = (m_{is}, \mu_{ijst})$ defined by

$$m_{is} = \mathrm{Tr}[\rho_{h,w} X_{is}] \qquad \text{and} \qquad \mu_{ijst} = \mathrm{Tr}[\rho_{h,w} X_{is} X_{jt}]. \tag{12}$$

On the other hand, the natural coordinates of $\mathcal{S}_1$ are $\bar{h} = (\bar{h}_{is})$ in (5), while the expectation coordinates are $\bar{m} = (\bar{m}_{is})$ defined by

$$\bar{m}_{is} = \mathrm{Tr}[\tau_{\bar{h}} X_{is}]. \tag{13}$$

In this case, the correspondence between the two coordinate systems can explicitly be represented as

$$\bar{m}_{is} = \frac{\partial \psi_i(\bar{h}_i)}{\partial \bar{h}_{is}} = \frac{\bar{h}_{is}}{\|\bar{h}_i\|} \tanh(\|\bar{h}_i\|) \tag{14}$$

or as

$$\bar{h}_{is} = \frac{\bar{m}_{is}}{\|\bar{m}_i\|} \tanh^{-1}(\|\bar{m}_i\|), \tag{15}$$

where $\|\bar{m}_i\| \stackrel{\text{def}}{=} \sqrt{\sum_s (\bar{m}_{is})^2}$.

4. Metric and affine connections on a state manifold

In classical information geometry [8], a Riemannian metric, called the Fisher metric, and a one-parameter family of affine connections, called the α-connections ($\alpha \in \mathbb{R}$), are canonically defined on an arbitrary manifold of probability distributions. In particular, the $(\alpha = 1)$-connection and the $(\alpha = -1)$-connection, which are also called the e-connection and m-connection, respectively, together with the Fisher metric have been shown very useful

[1] It should be noted, however, (10) is merely one of the possible definitions of quantum exponential family. Our definition has the advantage that it is closely related to the quantum relative entropy (30) and is completely analogous to the classical exponential family from a purely geometrical point of view. On the other hand, it does not fit well to the framework of quantum estimation theory, which needs another definition of QEF such as that based on symmetric logarithmic derivatives (see section 7.4 of [8]).

in many problems in statistics and other fields. In the quantum case, on the other hand, we have infinitely many mathematical equivalents of the Fisher metric and the α-connections including the e- and m-connections defined on a manifold of quantum states. We introduce, in the present section, an example of quantum Fisher metric and e-, m-connections, and describe their properties, mainly following [8]. Our choice of information geometrical structure is naturally linked with QEF and quantum relative entropy. For general terms of differential geometry such as manifold, Riemannian metric and affine connection, refer, for example, to [9].

Let $\mathcal{H}$ be a finite dimensional Hilbert space. We consider a d-dimensional parametric family

$$\mathcal{M} = \{\rho_\theta \mid \theta = (\theta^1, \dots, \theta^d) \in \Theta\}, \qquad \Theta \subset \mathbb{R}^d$$

of faithful states on $\mathcal{H}$. Then, $\theta = (\theta^i)$; $i = 1, \dots, d$ can be considered as a coordinate system and $\mathcal{M}$ becomes a submanifold of the manifold of faithful states $\mathcal{S}$ on $\mathcal{H}$. In the following, we discuss the information geometrical structure of $\mathcal{M}$ including the case $\mathcal{M} = \mathcal{S}$. As a first step, a Riemannian metric $g = [g_{ij}]$ is defined on $\mathcal{M}$ by

$$\begin{aligned} g_{ij}(\theta) &= g(\partial_i, \partial_j), \qquad \text{where} \quad \partial_i \stackrel{\text{def}}{=} \frac{\partial}{\partial \theta^i} \\ &= \int_0^1 \mathrm{Tr}\left[\rho_\theta^\lambda (\partial_i \log \rho_\theta) \rho_\theta^{1-\lambda} (\partial_j \log \rho_\theta)\right] \mathrm{d}\lambda \\ &= \mathrm{Tr}[(\partial_i \rho_\theta)(\partial_j \log \rho_\theta)]. \end{aligned} \tag{16}$$

This is a quantum version of the Fisher information metric and is called the *BKM (Bogoliubov–Kubo–Mori) metric*. Next, two torsion-free affine connections, the *exponential connection* (or *e-connection* for short) $\nabla^{(\mathrm{e})}$ and the *mixture connection* (or *m-connection* for short) $\nabla^{(\mathrm{m})}$ are defined on $\mathcal{M}$ as follows:

$$\Gamma^{(\mathrm{e})}_{ij,k}(\theta) \stackrel{\text{def}}{=} g\big(\nabla^{(\mathrm{e})}_{\partial_i} \partial_j, \partial_k\big) = \mathrm{Tr}[(\partial_i \partial_j \log \rho_\theta)(\partial_k \rho_\theta)] \tag{17}$$

and

$$\Gamma^{(\mathrm{m})}_{ij,k}(\theta) \stackrel{\text{def}}{=} g\big(\nabla^{(\mathrm{m})}_{\partial_i} \partial_j, \partial_k\big) = \mathrm{Tr}[(\partial_i \partial_j \rho_\theta)(\partial_k \log \rho_\theta)], \tag{18}$$

where g is the BKM metric. Note that both $\nabla^{(\mathrm{e})}$ and $\nabla^{(\mathrm{m})}$ are mappings (covariant derivatives) which map two vector fields X, Y to $\nabla^{(\mathrm{e})}_X Y$ and to $\nabla^{(\mathrm{m})}_X Y$, respectively. The coefficients $\Gamma^{(\mathrm{e})}_{ij,k}$ give a coordinate representation of the connection $\nabla^{(\mathrm{e})}$ relative to the metric g, while $\Gamma^{(\mathrm{e})k}_{ij}$ defined by $\nabla^{(\mathrm{e})}_{\partial_i} \partial_j = \sum_k \Gamma^{(\mathrm{e})k}_{ij} \partial_k$ purely represents $\nabla^{(\mathrm{e})}$. They are related to each other by $\Gamma^{(\mathrm{e})}_{ij,k} = \sum_l \Gamma^{(\mathrm{e})l}_{ij} g_{kl}$. Similarly, we have $\Gamma^{(\mathrm{m})k}_{ij}$ for $\nabla^{(\mathrm{m})}$ such that $\nabla^{(\mathrm{m})}_{\partial_i} \partial_j = \sum_k \Gamma^{(\mathrm{m})k}_{ij} \partial_k$ and $\Gamma^{(\mathrm{m})}_{ij,k} = \sum_l \Gamma^{(\mathrm{m})l}_{ij} g_{kl}$.

These two connections $\nabla^{(\mathrm{e})}$ and $\nabla^{(\mathrm{m})}$ are dual with respect to the BKM metric (16) in the sense that, for any vector fields X, Y, Z,

$$Xg(Y, Z) = g\big(\nabla^{(\mathrm{e})}_X Y, Z\big) + g\big(Y, \nabla^{(\mathrm{m})}_X Z\big), \tag{19}$$

or equivalently in the component form

$$\partial_i g_{jk} = \Gamma^{(\mathrm{e})}_{ij,k} + \Gamma^{(\mathrm{m})}_{ik,j}. \tag{20}$$

This kind of duality for affine connections plays a key role in the classical and quantum information geometry. Another notable relation between the two connections is

$$\Gamma^{(\mathrm{m})}_{ij,k} - \Gamma^{(\mathrm{e})}_{ij,k} = T_{ijk}, \tag{21}$$

where

$$T_{ijk}(\theta) \stackrel{\text{def}}{=} 2\,\mathrm{Re}\iint_{0\leqslant\nu\leqslant\lambda\leqslant 1} \mathrm{Tr}\left[\rho_\theta^{\nu}(\partial_i \log\rho_\theta)\rho_\theta^{\lambda-\nu}(\partial_j \log\rho_\theta)\rho_\theta^{1-\lambda}(\partial_k \log\rho_\theta)\right] \mathrm{d}\nu\,\mathrm{d}\lambda. \qquad (22)$$

Let us now consider the case when $\mathcal{M}$ is a QEF (10) with natural coordinates $\theta = (\theta^i)$. It is then easy to check from (17) that the coefficients $\Gamma^{(\mathrm{e})}_{ij,k}$ or $\Gamma^{(\mathrm{e})k}_{ij}$ of the e-connection are all zero. In the context of differential geometry, this means that $\mathcal{M}$ is flat with respect to the connection $\nabla^{(\mathrm{e})}$ (e-flat, for short) and $\theta = (\theta^i)$ forms an affine coordinate system for $\nabla^{(\mathrm{e})}$ (e-affine coordinate system, for short). On the other hand, the coefficients $\Gamma^{(\mathrm{m})}_{ij,k}$ of the m-connection do not vanish with respect to the natural coordinates $\theta = (\theta^i)$. However, one of the remarkable consequences of the duality (19) is that, if one of the two connections $\nabla^{(\mathrm{e})}$ and $\nabla^{(\mathrm{m})}$ is flat, then the other is also flat, which is referred to as the *dually flatness* of the manifold with respect to the information geometrical structure $(g, \nabla^{(\mathrm{e})}, \nabla^{(\mathrm{m})})$. In the present case, the connection coefficients of $\nabla^{(\mathrm{m})}$ with respect to the expectation coordinates $\eta = (\eta_i)$ defined by (11) turn out to identically vanish. This means that $\mathcal{M}$ is m-flat with an m-affine coordinate system (η_i). Moreover, we have

$$g\left(\frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\eta_j}\right) = \delta_i^j \qquad (=1 \text{ if } i = j, 0 \text{ otherwise}) \qquad (23)$$

and

$$\eta_i = \frac{\partial\psi}{\partial\theta^i}, \qquad \theta^i = \frac{\partial\phi}{\partial\eta_i}, \qquad (24)$$

where ψ given in (10), is regarded as a function $\mathcal{M} \to \mathbb{R}$ by $\psi(\rho_\theta) = \psi(\theta)$, and $\phi : \mathcal{M} \to \mathbb{R}$ is defined by the relation

$$\phi(\rho) + \psi(\rho) = \sum_i \eta_i(\rho)\theta^i(\rho), \qquad \forall\rho \in \mathcal{M}. \qquad (25)$$

Note that equation (14) is an example of the first equation in (24). It can also be shown that

$$\phi(\rho) = -\mathrm{Tr}[\rho C] - S(\rho), \qquad (26)$$

where $S(\rho) \stackrel{\text{def}}{=} -\mathrm{Tr}[\rho\log\rho]$ is the von Neumann entropy. In particular, for the QEF $\mathcal{S}_k$ of states (1), we have $C = 0$ and hence $\phi(\rho) = -S(\rho)$. We note that the existence of m-affine coordinates $\eta = (\eta_i)$ and functions ψ, ϕ satisfying the relations (23), (24) and (25) is ensured as a general property of the dually flat space (see theorem 3.6 in [8]), although it is not difficult to directly verify these relations for a QEF (10).

Finally, let us rewrite (23) into a form which will be useful in later arguments. Noting that (23) is written as

$$g\left(\frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\eta_j}\right) = \frac{\partial\eta_i}{\partial\eta_j} \qquad (27)$$

and that $\left\{\left(\frac{\partial}{\partial\eta_j}\right)_\rho\right\}$ form a basis of the tangent space $T_\rho(\mathcal{M})$, we have

$$g\left(\left(\frac{\partial}{\partial\theta^i}\right)_\rho, \partial'\right) = \partial'\eta_i \qquad \forall\,\partial' \in T_\rho(\mathcal{M}). \qquad (28)$$

Similarly, we have

$$g\left(\left(\frac{\partial}{\partial\eta_i}\right)_\rho, \partial'\right) = \partial'\theta^i \qquad \forall\,\partial' \in T_\rho(\mathcal{M}), \qquad (29)$$

although we use only (28) in this paper.

5. Geometry of quantum relative entropy

In this section, we focus on the quantum relative entropy

$$D(\rho\|\sigma) \stackrel{\text{def}}{=} \mathrm{Tr}[\rho(\log\rho - \log\sigma)] \tag{30}$$

for two density operators $\rho, \sigma \in \mathcal{M}$ and describe its properties related to the dually flat structure $(g, \nabla^{(\mathrm{e})}, \nabla^{(\mathrm{m})})$ of a QEF $\mathcal{M}$ as the continuation of the previous section. First, we note that the relation

$$D(\rho\|\sigma) = \phi(\rho) + \psi(\sigma) - \sum_i \eta_i(\rho)\theta^i(\sigma) \tag{31}$$

holds for any $\rho, \sigma \in \mathcal{M}$ with ψ and ϕ defined in (24) and (25). From this, we have

$$D(\rho\|\sigma) + D(\sigma\|\tau) - D(\rho\|\tau) = \sum_i \{\eta_i(\rho) - \eta_i(\sigma)\}\{\theta^i(\tau) - \theta^i(\sigma)\}. \tag{32}$$

Moreover, it can be shown that (32), with the positivity

$$D(\rho\|\sigma) \geqslant 0, \qquad D(\rho\|\sigma) = 0 \qquad \text{iff} \quad \rho = \sigma, \tag{33}$$

completely characterizes the quantum relative entropy D.

Let us clarify the geometric meaning of the right-hand side of (32). In general, given a coordinate system ξ^i and an affine connection ∇ with coefficients Γ_{ij}^k, a geodesic with respect to ∇ is defined by the second-order ordinary differential equation $\ddot{\xi}^k + \sum_{i,j}\Gamma_{ij}^k\dot{\xi}^i\dot{\xi}^j = 0$. If, in addition, ξ^i is an affine coordinate system with respect to a flat ∇, the equation becomes $\ddot{\xi}^k = 0$ or equivalently $\xi_t^i = t\xi_0^i + (1-t)\xi_1^i$. In particular, an e-geodesic in the QEF (10) is given by

$$\theta_t^i = t\theta_0^i + (1-t)\theta_1^i. \tag{34}$$

This turns out to be equivalent to

$$\log\rho_t = t\log\rho_0 + (1-t)\log\rho_1 - \psi(t),$$

where $\psi(t)$ is the normalization constant. In other words, an e-geodesic of a QEF is itself a one-dimensional QEF. On the other hand, an m-geodesic is represented as

$$\eta_{ti} = t\eta_{0i} + (1-t)\eta_{1i}. \tag{35}$$

If we consider the case $\mathcal{M} = \mathcal{S}$, the m-geodesic can be written as

$$\rho_t = t\rho_0 + (1-t)\rho_1. \tag{36}$$

Such a family of states $\{\rho_t\}$ is called a (one dimensional) mixture family, which is related to the origin of the name 'mixture connection', but note that (35) is not generally represented as (36) unless $\mathcal{M} = \mathcal{S}$.

Let $\gamma : [0,1] \to \mathcal{M}$ be an m-geodesic such that $\gamma(0) = \sigma$, $\gamma(1) = \rho$ and $\delta : [0,1] \to \mathcal{M}$ be an e-geodesic such that $\delta(0) = \sigma$, $\delta(1) = \tau$. Then, from (34) and (35) we obtain

$$\dot{\gamma}(0) = \sum_i \{\eta_i(\rho) - \eta_i(\sigma)\}\left(\frac{\partial}{\partial\eta_i}\right)_\sigma \in T_\sigma(\mathcal{M}) \tag{37}$$

and

$$\dot{\delta}(0) = \sum_i \{\theta^i(\tau) - \theta^i(\sigma)\}\left(\frac{\partial}{\partial\theta^i}\right)_\sigma \in T_\sigma(\mathcal{M}). \tag{38}$$

Hence, from (23) we have

$$g(\dot{\gamma}(0), \dot{\delta}(0)) = \sum_i \{\eta_i(\rho) - \eta_i(\sigma)\}\{\theta^i(\tau) - \theta^i(\sigma)\} \tag{39}$$

which coincides with the right-hand side of (32). We thus obtain the following theorem:

Theorem 1 (Pythagorean relation)**.** *Let ρ, σ and τ be three points in the manifold $\mathcal{M}$ such that the m-geodesic connecting ρ and σ is orthogonal at σ to the e-geodesic connecting σ and τ with respect to the BKM Riemannian metric. Then, the generalized Pythagorean relation*

$$D(\rho\|\sigma) + D(\sigma\|\tau) = D(\rho\|\tau) \tag{40}$$

holds.

Next, we define the m- and e-projections. Let $\mathcal{M}$ be a QEF of the form (10) and $\mathcal{N}$ be a smooth submanifold of $\mathcal{M}$. For an arbitrary point $\rho \in \mathcal{M}$, let $D(\rho\|\cdot)|_{\mathcal{N}}$ be a function on $\mathcal{N}$ defined by $\mathcal{N} \ni \sigma \mapsto D(\rho\|\sigma)$. When this function is stationary (i.e., the derivative is zero for every direction in $\mathcal{N}$) at a point $\sigma \in \mathcal{N}$, we say that σ is an *m-projection* of ρ onto $\mathcal{N}$. Similarly, when $D(\cdot\|\rho)|_{\mathcal{N}}$ is stationary at $\sigma \in \mathcal{N}$, we say that σ is an *e-projection* of ρ onto $\mathcal{N}$. Then, we have the following two theorems which are closely related to theorem 1.

Theorem 2. *The necessary and sufficient condition for σ to be an m-projection (resp. e-projection) of ρ onto $\mathcal{N}$ is that the m-geodesic (resp. e-geodesic) connecting ρ and σ is orthogonal to $\mathcal{N}$ at σ.*

Theorem 3. *If $\mathcal{N}$ is e-autoparallel (resp. m-autoparallel) in $\mathcal{M}$ in the sense that $\mathcal{N}$ forms an affine subspace in e-affine coordinates (θ^i) (resp. m-affine coordinates) of $\mathcal{M}$, then an m-projection (resp. e-projection) is unique and attains the minimum of $D(\rho\|\cdot)|_{\mathcal{N}}$ (resp. $D(\cdot\|\rho)|_{\mathcal{N}}$).*

Finally, we note another property of D for later use. We have the Taylor expansion of $D(\rho\|\sigma)$ (see [8], p 55) as

$$D(\rho\|\sigma) \stackrel{\text{def}}{=} \frac{1}{2}\sum_{ij} g_{ij}(\rho)\Delta\theta^i\Delta\theta^j + \frac{1}{6}\sum_{ijk} h_{ijk}(\rho)\Delta\theta^i\Delta\theta^j\Delta\theta^k + \cdots, \tag{41}$$

where $\Delta\theta^i \stackrel{\text{def}}{=} \theta^i(\sigma) - \theta^i(\rho)$. Here, the second-order coefficients g_{ij} are the components of the BKM metric and the third-order coefficients h_{ijk} are determined from g_{ij} and the connection coefficients by

$$h_{ijk} \stackrel{\text{def}}{=} \partial_i g_{jk} + \Gamma^{(\mathrm{e})}_{jk,i} = \Gamma^{(\mathrm{e})}_{ij,k} + \Gamma^{(\mathrm{m})}_{ik,j} + \Gamma^{(\mathrm{e})}_{jk,i}, \tag{42}$$

where the second equality is due to (20).

6. Naive mean-field approximation and e-, m-projections

Suppose that we are interested in calculating the expectations $m_{is} = \mathrm{Tr}[\rho_{h,w} X_{is}]$ from given $(h, w) = (h_{is}, w_{ijst})$. Since the direct calculation is intractable in general when the system size is large, we need to employ a computationally efficient approximation method. The mean-field approximation is a well-known technique for this purpose. The simple idea behind the mean-field approximation for a $\rho_{h,w} \in \mathcal{S}_2$ is to use quantities obtained in the form of expectation with respect to some relevant $\tau_{\bar{h}} \in \mathcal{S}_1$. Tanaka [4, 5] has elucidated the essence of the naive mean-field approximation for classical spin models in terms of e-, m-projections. Our aim is to extend this idea to quantized spin models.

In the following arguments, we regard $\mathcal{S}_2$ as a QEF with the natural coordinates $(\theta^\alpha) = (h_{is}, w_{ijst})$ and the expectation coordinates $(\eta_\alpha) = (m_{is}, \mu_{ijst})$ (see (12)), where α is an index denoting $\alpha = (i, s)$ or $\alpha = (i, j, s, t)$. First, let us consider the m-projection onto $\mathcal{S}_1$ and show that it preserves the expectations m_{is}. Note that since $\mathcal{S}_1$ is e-autoparallel in $\mathcal{S}_2$, the m-projection is unique and attains the minimum of D by theorem 3. Suppose $\rho = \rho_{h,w} \in \mathcal{S}_2$ is given and $\tau \in \mathcal{S}_1$ be its m-projection. Let γ be the m-geodesic such that $\gamma(1) = \rho, \gamma(0) = \tau$. Then, from (35) we have

$$\eta_\alpha(\gamma(t)) = t\eta_\alpha(\rho) + (1-t)\eta_\alpha(\tau), \tag{43}$$

and, in particular,

$$m_{is}(\gamma(t)) = tm_{is}(\rho) + (1-t)m_{is}(\tau). \tag{44}$$

Hence, substituting $\theta^i := h_{is}, \eta_i := m_{is}, \partial' := \dot{\gamma}(0)$ and $\rho := \tau$ into (28) we get

$$g\left(\left(\frac{\partial}{\partial h_{is}}\right)_\tau, \dot{\gamma}(0)\right) = \left.\frac{\mathrm{d}m_{is}(\gamma(t))}{\mathrm{d}t}\right|_{t=0} = m_{is}(\rho) - m_{is}(\tau). \tag{45}$$

Since $T_\tau(\mathcal{S}_1) = \mathrm{span}\left\{\left(\frac{\partial}{\partial h_{is}}\right)_\tau\right\}$, it follows from (45) and theorem 2 that $m_{is}(\rho) = m_{is}(\tau)$. This means that the expectation values do not change if we use the m-projection.

Next, we show that the naive mean-field equation is derived by considering e-projection. Suppose that $\tau = \tau_{\bar{h}} \in \mathcal{S}_1$ is an e-projection of $\rho = \rho_{h,w} \in \mathcal{S}_2$ onto $\mathcal{S}_1$, and let γ be the e-geodesic such that $\gamma(1) = \rho, \gamma(0) = \tau$. Note that from (34)

$$\left.\frac{\mathrm{d}\theta^\alpha(\gamma(t))}{\mathrm{d}t}\right|_{t=0} = \theta^\alpha(\rho) - \theta^\alpha(\tau), \tag{46}$$

i.e.,

$$\left.\frac{\mathrm{d}h_{is}(\gamma(t))}{\mathrm{d}t}\right|_{t=0} = h_{is}(\rho) - h_{is}(\tau) = h_{is} - \bar{h}_{is} \tag{47}$$

$$\left.\frac{\mathrm{d}w_{ijst}(\gamma(t))}{\mathrm{d}t}\right|_{t=0} = w_{ijst}(\rho) = w_{ijst} \qquad \text{since} \quad w_{ijst}(\tau) = 0. \tag{48}$$

Now, recall that $(\bar{m}_{is})$ defined by (13) form a coordinate system of $\mathcal{S}_1$, so that $T_\tau(\mathcal{S}_1) = \mathrm{span}\left\{\left(\frac{\partial}{\partial \bar{m}_{is}}\right)_\tau\right\}_{is}$. Hence it follows from theorem 2 that, $\forall\, i, s$,

$$\begin{aligned}
0 &= g\left(\dot{\gamma}(0), \left(\frac{\partial}{\partial \bar{m}_{is}}\right)_\tau\right) \\
&= \sum_\alpha \left.\frac{\mathrm{d}\theta^\alpha(\gamma(t))}{\mathrm{d}t}\right|_{t=0} g\left(\left(\frac{\partial}{\partial \theta^\alpha}\right)_\tau, \left(\frac{\partial}{\partial \bar{m}_{is}}\right)_\tau\right) \\
&= \sum_\alpha \left.\frac{\mathrm{d}\theta^\alpha(\gamma(t))}{\mathrm{d}t}\right|_{t=0} \left(\frac{\partial \eta_\alpha}{\partial \bar{m}_{is}}\right)_\tau \qquad \text{(from (28))} \\
&= \sum_{j,t}(h_{jt} - \bar{h}_{jt})\frac{\partial m_{jt}}{\partial \bar{m}_{is}} + \sum_{j<k}\sum_{t,u} w_{jktu}\frac{\partial \mu_{jktu}}{\partial \bar{m}_{is}} \\
&= h_{is} - \bar{h}_{is} + \sum_{(i<)k,u} w_{iksu}\bar{m}_{ku} + \sum_{(i>)j,t} w_{jits}\bar{m}_{jt} \\
&\quad (\text{since } m_{jt} = \bar{m}_{jt} \text{ and } \mu_{jktu} = \bar{m}_{jt}\bar{m}_{ku} \text{ on } \mathcal{S}_1) \\
&= h_{is} - \bar{h}_{is} + \sum_{j,t} w_{ijst}\bar{m}_{jt},
\end{aligned}$$

where the last equality follows by letting $w_{iist} \stackrel{\mathrm{def}}{=} 0$ and $w_{ijst} \stackrel{\mathrm{def}}{=} w_{jits}$ $(i > j)$. We thus obtain

$$\bar{h}_{is} = h_{is} + \sum_{j,t} w_{ijst}\bar{m}_{jt}. \tag{49}$$

Both (14) (or (15)) and (49) together give the naive mean-field equation for QBMs. It should be remarked that this naive mean-field equation may have several solutions $\{\bar{h}_{is}\}$ for a given set of $\{h_{is}, w_{ijst}\}$ as in the classical case, which corresponds to the fact that e-projection onto an e-autoparallel submanifold is not unique in general.

In this section, we have shown some properties of m- and e-projections based on the geometrical characterization given in theorem 2, but note that the same properties can also be derived in several different ways; for instance, we can use the relations (24), (25) and (31) for ψ and ϕ to derive them.

7. Plefka expansion and the higher-order mean-field approximations

Although the naive mean-field approximation is used extensively as a common tool to compute characteristic quantities of multi-particle systems, it is necessary to consider higher-order mean-field approximations to improve the accuracy in some situations. In this section, we discuss a method to derive higher-order mean-field approximations which utilizes a Taylor expansion of the quantum relative entropy. This coincides with the so-called *Plefka expansion* of the Gibbs potential as pointed out at the end of the section. We elucidate the correspondence of the coefficients of the Taylor expansion to the information geometrical quantities such as the metric and the e-, m-connections.

We start the discussion by recalling that the elements of $\mathcal{S}_2$ are parametrized as $\rho_{h,w}$ by $h = (h_{is})$ and $w = (w_{ijst})$. This means that $(\theta^\alpha) = (h, w)$ forms a coordinate system of the manifold $\mathcal{S}_2$. In viewing $\mathcal{S}_2$ as a QEF, (h, w) is a natural coordinate system, while the corresponding expectation coordinate system is given by $(\eta_\alpha) = (m, \mu)$ with $m = (m_{is})$ and $\mu = (\mu_{ijst})$. Let us now define a third coordinate system $(\xi^\alpha) \stackrel{\text{def}}{=} (m, w)$. The elements of $\mathcal{S}_2$ are then parametrized by (m, w), which we denote by $\hat{\rho}_{m,w}$ to avoid confusion with $\rho_{h,w}$. Note that

$$\mathcal{S}_2 = \{\rho_{h,w} \mid (h, w) : \text{free}\} = \{\hat{\rho}_{m,w} \mid (m, w) : \text{free}\} \tag{50}$$

and that

$$\hat{\rho}_{m,w} = \rho_{h,w} \iff \forall\, i, \forall\, s, m_{is} = \mathrm{Tr}[\rho_{h,w} X_{is}]. \tag{51}$$

For an arbitrarily fixed w, a submanifold of $\mathcal{S}_2$ is defined by

$$\mathcal{F}(w) \stackrel{\text{def}}{=} \{\rho_{h,w} \mid h : \text{free}\} = \{\hat{\rho}_{m,w} \mid m : \text{free}\}. \tag{52}$$

As a special case we have

$$\mathcal{F}(0) = \{\rho_{h,0} \mid h : \text{free}\} = \{\hat{\rho}_{m,0} \mid m : \text{free}\} = \mathcal{S}_1$$

which is the manifold of product states. We see that the family $\{\mathcal{F}(w)\}_w$ forms a foliation of $\mathcal{S}_2$ as

$$\mathcal{S}_2 = \bigcup_w \mathcal{F}(w). \tag{53}$$

Similarly, for an arbitrarily fixed m we define

$$\begin{aligned}\mathcal{A}(m) &\stackrel{\text{def}}{=} \{\hat{\rho}_{m,w} \mid w : \text{free}\}\\ &= \{\rho \in \mathcal{S}_2 \mid \forall\, i, \forall\, s, m_{is} = \mathrm{Tr}[\rho X_{is}]\},\end{aligned} \tag{54}$$

which yields another foliation of $\mathcal{S}_2$ as

$$\mathcal{S}_2 = \bigcup_m \mathcal{A}(m). \tag{55}$$

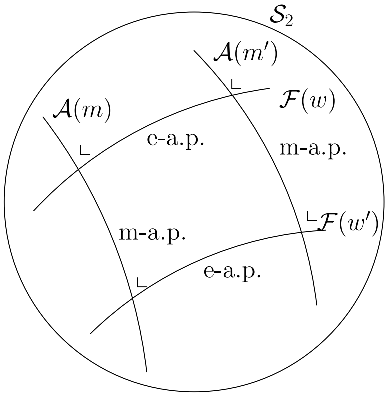

Figure 1. Mutually dual foliations of $\mathcal{S}_2$. Here, 'e-a.p.' and 'm-a.p.' stand for 'e-autoparallel' and 'm-autoparallel', respectively.

These foliations have several special properties. First, for any w, $\mathcal{F}(w)$ is defined by fixing w which is a part of e-affine coordinates $(\theta^\alpha) = (h, w)$ of $\mathcal{S}_2$. This implies that $\mathcal{F}(w)$ is e-autoparallel in $\mathcal{S}_2$ in the sense mentioned in theorem 3. On the other hand, each $\mathcal{A}(m)$ is m-autoparallel in $\mathcal{S}_2$. Furthermore, $\forall\, w, \forall\, m$,

$$\mathcal{F}(w) \perp \mathcal{A}(m) \quad \text{at} \quad \sigma \in \mathcal{F}(w) \cap \mathcal{A}(m). \tag{56}$$

To see this, we note that the tangent spaces of $\mathcal{F}(w)$ and $\mathcal{A}(m)$ at σ are given by

$$T_\sigma(\mathcal{F}(w)) = \text{span}\left\{\left(\frac{\partial}{\partial h_{is}}\right)_\sigma\right\}_{is} \tag{57}$$

and

$$T_\sigma(\mathcal{A}(m)) = \text{span}\left\{\left(\frac{\partial}{\partial \mu_{jktu}}\right)_\sigma\right\}_{jktu}. \tag{58}$$

The inner product $g\big(\frac{\partial}{\partial h_{is}}, \frac{\partial}{\partial \mu_{jktu}}\big)$ is a special case of $g\big(\frac{\partial}{\partial \theta^\alpha}, \frac{\partial}{\partial \eta_\beta}\big)$ with $\alpha \neq \beta$, hence is zero from (23), which proves (56). These properties mean that $\{\mathcal{F}(w)\}_w$ and $\{\mathcal{A}(m)\}_m$ jointly give an example of mutually dual foliations (see figure 1) defined in [8] (pp 75–76). It is now easy to see from theorem 1 that for any points $\rho \in \mathcal{A}(m)$ and $\tau \in \mathcal{F}(w)$ with the intersecting point $\sigma \in \mathcal{A}(m) \cap \mathcal{F}(w)$, the Pythagorean relation (40) holds. We also note that, for any w and m, both $\mathcal{F}(w)$ and $\mathcal{A}(m)$ are dually flat with respect to their e-, m-connections and the BKM metrics. This is obvious for $\mathcal{F}(w)$ because $\mathcal{F}(w)$ itself is a QEF. On the other hand, since $\mathcal{A}(m)$ is m-autoparallel in $\mathcal{S}_2$ which is m-flat, we can easily see that $\mathcal{A}(m)$ is also m-flat, and hence is dually flat as mentioned in section 4. Actually, (μ_{ijst}) and (w_{ijst}) restricted to $\mathcal{A}(m)$ turn out m-affine and e-affine coordinate systems respectively.

Let us now restate the problem which motivates both the naive mean-field approximation and its higher-order extension. Given $h = (h_{is})$ and $w = (w_{ijst})$ arbitrarily, consider the problem of calculating the expectations $\text{Tr}[\rho_{h,w} X_{is}]$ or their approximations from (h, w). From

$$\begin{aligned}
\forall\, i, \forall\, s, \frac{\partial}{\partial m_{is}} D(\hat{\rho}_{m,w} \| \rho_{h,w}) = 0 &\iff D(\hat{\rho}_{m,w} \| \rho_{h,w}) = \min_{m'} D(\hat{\rho}_{m',w} \| \rho_{h,w}) \\
&\iff \hat{\rho}_{m,w} = \operatorname*{argmin}_{\sigma \in \mathcal{F}(w)} D(\sigma \| \rho_{h,w}) = \rho_{h,w} \\
&\iff \forall\, i, \forall\, s, m_{is} = \text{Tr}[\rho_{h,w} X_{is}],
\end{aligned}$$

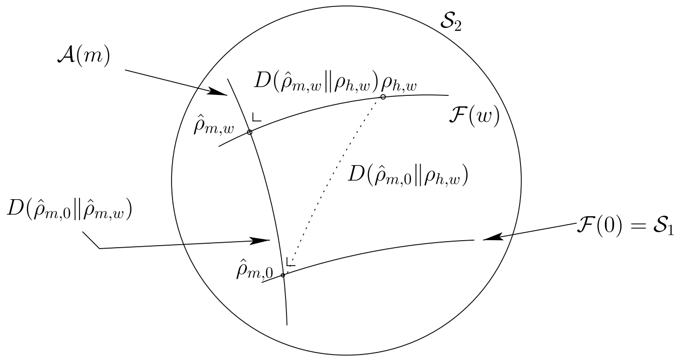

Figure 2. Pythagorean relation $D(\hat\rho_{m,0}\|\rho_{h,w}) = D(\hat\rho_{m,0}\|\hat\rho_{m,w}) + D(\hat\rho_{m,w}\|\rho_{h,w})$.

where the last equivalence follows from (51), we have the expectations m_{is} as the solution of the equation

$$\frac{\partial}{\partial m_{is}} D(\hat\rho_{m,w}\|\rho_{h,w}) = 0. \tag{59}$$

Of course, this method is practical only when the relative entropy $D(\hat\rho_{m,w}\|\rho_{h,w})$ is not too complicated as a function of the variables $m = (m_{is})$, which cannot be expected in general when n, the number of elements in the system, is large. On the other hand, if we let $w = 0$ in the first argument of $D(\hat\rho_{m,w}\|\rho_{h,w})$, then the resulting $D(\hat\rho_{m,0}\|\rho_{h,w})$ becomes the sum of simple functions of $m = (m_{is})$, and hence the equation

$$\frac{\partial}{\partial m_{is}} D(\hat\rho_{m,0}\|\rho_{h,w}) = 0 \tag{60}$$

is much more tractable than the original one in (59). When $\|w\|$ is sufficiently small so that $D(\hat\rho_{m,w}\|\rho_{h,w})$ is well approximated by $D(\hat\rho_{m,0}\|\rho_{h,w})$, the solution of (60) will give a good approximation for the true expectations. This is nothing but the idea of naive mean-field approximation. Actually, equation (60) means that $\hat\rho_{m,0}$ is an e-projection of $\rho_{h,w}$ onto $\mathcal{F}(0) = \mathcal{S}_1$, which turns out to be equivalent to (49) as shown in the previous section.

Now that the accuracy of the naive mean-field approximation depends on how close the function $D(\hat\rho_{m,w}\|\rho_{h,w})$ is to its substitute $D(\hat\rho_{m,0}\|\rho_{h,w})$, it is natural to expect that the approximation can be improved by properly retrieving the difference $D(\hat\rho_{m,0}\|\rho_{h,w}) - D(\hat\rho_{m,w}\|\rho_{h,w})$ up to a certain order of w. This is the information geometrical interpretation of the idea due to Plefka [10, 11], and we call the expansion of the difference with respect to w the *Plefka expansion* following Tanaka [5] who originally gave a similar interpretation in the classical case. From the information geometrical viewpoint, the gist of this approach is the fact that the Pythagorean relation (40) holds for the three points $\rho_{h,w}$, $\hat\rho_{m,w}$ and $\hat\rho_{m,0}$ (see figure 2) so that we have

$$D(\hat\rho_{m,0}\|\rho_{h,w}) - D(\hat\rho_{m,w}\|\rho_{h,w}) = D(\hat\rho_{m,0}\|\hat\rho_{m,w}). \tag{61}$$

The problem is thus reduced to the expansion of $D(\hat\rho_{m,0}\|\hat\rho_{m,w})$ with respect to w. Noting that $\hat\rho_{m,0}$ and $\hat\rho_{m,w}$ are the points on the manifold $\mathcal{A}(m)$ for which the coupling coefficients $w = (w_{ijst})$ form a coordinate system, the expansion formula (41) with (42) is applied to yield the Plefka expansion

$$D(\hat\rho_{m,0}\|\hat\rho_{m,w}) = \frac{1}{2}\sum_{IJ} g_{IJ} w_I w_J + \frac{1}{6}\sum_{IJK} h_{IJK} w_I w_J w_K + \cdots \tag{62}$$

with

$$h_{IJK} = \hat{\partial}_I g_{JK} + \Gamma^{(\mathrm{e})}_{JK,I} = \Gamma^{(\mathrm{e})}_{IJ,K} + \Gamma^{(\mathrm{m})}_{IK,J} + \Gamma^{(\mathrm{e})}_{JK,I}, \tag{63}$$

where the indices I, J, K represent quadruplets of indices such as (i, j, s, t). Here, g_{IJ}, $\Gamma^{(\mathrm{e})}_{IJ,K}$ and $\Gamma^{(\mathrm{m})}_{IJ,K}$ are respectively the components of the BKM metric, the e-connection and the m-connection of the manifold $\mathcal{A}(m)$, and $\hat{\partial}_I$ denotes $\frac{\partial}{\partial w_I}$ for the coordinates (w_I) of $\mathcal{A}(m)$, all evaluated at the point $\hat{\rho}_{m,0}$.

More specifically, it follows from (16) that

$$g_{IJ} = \int_0^1 \mathrm{Tr}\left[\hat{\rho}^{\lambda}(\hat{\partial}_I \log\hat{\rho})\hat{\rho}^{1-\lambda}(\hat{\partial}_J \log\hat{\rho})\right] \mathrm{d}\lambda, \tag{64}$$

where $\hat{\rho} = \hat{\rho}_{m,0}$ and $\hat{\partial}_I \log\hat{\rho} = \frac{\partial}{\partial w_I} \log\hat{\rho}_{m,w}\big|_{w=0}$. By some calculations (see appendix), we have

$$\hat{\partial}_I \log\hat{\rho} = (X_{is} - m_{is})(X_{jt} - m_{jt}) \tag{65}$$

for $I = (i, j, s, t)$. As for the third-order coefficients h_{IJK} in (63), we first note that $w = (w_I)$ is an e-affine coordinate system of $\mathcal{A}(m)$ as mentioned before and hence $\Gamma^{(\mathrm{e})}_{IJ,K} = 0$. As a consequence, we have

$$\begin{aligned} h_{IJK} &= \hat{\partial}_I g_{JK} = \Gamma^{(\mathrm{m})}_{IK,J} \\ &= 2\mathrm{Re}\iint_{0\leqslant\nu\leqslant\lambda\leqslant1} \mathrm{Tr}\left[\hat{\rho}^{\nu}(\hat{\partial}_I \log\hat{\rho})\hat{\rho}^{\lambda-\nu}(\hat{\partial}_J \log\hat{\rho})\hat{\rho}^{1-\lambda}(\hat{\partial}_K \log\hat{\rho})\right] \mathrm{d}\nu\, \mathrm{d}\lambda, \end{aligned} \tag{66}$$

where we have invoked (21) and (22).

If we succeed in obtaining explicit expressions for g_{IJ}, h_{IJK} and if they are not too complicated as functions of the mean variables $m = (m_{is})$, we can take

$$D(\hat{\rho}_{m,0}\|\rho_{h,w}) - \frac{1}{2}\sum_{IJ} g_{IJ} w_I w_J$$

or

$$D(\hat{\rho}_{m,0}\|\rho_{h,w}) - \frac{1}{2}\sum_{IJ} g_{IJ} w_I w_J - \frac{1}{6}\sum_{IJK} h_{IJK} w_I w_J w_K$$

as a substitute of $D(\hat{\rho}_{m,w}\|\rho_{h,w}) = D(\hat{\rho}_{m,0}\|\rho_{h,w}) - D(\hat{\rho}_{m,0}\|\hat{\rho}_{m,w})$ in equation (59) to improve the naive mean-field approximation (60). See [12, 13] for the classical case in this direction. In the quantum case, it is still a hard problem to carry out this program in spite of the remarkable progress made by [11].

Before closing this section, we verify the equivalence between our discussion and the original formulation of Plefka for expansion of the Gibbs potential. Let us define a function $\chi : \mathcal{S}_2 \to \mathbb{R}$ by

$$\chi(\rho) \overset{\text{def}}{=} \psi(\rho) - \sum_{i,s} m_{is}(\rho) h_{is}(\rho) \tag{67}$$

$$= S(\rho) + \sum_{i<j,s,t} \mu_{ijst}(\rho) w_{ijst}(\rho), \qquad \forall\, \rho \in \mathcal{S}_2 \tag{68}$$

where the second equality follows from (25) and $\phi(\rho) = -S(\rho)$. Noting that equation (24) yields

$$\mathrm{d}\psi = \sum_{\alpha} \eta_\alpha \mathrm{d}\theta^\alpha = \sum_{i,s} m_{is} \mathrm{d}h_{is} + \sum_{i<j,s,t} \mu_{ijst} \mathrm{d}w_{ijst}, \tag{69}$$

we obtain from equation (67) that

$$\begin{aligned} \mathrm{d}\chi &= \mathrm{d}\psi - \sum_{i,s} m_{is}\mathrm{d}h_{is} - \sum_{i,s} h_{is}\mathrm{d}m_{is} \\ &= \sum_{i<j,s,t} \mu_{ijst}\mathrm{d}w_{ijst} - \sum_{i,s} h_{is}\mathrm{d}m_{is}. \end{aligned} \tag{70}$$

This shows that it is natural to represent χ as a function of independent variables (m, w) by $\chi(\hat{\rho}_{m,w})$, which corresponds to what is called the Gibbs potential in [10, 11]. Now, it is immediate from (31) and (67) that, for any $(m, w) = (m_{is}, w_{ijst})$,

$$\chi(\hat{\rho}_{m,w}) = S(\hat{\rho}_{m,0}) + \sum_{i<j,s,t} m_{is}m_{jt}w_{ijst} + D(\hat{\rho}_{m,0}\|\hat{\rho}_{m,w}). \tag{71}$$

This implies that the expansions of $\chi(\hat{\rho}_{m,w})$ and $D(\hat{\rho}_{m,0}\|\hat{\rho}_{m,w})$ with respect to w are equivalent except for 0 th and first-order terms.

8. Discussion and conclusions

In this paper, we have derived the naive mean-field equations explicitly, followed by a comprehensive discussion of the higher-order approximations for QBMs from the viewpoint of information geometry. It has been shown that the fundamental concepts of information geometry such as the e-, m-connections and dualistic structure of exponential families play a major role. We have also established the correspondence of the information geometrical quantities such as the metric and e-, m-connections to the coefficients of the Plefka expansion and thus to higher-order approximations. Although our approach does not directly contribute to the explicit calculation of the coefficients, it will be interesting and important to investigate, for instance, the ingenious calculations of Plefka [11] for some higher-order terms in the light of information geometry.

Finally, we mention some possible extensions of the present quantum information geometrical formulation of the mean-field approximation beyond the QBMs. An immediate application is to employ this method for the mean-field approximation of a state (1) with $k \geqslant 3$. Another important extension is to consider the q-state quantum spin model in which each element has a local Hilbert space $\mathbb{C}^q$ and the whole system corresponds to $(\mathbb{C}^q)^{\otimes n}$. It would be useful to find other applications of this framework for quantum statistical models.

Appendix. Derivation of equation (65)

It follows from (70) that

$$\left(\frac{\partial\chi}{\partial w_{ijst}}\right)_{\xi} = \mu_{ijst} = m_{is}m_{jt} \qquad (\text{when } w = 0), \tag{A.1}$$

$$\left(\frac{\partial\chi}{\partial m_{is}}\right)_{\xi} = -h_{is}, \tag{A.2}$$

where $(\cdot)_{\xi}$ means that the partial differentiations are those with respect to the coordinate system $\xi = (m, w)$. Note that $\hat{\partial}_I$ in (65) for $I = (i, j, s, t)$ is $\left(\frac{\partial}{\partial w_{ijst}}\right)_{\xi}$ evaluated at $w = 0$. Now, from equations (3) and (67), we obtain

$$\begin{aligned}\log \hat{\rho}_{m,w} &= \sum_{ku} h_{ku} X_{ku} + \sum_{kluv} w_{kluv} X_{ku} X_{lv} - \psi \\ &= \sum_{ku} h_{ku}(X_{ku} - m_{ku}) + \sum_{kluv} w_{kluv} X_{ku} X_{lv} - \chi \end{aligned} \tag{A.3}$$

and

$$\hat{\partial}_I \log \hat{\rho}_{m,w} = \sum_{ku} \left(\frac{\partial h_{ku}}{\partial w_I}\right)_{\xi} (X_{ku} - m_{ku}) + X_{is} X_{jt} - \left(\frac{\partial \chi}{\partial w_I}\right)_{\xi}. \tag{A.4}$$

Here,

$$\begin{aligned}\left(\frac{\partial h_{ku}}{\partial w_I}\right)_{\xi} &= -\left(\frac{\partial^2 \chi}{\partial w_I \partial m_{ku}}\right)_{\xi} \qquad \text{(from equation (A.2))} \\ &= -\left(\frac{\partial}{\partial m_{ku}}\left(\frac{\partial \chi}{\partial w_I}\right)_{\xi}\right)_{\xi} \\ &= -\frac{\partial m_{is} m_{jt}}{\partial m_{ku}} \qquad \text{(from equation (A.1))} \\ &= -\delta_{ik}\delta_{su} m_{jt} - \delta_{jk}\delta_{tu} m_{is}.\end{aligned}$$

Substituting this and (A.1) in (A.4), we have

$$\begin{aligned}\hat{\partial}_I \log \hat{\rho}_{m,w} &= -(X_{is} - m_{is}) m_{jt} - (X_{jt} - m_{jt}) m_{is} + X_{is} X_{jt} - m_{is} m_{jt} \\ &= (X_{is} - m_{is})(X_{jt} - m_{jt}),\end{aligned}$$

which completes the derivation.

References

[1] Ackley D H, Hinton G E and Sejnowski T J 1985 *Cogn. Sci.* **9** 147–69

[2] Yapage N and Nagaoka H 2005 *Proc. ERATO Conf. on Quantum Information Science (EQIS'05) (Tokyo, Japan)* pp 204–5

[3] Yapage N and Nagaoka H 2006 *Proc. Asian Conf. on Quantum Information Science (AQIS'06) (Beijing, China)* pp 143–4

[4] Tanaka T 1996 *IEICE Trans. Fundam.* E79-A **5** 709–15

[5] Tanaka T 2000 *Neural Comput.* **12** 1951–68

[6] Bhattacharyya C and Keerthi S S 2000 *J. Phys. A: Math. Gen.* **33** 1307–12

[7] Amari S, Ikeda S and Shimokawa H 2001 *Advanced Mean Field Methods—Theory and Practice* ed M Opper and D Saad (Cambridge, MA: MIT Press) pp 241–57

[8] Amari S and Nagaoka H 2000 *Methods of Information Geometry* (Providence, RI/Oxford: American Mathematical Society/Oxford University Press)

[9] Kobayashi S and Nomizu K 1963 *Foundations of Differential Geometry* vol 1 (New York: Interscience)

[10] Plefka T 1982 *J. Phys. A: Math. Gen.* **15** 1971–8

[11] Plefka T 2006 *Phys. Rev.* E **73** 016129

[12] Thouless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593–601

[13] Nakanishi K and Takayama H 1997 *J. Phys. A: Math. Gen.* **30** 8085–94