数理解析研究所講究録
第 1692 巻 2010 年 112-132

Gradient systems on the quantum information space and engineering algorithms [1]

Hiromi Yuya
Graduate School of Systems Information Science
Future University Hakodate

Yoshio Uwano[2]
Department of Complex Systems
Future University Hakodate

Abstract

As a trial to associate algorithms with quantum mechanical postulates in a sense, the Karmarkar flow and an averaged learning equation of Hebb type are shown to be generalized to gradient systems on the quantum information space. The former is a continuous analogue of Karmarkar's affine scaling algorithm for linear programming and the latter is of a principal component analizer so that both systmems are sound in engineering.

1 Introduction

Algorithms are looked upon as discrete-time dynamical systems in a sense that they provide the evolution-rules of systems under consideration. Quite often, continuous analogue of algorithms are investigated to obtain deeper understanding of those algorithms. For example, the infinitesimal limit in the scaling parameter is applied to the celebrated Karmarkar projective scaling algorithm [1] to obtain the differential equation named the Karmarkar flow, which Karmarkar, the founder, has studied by himself [2]. As another example, Oja's rule of neuronal learning process is worth referred to, whose approximation in very small time-interval is introduced by Oja [3], the proposer of the rule. The differential equation arising from the approximation will be referred to as the averaged learning equation of Hebb type, which will be abbreviated to ALEH in this article.

In the middle of 1990's, Nakamura made a series of studies [4, 5, 6, 7] on differential equations relevant to engineering algorithms including the Karmarkar flow [2, 6] and the ALEH [3, 7] from integrability viewpoint: A Lax-type structure, gradient-system structure and Hamiltonian structure. More than a decade after Nakamura's works, an integrable gradient system on the statistical manifold of multinomial distributions was shown by Uwano [8, 9] to admit a very natural counterpart on the quantum information space (QIS). The gradient system associated with the negative von Neumann entropy on the

[1]Based on the talk by HY with the same title given in RIMS Workshop "*Geometric Mechanics*" (December 21-22, 2009).

[2]e-mail: uwano@fun.ac.jp

QIS was shown to be a generalization of the gradient system associated with the negative Shannon entropy on the statistical manifold of multinomial distributions.

The encounter of the works by Nakamura [6] and by Uwano [8, 9] gave a strong motivation to the authors to find other dynamical systems from Nakamura's list [4, 5, 6, 7] that allow natural counterpart on the QIS. The aim of the present article is to report that the Karmarkar flow and the ALEH can be realized as gradient systems on the QIS. The results on the Karmarkar flow in this article will be partially published soon in the paper [10]. The results given in this paper could be placed as a clue to algorithms with a quantum postulate since the dynamical systems dealt with are closely related to engineering algorithms and since the manifold, the QIS, to describe their generalization is highly quantum mechanical. Further, it could also be a clue to enlarge target area for seeking a novel quantum algorithms other than Shor's [11] and Grover's [12] both of which come from theory of computing. The contents of the present article are outlined in what follows.

Section 2 is for preliminaries to geometry and dynamics used in this article. The QIS is introduced as the space of $m \times m$ regular density matrices endowed with the quantum SLD (symmetric logarithmic derivative) Fisher metric. After the QIS, an explicit form of the gradient gradient equation is derived through a differential geometric calculus. Section 3 is one of the core parts, where the Karmarkar flow is realized as a gradient system on the QIS. A gradient system form of the Karmarkar flow is reviewed along with Nakamura's paper [6]. After the review, the gradient system realizing the Karmarkar flow on the QIS is derived by using preliminaries in Section 2. Section 4 is another of the core parts, where the averaged learning equation of Hebb type (ALEH) is realized as a gradient system on the QIS. The ALEH is introduced along the papers by Oja and by Nakamura [3, 7], whose gradient system form is given together. After the review, the ALEH is realized the gradient system on the QIS. In addition to the deriving method for the Karmarkar flow on the QIS, a symmetry of the ALHE has to be revealed to ensure the generalization. Section 5 is for concluding remarks.

2 Gradient systems on the QIS

2.1 The QIS

Following Uwano [8, 9], we introduce the quantum information space (QIS), which is realized as the space of the regular density matrices endowed with the quantum SLD (symmetric logarithmic derivative) Fisher metric in what follows.

Let us consider the space of $m \times m$ regular density matrices,

$$\dot{Q}_m = \{\mathrm{P} \in M(m,m) \,|\, \mathrm{P}^\dagger = \mathrm{P},\ \mathrm{tr}\,\mathrm{P} = 1,\ \mathrm{P} : \text{positive definite}\}, \tag{1}$$

where $M(m,m)$ is the set of $m \times m$ complex matrices. The tangent space of $\dot{Q}_m$ at $\mathrm{P} \in \dot{Q}_m$ is defined to be

$$T_\mathrm{P}\dot{Q}_m = \{\Xi \in M(m,m) \,|\, \Xi^\dagger = \Xi,\ \mathrm{tr}\,\Xi = 0\}. \tag{2}$$

By the symmetric logarithmic derivative (SLD), denoted by $\mathcal{L}_\mathrm{P}(\Xi)$, at $\mathrm{P} \in \dot{Q}_m$ subject to

$$\mathrm{P}\mathcal{L}_\mathrm{P}(\Xi) + \mathcal{L}_\mathrm{P}(\Xi)\,\mathrm{P} = 2\Xi \qquad (\Xi \in T_\mathrm{P}\dot{Q}_m) \tag{3}$$

the quantum SLD Fisher metric, denoted by $(\!(\cdot,\cdot)\!)^{QF}$, is defined to be

$$(\!(\Xi,\Xi')\!)^{QF}_{\mathrm{P}} = \frac{1}{2}\mathrm{tr}\ [\mathrm{P}\,(\mathcal{L}_{\mathrm{P}}(\Xi)\mathcal{L}_{\mathrm{P}}(\Xi') + \mathcal{L}_{\mathrm{P}}(\Xi')\mathcal{L}_{\mathrm{P}}(\Xi))] \qquad (\Xi,\Xi' \in T_{\mathrm{P}}\dot{Q}_m)\,, \tag{4}$$

(see [8, 9, 13]). Since $\mathrm{P} \in \dot{Q}_m$ admits the expression,

$$\begin{aligned} &\mathrm{P} = \Upsilon\Theta\Upsilon^{\dagger}, \quad \Upsilon \in \mathrm{U}(m), \\ &\Theta = \mathrm{diag}\,(\theta_1,\dots,\theta_m) \quad \text{with} \quad \mathrm{tr}\,\Theta = 1, \quad \theta_k > 0\ (k = 1,2,\cdots,m), \end{aligned} \tag{5}$$

where $U(m)$ denotes the group of $m \times m$ unitary matrices, the SLD Fisher metric, denoted by $(\!(\Xi,\Xi')\!)^{QF}_{\mathrm{P}}$, has an explicit expression,

$$(\!(\Xi,\Xi')\!)^{QF}_{\mathrm{P}} = 2\sum_{j,k=1}^{m} \frac{\overline{\mathrm{X}}_{jk}\mathrm{X}'_{jk}}{\theta_j + \theta_k}, \tag{6}$$

where $\Xi,\Xi' \in T_{\mathrm{P}}\dot{Q}_m$ are the Hermitean matrices subject to

$$\Xi = \Upsilon\mathrm{X}\Upsilon^{\dagger}, \Xi' = \Upsilon\mathrm{X}'\Upsilon^{\dagger}. \tag{7}$$

To summarize, we reach to the definition of the QIS.

Definition 2.1 *The Riemannian manifold $(\dot{Q}_m, (\!(\cdot,\cdot)\!)^{QF})$ is called the quantum information space, which is often abbreviated to as the QIS henceforth.*

2.2 Gradient equation

Let us derive the gradient equation on the QIS associated with an arbitrary potential function. By differential geometric convention, functions are assumed to be of C^{∞} class if no other conditions are mentioned of. The gradient vector field for any given potential function ϕ is determined not only on the potential ϕ but also on the Riemannian metric $(\!(\cdot,\cdot)\!)^{QF}$ [14]: The gradient vector field denoted by grad ϕ is defined by

$$(\!((\mathrm{grad}\ \phi)(\mathrm{P}),\Xi')\!)^{QF}_{\mathrm{P}} = (\mathrm{d}\phi)_{\mathrm{P}}(\Xi') = \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \phi(\Gamma'(t)) \quad ({}^{\forall}\Xi' \in T_{\mathrm{P}}\dot{Q}_m), \tag{8}$$

where $\Gamma'(t)$ with $a \le t \le b$ $(a < 0 < b)$ in (8) is a C^{∞} curve in $\dot{Q}_m$ subject to

$$\Gamma' : t \in [a,\, b] \mapsto \Gamma'(t) \in \dot{Q}_m, \quad \Gamma'(0) = \mathrm{P}, \quad \left.\frac{\mathrm{d}\Gamma'}{\mathrm{dt}}\right|_{t=0} = \Xi'. \tag{9}$$

To calculate (8), we introduce the partial differentiations,

$$\frac{\partial}{\partial \mathrm{P}_{jk}} = \frac{1}{2}\Big(\frac{\partial}{\partial \xi_{jk}} - \mathrm{i}\frac{\partial}{\partial \eta_{jk}}\Big), \quad \frac{\partial}{\partial \overline{\mathrm{P}}_{jk}} = \frac{1}{2}\Big(\frac{\partial}{\partial \xi_{jk}} + \mathrm{i}\frac{\partial}{\partial \eta_{jk}}\Big) \quad (1 \le j < k \le m), \tag{10}$$

where the i indicates the imaginary unit throughout the present paper. In terms of those differentiations, we prepare the matrix-valued operator $\mathcal{M}$ to ϕ to be

$$(\mathcal{M}(\phi))_{jk} = \begin{cases} \dfrac{\partial\phi}{\partial\overline{\mathrm{P}}_{jk}} = \overline{\dfrac{\partial\phi}{\partial\mathrm{P}_{jk}}} & (1 \le j < k \le m) \\[2ex] \dfrac{\partial\phi}{\partial\mathrm{P}_{kj}} & (1 \le k < j \le m) \\[2ex] \dfrac{\partial\phi}{\partial\mathrm{P}_{jj}} & (j = 1,2,\cdots,m). \end{cases} \tag{11}$$

The rhs of (8) is calculated with (9)-(11) to be

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \phi(\Gamma'(t)) \\
&= \sum_{1\le a<b\le m} \left\{ \frac{\partial \phi}{\partial \xi_{ab}}(\mathrm{P})\Re(\Xi'_{ab}) + \frac{\partial \phi}{\partial \psi_{ab}}(\mathrm{P})\Im(\Xi'_{ab}) \right\} + \sum_{c=1}^{m} \frac{\partial \phi}{\partial \zeta_c}(\mathrm{P})\,\Xi'_{cc} \\
&= \sum_{1\le a<b\le m} \left\{ \frac{\partial \phi}{\partial \xi_{ab}}(\mathrm{P})\frac{1}{2}\left(\Xi'_{ab} + \overline{\Xi'_{ab}}\right) + \frac{\partial \phi}{\partial \psi_{ab}}(\mathrm{P})\frac{1}{2\mathrm{i}}\left(\Xi'_{ab} - \overline{\Xi'_{ab}}\right) \right\} + \sum_{c=1}^{m} \frac{\partial \phi}{\partial \zeta_c}(\mathrm{P})\,\Xi'_{cc} \\
&= \sum_{1\le a<b\le m} \left\{ \frac{\partial \phi}{\partial \rho_{ab}}(\mathrm{P})\Xi'_{ab} + \frac{\partial \phi}{\partial \overline{\rho}_{ab}}(\mathrm{P})\overline{\Xi'_{ab}} \right\} + \sum_{c=1}^{m} \frac{\partial \phi}{\partial \zeta_c}(\mathrm{P})\,\Xi'_{cc} \\
&= \sum_{1\le a<b\le m} \{(\mathcal{M}(\phi))_{ba}\Xi'_{ab} + (\mathcal{M}(\phi))_{ab}\Xi'_{ba}\} + \sum_{c=1}^{m}(\mathcal{M}(\phi))_{cc}\Xi'_{cc} \\
&= \mathrm{tr}\ (\mathcal{M}(\phi)\Xi'),
\end{aligned} \tag{12}$$

where the symbols, $\Re$ and $\Im$, stand for the real part and the imaginary part of complex numbers, respectively.

We move on to calculate the lhs of (8). Let us express the gradient vector $(\mathrm{grad}\ \phi)(\mathrm{P})$ at $\mathrm{P} \in \dot{Q}_m$ to be

$$(\mathrm{grad}\ \phi)(\mathrm{P}) = \Upsilon\Psi\Upsilon^\dagger \tag{13}$$

where $\Upsilon \in U(m)$ is chosen in (5) to P. The Ψ is a traceless Hermitean matrix. On recalling the explicit expression (6) of the quantum Fisher metric, the lhs of (8) is then calculated to be

$$\begin{aligned}
\langle\!\langle (\mathrm{grad}\ \phi)(\mathrm{P}), \Xi' \rangle\!\rangle_{\mathrm{P}}^{QF} &= 2\sum_{j,k=1}^{m} \frac{\overline{\Psi}_{jk}X'_{jk}}{\theta_j + \theta_k} = 2\sum_{j,k=1}^{m} \frac{\Psi_{kj}}{\theta_j + \theta_k}\Big(\sum_{a,b=1}^{m}(\Upsilon^\dagger)_{ja}\Xi'_{ab}\Upsilon_{bk}\Big) \\
&= 2\sum_{a,b=1}^{m}\Big(\sum_{j,k=1}^{m}\Upsilon_{bk}\tilde{\Psi}_{kj}(\Upsilon^\dagger)_{ja}\Big)\Xi'_{ab} = 2\mathrm{tr}\ \Big((\Upsilon\tilde{\Psi}\Upsilon^\dagger)\Xi'\Big),
\end{aligned} \tag{14}$$

where $\tilde{\Psi}$ is the Hermitean matrix subject to

$$\tilde{\Psi}_{jk} = \frac{\Psi_{jk}}{\theta_j + \theta_k} \quad (j,k = 1,2,\cdots,m). \tag{15}$$

Since the bilinear form,

$$(\Xi, \Xi') \in T_{\mathrm{P}}\dot{Q}_m \times T_{\mathrm{P}}\dot{Q}_m \mapsto \mathrm{tr}\ (\Xi^\dagger\Xi') \in \mathbf{C}, \tag{16}$$

is well-known to be non-degenerate, we have the equation

$$\mathcal{M}(\phi) = 2\Upsilon\tilde{\Psi}\Upsilon^\dagger + 2\nu I \tag{17}$$

from (12) and (14), where ν is a constant determined soon below. From (13) and (17), The entries of Ψ turns out to take the form

$$\begin{aligned}
\Psi_{jk} &= \frac{1}{2}(\theta_j + \theta_k)(\Upsilon^\dagger\mathcal{M}(\phi)\Upsilon - 2\nu I)_{jk} \\
&= \frac{1}{2}\left(\Theta\Upsilon^\dagger\mathcal{M}(\phi)\Upsilon + \Upsilon^\dagger\mathcal{M}(\phi)\Upsilon\Theta - 4\nu\Theta\right)_{jk} \quad (j,k = 1,2,\cdots),
\end{aligned} \tag{18}$$

where δ_{jk} denotes the Kronecker delta. Through (5) and (13), we get

$$(\text{grad } \phi)(\text{P}) = \frac{1}{2}\Big(\text{P}\mathcal{M}(\phi) + \mathcal{M}(\phi)\text{P}\Big) - 2\nu\text{P}. \tag{19}$$

Since tangent vectors in $T_{\text{P}}\dot{Q}_m$ have the traceless property, we can determine the value of ν. On taking the trace in both sides of (19), we have the value ofν,

$$\nu = \frac{1}{2}\text{tr}\,\Big(\text{P}\mathcal{M}(\phi)\Big). \tag{20}$$

Therefore, the gradient equation associated with the potential function ϕ takes the form

$$\frac{\text{dP}}{\text{dt}} = -(\text{grad } \phi)(\text{P}) = -\frac{1}{2}\Big(\text{P}\mathcal{M}(\phi) + \mathcal{M}(\phi)\text{P}\Big) + \Big(\text{tr}\,\big(\text{P}\mathcal{M}(\phi)\big)\Big)\text{P}. \tag{21}$$

To summarize, we have the following lemma.

Lemma 2.2 *[10, 15, 16] A gradient system on the QIS $(\dot{Q}_m, ((\cdot,\cdot))^{QF})$ associated with a potential function ϕ is governed by (21), where $\mathcal{M}(\phi)$ is the Hermitean matrix defined by (11).*

Remark As a related work on the gradient equation on the QIS, the paper [17] by Braunstein would be worth being cited. In [17], the gradient equation on the QIS is discussed in tensorial local-coordinate framework, in contrast with our global description. As an intuitive example, the gradient equation leading a quantum inference problem is given, which is looked upon as a special case of our ALEH equation on the QIS discussed in the succeeding section.

3 The Karmarkar flow realized on the QIS

3.1 The Karmarkar flow

Karmarkar's projective scaling algorithm is defined for the canonical linear programming problem of *unconstrained case*,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & e^T x = 1, \quad x_j \geq 0 \ \ (j = 1, 2, \cdots, m) \end{array} \tag{22}$$

where c is a nonvanishing vector, $x = (x_j)$ the real-valued variables and e the vector of all whose entries are one [2, 6];

$$e = (1, 1, \cdots, 1)^T. \tag{23}$$

The account for *unconstrained* is that we usually pose additional linear constraints expressed as $Ax = b$ to the canonical problem, where A and b are a matrix and a vectors suitably chosen. Due to the constraint $e^T x = 1$ in (22), Karmarkar's algorithm for our canonical problem is described on the $m-1$ dimensional canonical simplex,

$$\mathcal{S}_m = \Big\{ x \in \mathbf{R}^m \,\Big|\, \sum_{j=1}^{m} x_j = 1,\ x_j \geq 0\,(j = 1, 2, \cdots, m) \Big\}. \tag{24}$$

Accoding to Karmarkar, the continuous-time version of the Karmarkar's algorithm takes the form

$$\frac{\mathrm{d}x_j}{\mathrm{dt}} = -c_j x_j^2 + x_j\Big(\sum_{k=1}^{m} c_k x_k^2\Big) \quad (j = 1, 2, \cdots, m). \tag{25}$$

The dynamical system governed by (25) is what we call the Karmarkar flow. In present article, both of the differential equation (25) itself and the family of trajectories governed by (25) will be also referred to as the Karmarkar flow.

Following Nakamura [6], we show that the Karmarkar flow (25) admits the gradient system form. To be free from boundary of $\mathcal{S}_m$, we restrict (25) to the regular part

$$\dot{\mathcal{S}}_m = \Big\{x \in \mathbf{R}^m \;\Big|\; \sum_{j=1}^{m} x_j = 1,\ x_j > 0\,(j = 1, 2, \cdots, m)\Big\}, \tag{26}$$

of the simplex $\mathcal{S}_m$. With $\dot{\mathcal{S}}_m$, we endow the Riemannian metric

$$((u, u'))_x^{Smp} = \sum_{j=1}^{m} \frac{u_j u'_j}{x_j} \quad (u, u' \in T_x\dot{\mathcal{S}}_m) \tag{27}$$

where $T_x\dot{\mathcal{S}}_m$ denotes the tangent space of $\dot{\mathcal{S}}_m$ at $x \in \dot{\mathcal{S}}_m$ defined to be

$$T_x\dot{\mathcal{S}}_m = \Big\{u \in \mathbf{R}^m \;\Big|\; \sum_{j=1}^{m} u_j = 0\Big\}. \tag{28}$$

If we take the potential function

$$k(x) = \frac{1}{2}x^T C x \quad (x \in \dot{\mathcal{S}}_m), \qquad C = \mathrm{diag}\,(c_1, c_2, \cdots, c_m), \tag{29}$$

the dynamical system (25) is brought into the gradient equation

$$\frac{\mathrm{d}x}{\mathrm{dt}} = -(\mathrm{grad}\ k)(x) \tag{30}$$

on $\dot{\mathcal{S}}_m$, where the symbol $(\mathrm{grad}\ k)(x)$ denotes the gradient vector at $x \in \dot{\mathcal{S}}_m$ for $k(x)$. The $(\mathrm{grad}\ k)(x)$ is defined to satisfy

$$(((\mathrm{grad}\ k)(x), u'))_x^{Smp} = \frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} k(\sigma(t)) \quad ({}^\forall u' \in T_x\dot{\mathcal{S}}_m), \tag{31}$$

where $\sigma(t)$ with $a \le t \le b$ $(a < 0 < b)$ is a curve in $\dot{\mathcal{S}}_m$ subject to

$$t \in [a,\, b] \mapsto \sigma(t) \in \dot{\mathcal{S}}_m, \quad \sigma(0) = x, \quad \frac{\mathrm{d}\sigma}{\mathrm{dt}}\Big|_{t=0} = u' \tag{32}$$

(cf. (8) with (9)). By straightforward calculation, we obtain the gradient form

$$\frac{\mathrm{d}x_j}{\mathrm{dt}} = -\big((\mathrm{grad}\ k)(x)\big)_j = -c_j x_j^2 + x_j\Big(\sum_{k=1}^{m} c_k x_k^2\Big) \quad (j = 1, 2, \cdots, m). \tag{33}$$

Proposition 3.1 *[6] The Karmarkar flow (25) admits the gradient system form (30) associated with the potential function $k(x)$ given by (29).*

3.2 The gradient system realizing the Karmarkar flow on the QIS

As a crucial key in realizing the Karmarkar flow as a gradient system on the QIS, we study the Riemannian submanifold,

$$D_m = \left\{\Theta \in \dot{Q}_m \,\middle|\, \Theta = \text{diag}\,(\theta_1, \cdots, \theta_m),\ \text{tr}\,\Theta = 1,\ \theta_j > 0\ \ (j = 1, 2, \cdots, m)\right\}, \tag{34}$$

of the QIS, whose tangent space at Θ is given by

$$T_\Theta D_m = \left\{\mathrm{Z} \in M(m, m) \,\middle|\, \mathrm{Z} = \text{diag}\,(\zeta_1, \cdots, \zeta_m),\ \text{tr}\,\mathrm{Z} = 0\right\}. \tag{35}$$

To show that D_m is isometrically diffeomorphic to the canonical simplex $\dot{\mathcal{S}}_m$ endowed with the metric $((\cdot,\cdot))^{Smp}$, we consider the inclusion map,

$$\iota^{D_m} : \Theta \in D_m\,(\subset \dot{Q}_m) \mapsto \Theta \in \dot{Q}_m, \tag{36}$$

whose differential is given by

$$\iota^{D_m}_{*,\Theta}(\mathrm{Z}) = \mathrm{Z} \in T_\Theta \dot{Q}_m \quad (\mathrm{Z} \in T_\Theta D_m). \tag{37}$$

The submanifold D_m is then allowed to have the Riemannian metric $((\cdot,\cdot))^D$ defined by

$$\begin{aligned} ((\mathrm{Z}, \mathrm{Z}'))^D_\Theta &= ((\iota^{D_m}_{*,\Theta}(\mathrm{Z}), \iota^{D_m}_{*,\Theta}(\mathrm{Z}')))^{QF}_\Theta \\ &= ((\mathrm{Z}, \mathrm{Z}'))^{QF}_\Theta \quad (\mathrm{Z}, \mathrm{Z}' \in T_\Theta D_m), \end{aligned} \tag{38}$$

which makes D_m the Riemannian submanifold of the QIS $(\dot{Q}_m, ((\cdot,\cdot))^{QF})$. As for the differentiable one-to-one onto map

$$\alpha : x \in \dot{\mathcal{S}}_m \mapsto \text{diag}\,(x_1, \cdots, x_m) \in D_m \subset \dot{Q}_m \tag{39}$$

of $\dot{\mathcal{S}}_m$ to D_m, we show the following.

Lemma 3.2 *The Riemannian submanifold $(D_m, ((\cdot,\cdot))^D)$ of the QIS $(\dot{Q}_m, ((\cdot,\cdot))^{QF})$ is isometrically diffeomorphic to the canonical simplex $(\dot{\mathcal{S}}_m, ((\cdot,\cdot))^{Smp})$.*

Proof. From the definition (39) of α, we immediately see that α is a one-to-one and onto map, so that we have only to show an isometric property of α below. In fact, since the differential, $\alpha_{*,x}$ of the map α at $x \in \dot{\mathcal{S}}_m$ takes the form

$$\alpha_{*,x}(u) = \text{diag}\,(u_1, \cdots, u_m) \in T_\Theta D_m \quad (u \in T_x \dot{\mathcal{S}}_m). \tag{40}$$

we have the identity,

$$\begin{aligned} &((\mu_{*,x}(u), \mu_{*,x}(u')))^D_{\mu(x)} = ((\mu_{*,x}(u), \mu_{*,x}(u')))^{QF}_{\mu(x)} \\ &= 2\sum_{j,k=1}^{m} \frac{\overline{(\mu_{*,x}(u))_{jk}}\,(\mu_{*,x}(u'))_{jk}}{x_j + x_k} = \sum_{j=1}^{m} \frac{u_j u'_j}{x_j} = ((u, u'))^{Smp}_x, \end{aligned} \tag{41}$$

which shows that α is an isometry. This completes the proof.

Using the isometric diffeomorphism α, we transfer the gradient vector field grad k for the Karmarkar flow on $\dot{\mathcal{S}}_m$ to the vector field denoted by $\alpha_*(\text{grad } k)$ on D_m, which is determined through

$$\big(\alpha_*(\text{grad } k)\big)(\alpha(x)) = \alpha_{*,x}\big((\text{grad } k)(x)\big) \quad (x \in \dot{\mathcal{S}}_m). \tag{42}$$

Let us seek the gradient vector field, denoted by grad $\kappa(\mathrm{P})$, on the QIS subject to

$$(\text{grad } \kappa)(\Theta) = \alpha_{*,x}\big((\text{grad } k)(x)\big) \quad (\Theta = \alpha(x) \in D_m \subset \dot{Q}_m), \tag{43}$$

where κ denotes the associated potential function. We draw a necessary condition for the potential function κ. To do this, let us consider a curve $r(t)$ subject to $r(t) = \alpha(\sigma(t))$, where t is in a sufficiently small interval $[a, b]$ with $a < 0 < b$ and $\sigma(t)$ satisfies (32). The curve $r(t)$ then satisfies

$$\begin{aligned} &t \in [a,\, b] \mapsto r(t) = \alpha(\sigma(t)) \in D_m, \quad r(0) = \alpha(\sigma(0)), \\ &\left.\frac{\mathrm{d}r}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\alpha(\sigma(t))\right|_{t=0} = \alpha_{*,x}(u'). \end{aligned} \tag{44}$$

Owing to the diffeomorphism α of $\dot{\mathcal{S}}_m$ to D_m, we can write down any $\mathrm{Z}' \in T_{\alpha(x)}D_m$ in the form

$$\mathrm{Z}' = \alpha_{*,x}(u') \tag{45}$$

with $u' \in T_x\dot{\mathcal{S}}_m$. Then, Equation (43) is put together with (37) and (38) to show

$$\begin{aligned} &\langle\!\langle (\text{grad } \kappa)(\Theta), \mathrm{Z}' \rangle\!\rangle_\Theta^{QF} = \langle\!\langle \alpha_{*,x}((\text{grad } k)(x)), \mathrm{Z}' \rangle\!\rangle_\Theta^{QF} \\ = \;& \langle\!\langle \alpha_{*,x}((\text{grad } k)(x)), \mathrm{Z}' \rangle\!\rangle_\Theta^{D} = \langle\!\langle \alpha_{*,x}((\text{grad } k)(x)), \alpha_{*,x}(u') \rangle\!\rangle_\Theta^{D} \\ = \;& \langle\!\langle (\text{grad } k)(x), u' \rangle\!\rangle_x^{Smp} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} k(\sigma(t)). \end{aligned} \tag{46}$$

The lhs of eq.(46) admits an alternative calculation,

$$\begin{aligned} \langle\!\langle \text{grad } \kappa(\Theta), \mathrm{Z}' \rangle\!\rangle_\Theta^{QF} &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \kappa(r(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \kappa(\alpha(\sigma(t))) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\kappa \circ \alpha)(\sigma(t)), \end{aligned} \tag{47}$$

which leads us to

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} k(\sigma(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\kappa \circ \alpha)(\sigma(t)). \tag{48}$$

Namely, (48) implies a necessary condition

$$(\kappa \circ \alpha)(x) - k(x) = \text{constant} \quad (x \in \dot{\mathcal{S}}_m). \tag{49}$$

for (48), which suggests us to take κ to be

$$\kappa(\mathrm{P}) = \frac{1}{2}\mathrm{tr}\,\left(C\mathrm{P}^2\right), \tag{50}$$

where C is the diagonal matrix given in (29). The matrix $\mathcal{M}(\kappa)$ given by (11) with $\phi(\mathrm{P}) = \kappa(\mathrm{P})$ is calculated to be

$$\mathcal{M}(\kappa) = \frac{1}{2}\big(C\mathrm{P} + \mathrm{P}C\big). \tag{51}$$

Putting (50) and (51) into (21), we obtain the gradient system subject to the equation

$$\frac{\mathrm{dP}}{\mathrm{dt}} = -(\mathrm{grad}\ \kappa)(\mathrm{P}) = -\frac{1}{4}(\mathrm{P}^2C + 2\mathrm{P}C\mathrm{P} + C\mathrm{P}^2) + \Big(\mathrm{tr}\ (C\mathrm{P}^2)\Big)\mathrm{P}. \tag{52}$$

Since the gradient vector $(\mathrm{grad}\ \kappa)(\Theta)$ at $\Theta \in D_m \subset \dot{Q}_m$ is shown to be tangent to D_m, we can restrict (52) to the submanifold D_m of $\dot{Q}_m$. The restriction gives the differential equation

$$\frac{\mathrm{d}\Theta}{\mathrm{dt}} = -C\Theta^2 + \Big(\mathrm{tr}\ (C\Theta^2)\Big)\Theta \tag{53}$$

on D_m, which coincides with the Karmarkar flow. Thus, we reach to the following theorem.

Theorem 3.3 *[15, 10] The gradient system $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \kappa)$ on the QIS realizes the Karmarkar flow on the submanifold D_m of the QIS.*

3.3 Trajectories of the Karmarkar flow on the QIS

The solution of the Karmarkar flow on the simplex $\dot{\mathcal{S}}_m$ is given by Nakamura [6]. In a matrix form, Nakamura's solution is brought into the solution of (52) with the initial condition $\mathrm{P}(0) = \Theta(0)$,

$$\Theta(t) = \Theta(0)\Big(\xi(t)C\Theta(0) + I\Big)^{-1} \cdot \left\{\mathrm{tr}\left(\Theta(0)\Big(\xi(t)C\Theta(0) + I\Big)^{-1}\right)\right\}^{-1}, \tag{54}$$

where $\xi(t)$ satisfies

$$t = \mathrm{tr}\left(C^{-1}\log\Big(\xi(t)C\Theta(0) + I\Big)\right). \tag{55}$$

Let us start with checking the realized Karmarkar flow satisfies (54). By (55), we obtain the equation

$$\frac{\mathrm{d}\xi}{\mathrm{dt}} = \left\{\mathrm{tr}\left(\Theta(0)\Big(\xi(t)C\Theta(0) + I\Big)^{-1}\right)\right\}^{-1}. \tag{56}$$

Then, equation (54) takes the form

$$\Theta(t) = \frac{\mathrm{d}\xi}{\mathrm{dt}}\Theta(0)\Big(\xi(t)C\Theta(0) + I\Big)^{-1}, \tag{57}$$

whose derivative is expressed as

$$\frac{\mathrm{d}\Theta}{\mathrm{dt}} = \frac{\mathrm{d}\xi^2}{\mathrm{dt}^2}\Theta(0)\Big(\xi(t)C\Theta(0) + I\Big)^{-1} + \frac{\mathrm{d}\xi}{\mathrm{dt}}\Theta(0)\frac{\mathrm{d}}{\mathrm{dt}}\left\{\Big(\xi(t)C\Theta(0) + I\Big)^{-1}\right\}. \tag{58}$$

From (57), since we obtain the equations

$$\frac{\mathrm{d}\xi^2}{\mathrm{dt}^2} = \frac{\mathrm{d}\xi}{\mathrm{dt}}\mathrm{tr}\,\Big(C\Theta(t)^2\Big)$$
$$\frac{\mathrm{d}}{\mathrm{dt}}\left\{\Big(\xi(t)C\Theta(0)+I\Big)^{-1}\right\} = -C\Theta(t)\Big(\xi(t)C\Theta(0)+I\Big)^{-1}, \qquad (59)$$

equation (58) is calculated to be

$$\frac{\mathrm{d}\Theta}{\mathrm{dt}} = -C\Theta^2 + \Big(\mathrm{tr}\,(C\Theta^2)\Big)\Theta$$

Thus, we have shown that the solution of the realized Karmarkar flow on D_m is expressed by (54). To summarize, we have the following proposition.

Proposition 3.4 *The solution of the gradient system* $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \kappa)$ *running on the submanifold* D_m *of the QIS is expressed as (54).*

We move to the next stage, where we consider the solution of the realized Karmarkar flow on the QIS with a special object function. Namely, we study the case of $C = -2I$ in (52). In the case, the gradient system $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \kappa)$ also realizes the eigenvalue problem of the anti-Hermitian matrices. On looking carefully at the solution (54) on D_m, we conjecture that the solution on the QIS with $C = -2I$ takes the form

$$\mathrm{P}(t) = \mathrm{P}(0)\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\cdot\left\{\mathrm{tr}\,\left(\mathrm{P}(0)\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\right)\right\}^{-1}, \qquad (60)$$

where τ is defined by

$$t = -\frac{1}{2}\mathrm{tr}\,\left(\log\Big(-2\tau(t)\mathrm{P}(0)+I\Big)\right). \qquad (61)$$

We are to confirm that our conjecture is correct. From (61), we get the derivation of τ

$$\frac{\mathrm{d}\tau}{\mathrm{dt}} = \left\{\mathrm{tr}\,\left(\mathrm{P}(0)\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\right)\right\}^{-1}, \qquad (62)$$

which is combined with (60) to show

$$\mathrm{P}(t) = \frac{\mathrm{d}\tau}{\mathrm{dt}}\mathrm{P}(0)\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}. \qquad (63)$$

The derivation of (62) and of (63) yield

$$\frac{\mathrm{dP}}{\mathrm{dt}} = \frac{\mathrm{d}\tau^2}{\mathrm{dt}^2}\mathrm{P}(0)\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1} + \frac{\mathrm{d}\tau}{\mathrm{dt}}\mathrm{P}(0)\frac{\mathrm{d}}{\mathrm{dt}}\left\{\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\right\} \qquad (64)$$

and

$$\frac{\mathrm{d}\tau^2}{\mathrm{dt}^2} = -2\frac{\mathrm{d}\tau}{\mathrm{dt}}\mathrm{tr}\,\Big(\mathrm{P}(t)^2\Big), \qquad (65)$$

respectively. Further, since (63) provides

$$\frac{\mathrm{d}}{\mathrm{dt}}\left\{\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\right\}=2\Big(-2\tau(t)\mathrm{P}(0)+I\Big)^{-1}\mathrm{P}(t). \qquad (66)$$

Equations (64), (65) and (66) are put into Eq.(63) to show

$$\frac{\mathrm{dP}}{\mathrm{dt}}=2\mathrm{P}^2-2\Big(\mathrm{tr}\,(\mathrm{P}^2)\Big)\mathrm{P}$$

Thus, we found that the solution of the Karmarkar flow on the QIS with $C=-2I$ is expressed by (60). To summarize, we have the following proposition.

Proposition 3.5 *The solution of the gradient system $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \kappa)$ on the QIS if $C=-2I$ in (52) is expressed as (60).*

We study the asymptotic behavior of the solution (60) in turn as $t\longrightarrow\infty$. Since τ tends to infinity as $t\longrightarrow\infty$, the limit of P as $t\longrightarrow\infty$ is calculated to be

$$\begin{aligned}\lim_{t\to\infty}\mathrm{P}(t)&=\lim_{t\to\infty}\left\{\frac{1}{\tau(t)}\mathrm{P}(0)\Big(-2\mathrm{P}(0)+\frac{1}{\tau(t)}I\Big)^{-1}\right.\\&\qquad\left.\cdot\left\{\mathrm{tr}\left(\mathrm{P}(0)\frac{1}{\tau(t)}\Big(-2\mathrm{P}(0)+\frac{1}{\tau(t)}I\Big)^{-1}\right)\right\}^{-1}\right\}\\&=\mathrm{P}(0)\Big(-2\mathrm{P}(0)\Big)^{-1}\left\{\mathrm{tr}\left(\mathrm{P}(0)\Big(-2\mathrm{P}(0)\Big)^{-1}\right)\right\}^{-1}=\frac{1}{m}\end{aligned}$$

from (60). Therefore, we reach to the following.

Proposition 3.6 *The gradient system $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \kappa)$ on the QIS if $C=-2I$ in (52) converges on the center-of-mass of the simplex as $t\longrightarrow\infty$.*

On closing this subsection, we remark on a physical meaning of the case $C=-2I$. In this case, the potential function κ is understood to be the Renyi potential which measures the 'purity' of the quantum states [18] associated with density matrices, P. In this sense, Proposition 3.6 looks very natural since in our gradient system with $C=-2I$ all the trajectories tend to the density matrices with the highest purity. The solutions of (52) with general C haven't been revealed explicitly yet.

4 The ALEH realized on the QIS

4.1 Learning equation

We start with an introduction of a standard neuronal model. Let $q(s)$ and $r(s)$ be the $\mathbf{R}^m$-valued functions, which carry the input signals to a synaptic neuron and the coupling strength coefficients of inputs, respectively. The membrane potential value expressed by the $\mathbf{R}$-valued function $p(s)$ is determined by the relation

$$p(s)=q^T(s)r(s). \qquad (67)$$

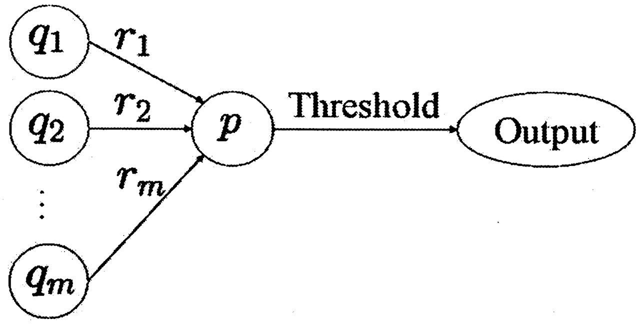

Figure 1: The neuronal model

Figure 4.1 presents a graphical description for our setting with (67).

In neural network theory, learning amounts to updating the coupling strength coefficient to improve the output along with a learning rule. This idea comes from Hebb's hypothesis [19]; *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.* According to our setting, Hebb's hypothesis is brought to the recurrence rule

$$r(s+1) = r(s) + \varepsilon p(s)q(s), \tag{68}$$

where ε stands for the learning rate parameter. Unfortunately, the rule (68) allows an unbounded growth of the norm $\|r(s)\|$ of the coupling strength coefficient vector $r(s)$, which is usually unacceptable. To avoid the unbounded growth of $r(s)$, we apply Oja's rule [3]

$$r(s+1) = \frac{r(s) + \varepsilon p(s)q(s)}{\|r(s) + \varepsilon p(s)q(s)\|}, \tag{69}$$

in this thesis, which admits the norm-preserving property,

$$\|r(s)\| = 1 \quad \text{for} \quad \|r(0)\| = 1. \tag{70}$$

We derive a differential equation from Oja's rule (69) with the learning equation (67) as follows. Under the relation (67), Eq.(69) is put in the form

$$r(s+1) = \frac{r(s) + \varepsilon q(s)q^T(s)r(s)}{\|r(s) + \varepsilon q(s)q^T(s)r(s)\|}, \tag{71}$$

which admits the Maclaurin expansion

$$r(s+1) = r(s) + \varepsilon\{q(s)q^T(s)r(s) - \left(r^T(s)q(s)q^T(s)r(s)\right) r(s)\} + O(\varepsilon^2) \tag{72}$$

for sufficiently small ε. The $O(\varepsilon^2)$ in (72) denotes the second-order infinitesimal. Elimination of the term $O(\varepsilon^2)$ from the rhs of (72) therefore provides us with the recurrence equation

$$r(s+1) = r(s) + \varepsilon\big\{q(s)q^T(s)r(s) - \left(r^T(s)q(s)q^T(s)r(s)\right) r(s)\big\}. \tag{73}$$

Equation (73) will be averaged in the succeeding subsection, which leads us to the ALEH through a continuous limit in time.

4.2 The ALEH

We are to average Eq.(73) under the assumption made as follows: The input signals $q(s)$ and the coupling strength coefficients $r(s)$ are assumed to be stochastic processes which are statistically independent to each other. Further, the stochastic process $q(t)$ is a stationary process [3]. Taking the expectation of (73) on the sample space, we obtain

$$\begin{aligned} & E[r(s+1) \mid r(s)] - E[r(s)] \\ = \; & \varepsilon\{E[q(s)q^T(s)]E[r(s)] - \big(E[r(s)]^T E[q(s)q^T(s)]E[r(s)]\big)E[r(s)]\}, \end{aligned} \tag{74}$$

where the symbol $E[\cdot]$ denotes expectation operation. From the stationaries of $q(s)$, the correlation matrix $E[q(s)q^T(s)]$ turns out to be invariant along s. Hence, when $E[q(s)q^T(s)]$ is diagonalized to be

$$A = \text{diag}\ (a_1, a_2, \cdots, a_m) = G^T E[q(s)q^T(s)]G, \tag{75}$$

with an orthogonal matrix G, both the orthogonal matrix G and the diagonal matrix A are invariant along s. The change of time-variables

$$t = \varepsilon s \tag{76}$$

with the introduction,

$$w(t) = (w_1(t), w_2(t), \cdots, w_m(t))^T = G^T r(t/\varepsilon), \tag{77}$$

of $w(t)$ brings Eq.(74) to the form

$$w(t+\varepsilon) - w(t) = \varepsilon\Big(Aw(t) - \big(w^T(t)Aw(t)\big)w(t)\Big). \tag{78}$$

In the case that the learning rate parameter ε is sufficiently small, Eq.(78) is understood as the forward Euler integration formula to the differential equation

$$\frac{\mathrm{d}w}{\mathrm{d}t} = Aw - (w^T Aw)w. \tag{79}$$

The equation (79) is what is dealt with as a continuous-time analogue of (78) [7, 3]. The (79) is referred to as the averaged learning equation of Hebb type, which will be abbreviated to the ALEH throughout this article.

We are to draw the gradient system form of the ALEH following Nakamura [7] in what follows. To do this, we start with the norm preserving property

$$\|w(t)\| = 1 \quad \text{for} \quad \|w(0)\| = 1 \tag{80}$$

of $w(t)$ subject to the ALEH (79), which is understood to be a counterpart to (70) of (67) with Oja's rule (69). Indeed, (80) is ensured by the calculation,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|^2 &= 2w^T(t)\frac{\mathrm{d}w}{\mathrm{d}t}(t) = 2w^T(t)\big(Aw - (w^T Aw)w\big) \\ &= 2\,(w^T Aw)\,(1 - \|w(t)\|^2) = 0, \end{aligned} \tag{81}$$

under $\|w(0)\| = 1$. Since the norm preserving property (80) is put together with (76) and (77) to yield the norm preserving property

$$\|w(t)\| = \|Gr(\varepsilon t)\| = \|r(\varepsilon t)\| = \|r(s)\| = 1 \tag{82}$$

of $r(s)$, which is understood as a counterpart of (70).

Owing to the norm preserving property (80), we can deal with the ALEH as a differential equation on the $(m-1)$-dimensional unit sphere,

$$S^{m-1} = \left\{ w \in \mathbf{R}^m \,\middle|\, \|w\| = 1 \right\}. \tag{83}$$

To avoid excessive calculus on boundaries of the QIS in the succeeding section, we will restrict our discussion henceforth on the dense submanifold

$$\dot{S}^{m-1} = S^{m-1} \backslash \cup_{k=1}^{m} \mathcal{N}_m^{(k)} \tag{84}$$

where $\mathcal{N}_m^{(k)}$s are measure-zero subsets of the sphere S^{m-1} defined by

$$\mathcal{N}_m^{(k)} = \{ w \in S^{m-1} \,|\, w_k = 0 \} \quad (k = 1, 2, \cdots, m). \tag{85}$$

With $\dot{S}^{m-1}$, we endow the standard Riemannian metric

$$((v, v'))_w^{Sph} = v^T v' \quad (v, v' \in T_w \dot{S}^{m-1},\ w \in \dot{S}^{m-1}), \tag{86}$$

where $T_w \dot{S}^{m-1}$ denotes the tangent space of $\dot{S}^{m-1}$ at w defined to be

$$T_w \dot{S}^{m-1} = \left\{ v \in \mathbf{R}^m \,\middle|\, w^T v = 0 \right\} \quad (w \in \dot{S}^{m-1}). \tag{87}$$

To show that the ALEH on $\dot{S}^{m-1}$ admits the gradient system form, we introduce the function

$$\ell(w) = -\frac{1}{2} w^T A w = -\frac{1}{2} \sum_{k=1}^{m} a_k {w_k}^2 \quad (w \in \dot{S}^{m-1}) \tag{88}$$

as the potential function for our gradient system form. The gradient vector field , denoted by grad ℓ, for the gradient system $(\dot{S}^{m-1}, ((\cdot,\cdot))^{Sph}, \ell)$ is defined by

$$((\mathrm{grad}\ \ell(w), v'))_w^{Sph} = \left. \frac{\mathrm{d}}{\mathrm{dt}} \right|_{t=0} \ell(\gamma(t)) \quad (v' \in T_w \dot{S}^{m-1}). \tag{89}$$

where γ with $a \le t \le b (a < 0 < b)$ is a curve in $\dot{S}^{m-1}$ subject to

$$t \in [a,\, b] \mapsto \gamma(t) \in \dot{S}^{m-1}, \quad \gamma(0) = w, \quad \left. \frac{\mathrm{d}\gamma}{\mathrm{dt}} \right|_{t=0} = v' \in T_w \dot{S}^{m-1}. \tag{90}$$

By straightforward calculation, we obtain the gradient system form

$$\frac{\mathrm{d}w}{\mathrm{dt}} = -(\mathrm{grad}\ \ell)(w) = Aw - (w^T A w) w. \tag{91}$$

We have the following proposition.

Proposition 4.1 *[7] The averaged learning equation of Hebb type (ALEH) (79) admits the gradient system form (91) associated with the potential function $\ell(w)$ given by (88).*

4.3 Geometric setting for the ALEH on the QIS

To realize the ALEH on the QIS, we prepare a pair of geometric devices. One is a metric-preserving immersion from $\dot{S}^{m-1}$ to the submanifold D_m of the QIS, where D_m consists of diagonal matrices in the QIS (see (34) for definition). Another is a natural $(\mathbf{Z}_2)^m$ action on $\dot{S}^{m-1}$.

In the case of the Karmarkar flow dealt with in section 2, the map α of $\dot{\mathcal{S}}_m$ to D_m is surjective and injective, so that the Karmarkar flow can be mapped to D_m through α immediately. By contrast, the metric-preserving immersion which we are going to apply to the ALEH is surjective but not injective since the disconnected manifold $\dot{S}^{m-1}$ is mapped onto the connected manifold D_m through the immersion. Due to the non-injectivity, we have to ensure that the gradient vectors at $w \in \dot{S}^{m-1}$ and $w' \in \dot{S}^{m-1}$ are mapped to the same tangent vector at $\Theta \in D_m$ if both w and w' are mapped to Θ. To check this, the natural $(\mathbf{Z}_2)^m$ action plays a key role in what follows.

Let us start with introducing the map β of $\dot{S}^{m-1}$ to D_m of the form

$$\beta : w = (w_1, w_2, \cdots, w_m)^T \in \dot{S}^{m-1} \mapsto \mathrm{diag}\,({w_1}^2, {w_2}^2, \cdots, {w_m}^2) \in D_m \tag{92}$$

together with the natural $(\mathbf{Z}_2)^m$ action of the form

$$\mu_\sigma : (w_1, w_2, \cdots, w_m)^T \in \dot{S}^{m-1} \mapsto (\sigma_1 w_1, \sigma_2 w_2, \cdots, \sigma_m w_m)^T \in \dot{S}^{m-1} \tag{93}$$

on $\dot{S}^{m-1}$, where $\sigma \in (\mathbf{Z}_2)^m$ takes the form

$$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_m)^T \in (\mathbf{Z}_2)^m \quad (\mathbf{Z}_2 = \{-1, 1\}). \tag{94}$$

The differential maps of β and μ at $w \in \dot{S}^{m-1}$ turn out to take the forms,

$$\beta_{*,w}(v) = 2\,\mathrm{diag}\,(w_1 v_1, w_2 v_2, \cdots, w_m v_m) \quad (v \in T_w \dot{S}^{m-1}) \tag{95}$$

and

$$(\mu_\sigma)_{*,w}(v) = (\sigma_1 v_1, \sigma_2 v_2, \cdots, \sigma_m v_m)^T \quad (v \in T_w \dot{S}^{m-1},\ \sigma \in (\mathbf{Z}_2)^m), \tag{96}$$

respectively. We show the following lemmas.

Lemma 4.2 *The map β given by (92) is an isometric immersion of $\dot{S}^{m-1}$ to D_m. The isometric property is thought of up to the constant multiple* 4.

Proof. Owing to (95), we easily see that β is an immersion (see [14] for the definition of immersion, for example). What we have to show is then an isometric property of β. Using (36), (37), (38), (86), (92) and (95), we confirm

$$\begin{aligned}
\langle\!\langle \beta_{*,w}(v), \beta_{*,w}(v') \rangle\!\rangle^D_{\beta(w)} &= \langle\!\langle (\iota^{D_m}_{*,\beta(w)} \circ \beta_{*,w})(v), (\iota^{D_m *}_{*,\beta(w)} \circ \beta_{*,w})(v') \rangle\!\rangle^{QF}_{\beta(w)} \\
&= \langle\!\langle \beta_{*,w}(v), \beta_{*,w}(v') \rangle\!\rangle^{QF}_{\beta(w)} \\
&= 2 \sum_{j,k=1}^{m} \frac{\overline{(\beta_{*,w}(v))_{jk}}\,(\beta_{*,w}(v'))_{jk}}{{w_j}^2 + {w_k}^2} \\
&= 4 \sum_{j=1}^{m} v_j v'_j = 4\,\langle\!\langle v, v' \rangle\!\rangle^{Sph}_w
\end{aligned} \tag{97}$$

for any $w \in \dot{S}^{m-1}$ and $v, v' \in T_w \dot{S}^{m-1}$, which proves our assertion.

Lemma 4.3 *The map β of $\dot{S}^{m-1}$ to D_m is invariant under the $(\mathbf{Z}_2)^m$ action σ given by (94). The invariance stands for the equation*

$$(\beta \circ \mu_\sigma)(w) = \beta(\mu_\sigma(w)) = \beta(w) \quad (w \in \dot{S}^{m-1},\ \sigma \in (\mathbf{Z}_2)^m). \tag{98}$$

Proof. On account of $(\sigma_j)^2 = 1$ $(j = 1, 2, \cdots, m)$, we immediately have

$$\begin{aligned}(\beta \circ \mu_\sigma)(w) = \beta(\mu_\sigma(w)) &= \text{diag}\left((\sigma_1 w_1)^2, (\sigma_2 w_2)^2, \cdots, (\sigma_m w_m)^2\right) \\ &= \text{diag}\,(w_1^2, w_2^2, \cdots, w_m^2) = \beta(w),\end{aligned} \tag{99}$$

which completes the proof.

We are now in a position to show the $(\mathbf{Z}_2)^m$ invariance of the gradient vector field grad ℓ given by (91). Namely, we show

$$\mu_{*,w}\left((\text{grad}\ \ell)(w)\right) = (\text{grad}\ \ell)(\mu_\sigma(w)) \quad (w \in \dot{S}^{m-1},\ \sigma \in (\mathbf{Z}_2)^m). \tag{100}$$

Indeed, on putting (91) and (95) together, the rhs of (100) is calculated to be

$$\begin{aligned}&(\text{grad}\ \ell)(\mu_\sigma(w)) = -A\mu_\sigma(w) + (\mu_\sigma(w)^T A \mu_\sigma(w))\mu_\sigma(w) \\ =\ & \text{diag}\,(\sigma_1, \sigma_2, \cdots, \sigma_m)\left\{-Aw + (w^T A w)w\right\} \\ =\ & \left(\sigma_1((\text{grad}\ \ell)(w))_1, \sigma_2((\text{grad}\ \ell)(w))_2, \cdots, \sigma_m((\text{grad}\ \ell)(w))_m\right)^T \\ =\ & (\mu_\sigma)_{*,w}(\text{grad}\ \ell(w)) \qquad (w \in \dot{S}^{m-1},\ \sigma \in (\mathbf{Z}_2)^m),\end{aligned} \tag{101}$$

which proves (100). To summarize, we have the following.

Lemma 4.4 *[10] The gradient vector field,* grad ℓ*, for the ALEH given by (91) is invariant under the $(\mathbf{Z}_2)^m$ action ((93).*

Lemmas 4.3 and 4.4 are put together with (95) and (96) to show the following.

Proposition 4.5 *[16] For the gradient vector field,* grad ℓ*, for the ALEH given by (91), there exists the unique vector field denoted by β_*(grad ℓ) on the submanifold D_m of the QIS which satisfies*

$$(\beta_*(\text{grad}\ \ell))(\Theta) = \beta_{*,\tilde{w}}\left((\text{grad}\ \ell))(\tilde{w})\right) \quad (\Theta \in D_m,\ \tilde{w} \in \beta^{-1}(\Theta)). \tag{102}$$

The $\beta^{-1}(\Theta)$ denotes the inverse image of $\Theta \in D_m$ given by

$$\begin{aligned}\beta^{-1}(\Theta) &= \{w \in \dot{S}^{m-1} \,|\, \beta(w) = \Theta\} \\ &= \{w \in \dot{S}^{m-1} \,|\, w_j = \sigma_j\sqrt{\Theta},\ \sigma = (\sigma_j) \in (\mathbf{Z}_2)^m\} \quad (\Theta \in D_m).\end{aligned} \tag{103}$$

Proof. We have only to show that the rhs of (102) is independent of the choice of $\tilde{w} \in \beta^{-1}(\Theta)$. In view of (103), we consider a pair of points, $\tilde{w}$ and $\tilde{w}'$, in $\beta^{-1}(\Theta)$, which turn to satisfy

$$\tilde{w}' = \mu_\sigma(\tilde{w}) \quad ({}^{\exists}\sigma \in (\mathbf{Z}_2)^m). \tag{104}$$

Further, differentiating (99), we have

$$\beta_{*,\mu_\sigma w} \circ (\mu_\sigma)_{*,w} = (\beta \circ \mu_\sigma)_{*,w} = \beta_{*,w} \quad (w \in \dot{S}^{m-1},\ \sigma \in (\mathbf{Z}_2)^m). \tag{105}$$

Then, for any $\tilde{w}, \tilde{w}' \in \beta^{-1}(\Theta)$ subject to (104), Equations.(101) and (105) are put together to show

$$\begin{aligned}\beta_{*,\tilde{w}'}\big((\text{grad } \ell)(\tilde{w}')\big) &= \beta_{*,\mu_\sigma(\tilde{w})}\big((\text{grad } \ell)(\mu_\sigma(\tilde{w}))\big) \\ &= \beta_{*,\mu_\sigma(\tilde{w})}\big((\mu_\sigma)_{*,\tilde{w}}((\text{grad } \ell)(\tilde{w}))\big) \\ &= \big(\beta_{*,\mu_\sigma(\tilde{w})} \circ (\mu_\sigma)_{*,\tilde{w}}\big)\big((\text{grad } \ell)(\tilde{w})\big) \\ &= \beta_{*,\tilde{w}}\big((\text{grad } \ell)(\tilde{w})\big). \end{aligned} \tag{106}$$

This ensures that the rhs of (102) is well-defined as a tangent vector of D_m at Θ. This completes the proof.

The vector field $\beta_*(\text{grad } \ell)$ on D_m can be understood to be a 'copy' of the gradient vector field grad ℓ in the following sense. Let us denote by $\mathcal{C}_m^\sigma$ ($\sigma \in (\mathbf{Z}_2)^m$) all the possible connected components of $\dot{S}^{m-1}$, which are described to be

$$\mathcal{C}_m^\sigma = \{w \in \dot{S}^{m-1} \,|\, w = \mu_\sigma w',\ w' \in \mathcal{C}_m^{\text{id}}\} \quad (\sigma \in (\mathbf{Z}_2)^m) \tag{107}$$

with

$$\mathcal{C}_m^{\text{id}} = \{w \in \dot{S}^{m-1} \,|\, w_j > 0 \ (j = 1, 2, \cdots, m)\}. \tag{108}$$

If the domain of the isometric immersion $\beta : \dot{S}^{m-1} \longrightarrow D_m$ is restricted to each of $\mathcal{C}_m^\sigma$ ($\sigma \in (\mathbf{Z}_2)^m$), then the restriction denoted by β^σ of β to $\mathcal{C}_m^\sigma$ turns out to be a diffeomorphism, namely, a differentiable one-to-one and onto map. Further, since the β^σ works in the same way as β to any $w \in \mathcal{C}_m^\sigma$, we have

$$\beta_{*,w}^\sigma\big((\text{grad } \ell)(w)\big) = \big(\beta_*(\text{grad } \ell)\big)(\beta^\sigma(w)) \quad (w \in \mathcal{C}_m^\sigma), \tag{109}$$

where $\beta_*(\text{grad } \ell)$ is the vector field on D_m defined by (102). This confirms our assertion. The vector field $\beta_*(\text{grad } \ell)$ on D_m is what should be realized by a gradient vector on the QIS in the succeeding section.

4.4 The gradient system on the QIS realizing the ALEH

We are to seek to a gradient vector field, denoted by grad $\lambda(\text{P})$, on the QIS which satisfies

$$\text{grad } \lambda(\Theta) = \beta_{*,w}(\text{grad } \ell(w)) \quad (\Theta = \beta(w) \in D_m \subset \dot{Q}_m), \tag{110}$$

where λ is the associated potential function. On an analogous line of thought to realize the Karmarkar flow in the QIS in section 3, we draw a necessary condition for the potential function for λ. To do this, we consider a curve $r(t)$ subject to $r(t) = \beta(\gamma(t))$, where t is in a sufficiently small interval $[a, b]$ with $a < 0 < b$ and $\sigma(t)$ satisfies (90). The curve $r(t)$ then satisfies

$$\begin{aligned} &t \in [a,\, b] \mapsto r(t) = \beta(\gamma(t)) \in D_m, \quad r(0) = \beta(w), \\ &\left.\frac{\mathrm{d}r}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\beta(\gamma(t))\right|_{t=0} = \beta_{*,w}(v'). \end{aligned} \tag{111}$$

Recalling that any $w \in \dot{S}^{m-1}$ belongs to a connected component $\mathcal{C}_m^\sigma$ of the QIS and that the restriction β^σ of the isometric immersion β is a diffeomorphism, we can write down any $Z' \in T_{\beta(w)}D_m$ in the form

$$\text{Z}' = \beta_{*,w}(v') = \beta_{*,w}^\sigma(v') \tag{112}$$

with $v' \in T_w\dot{S}^{m-1}$. Then, Equation (110) is put together with (37) and (38) to show

$$\begin{aligned} & ((\text{grad } \lambda(\Theta), Z'))_\Theta^{QF} = ((\beta_{*,w}(\text{grad } \ell(w)), Z'))_\Theta^{QF} \\ = & ((\beta_{*,w}(\text{grad } \ell(w)), Z'))_\Theta^{D} = ((\beta_{*,w}(\text{grad } \ell(w)), \beta_{*,w}(v')))_\Theta^{D} \\ = & 4((\text{grad } \ell(w), v'))_w^{Sph} = 4 \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \ell(\sigma(t)). \end{aligned} \tag{113}$$

Further, the definition of the gradient vector (see [14] and cf. (8)) gives rise to the other expression

$$\begin{aligned} ((\text{grad } \lambda(\Theta), Z'))_\Theta^{QF} &= \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \lambda(r(t)) = \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \lambda(\beta(\gamma(t))) \\ &= = \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} (\lambda \circ \beta)(\gamma(t)) \end{aligned} \tag{114}$$

of $((\text{grad } \lambda(\Theta), \mathrm{Z}'))_\Theta^{QF}$ than that given by (113), so that we have

$$\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \ell(\gamma(t)) = \left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} (\lambda \circ \beta)(\gamma(t)) \tag{115}$$

as an equivalent condition to (110). A necessary condition

$$(\lambda \circ \beta)(w) - 4\ell(w) = \text{constant} \quad (w \in \dot{S}^{m-1}). \tag{116}$$

for (115) suggests us to take λ to be

$$\lambda(\mathrm{P}) = -2\text{tr } (A\mathrm{P}), \tag{117}$$

as a candidate for the potential function satisfying (110), where A is the diagonal matrix given in (75). Indeed, we can confirm that (117) satisfies (116):

$$\lambda(\beta(w)) = -2\text{tr } (A\beta(w)) = -2\sum_{k=1}^{m} c_k {w_k}^2 = 4\left(-\frac{1}{2}\sum_{k=1}^{m} c_k {w_k}^2\right) = 4\ell(w). \tag{118}$$

The matrix $\mathcal{M}(\lambda)$ given by (11) with $\phi(\mathrm{P}) = \lambda(\mathrm{P})$ takes the form

$$\mathcal{M}(\lambda) = -2A, \tag{119}$$

Combining (21) with (119), the gradient system associated the potential function λ is expressed as

$$\frac{\mathrm{dP}}{\mathrm{dt}} = -(\text{grad } \lambda)(\mathrm{P}) = (\mathrm{P}A + A\mathrm{P}) - 2\Big(\text{tr } (A\mathrm{P})\Big)\mathrm{P}. \tag{120}$$

Since the gradient vector $(\text{grad } \lambda)(\Theta)$ at $\Theta \in D_m \subset \dot{Q}_m$ is tangent to D_m, we can restrict (120) to the submanifold D_m of $\dot{Q}_m$. The restriction gives

$$\frac{\mathrm{d}\Theta}{\mathrm{dt}} = 2A\Theta - 2\Big(\text{tr } (A\Theta)\Big)\Theta \in T_\Theta D_m \subset T_\Theta \dot{Q}_m, \tag{121}$$

which indicates the realization of the ALEH on the submanifold D_m of $\dot{Q}_m$ with ${w_j}^2 = \theta_j$ $(j = 1, 2, \cdots, m)$. Therefore, we reach the following theorem.

Theorem 4.6 *The gradient system $(\dot{Q}_m, ((\cdot,\cdot))^{QF}, \lambda)$ on the QIS realizes the ALEH on the submanifold D_m of the QIS.*

5 Concluding remarks

We have reported the realization of the engineering algorithms, the Karmarkar flow and the averaged learning equation of Hebb type (ALEH), as the gradient systems on the quantum information space (QIS). For ALEH, we have an alternative way to realize it on the QIS. The way is constructed by compositon of a pair of maps; one is the isometric surjection,

$$\Delta : w \in \dot{S}^{m-1} \mapsto \tilde{w} = (w_1^2, w_2^2, \cdots, w_m^2)^T \in \dot{\mathcal{S}}_m, \tag{122}$$

of $\dot{S}^{m-1}$ to $\dot{\mathcal{S}}_m$ and another is the map α of $\dot{\mathcal{S}}_m$ to D_m given by (39). We note here that Faybusovich used the map Δ in some works, which permits us to avoid a delicate treatment of the boundary of simplex for Karmarkar-type interior point algorithms [20, 21].

In the case of $C = -2I$, the gradient systems have following features. One is that the Karmarkar flow on the QIS is understood to be a generalization of an eigenvalue solver of anti-Hermitean matrices studied by Nakamura [4]. Another is that the potential function $\kappa(\mathrm{P})$ with $C = -2I$ turns out to express *purity* whose logarithm provides the quantum Renyi potential $\log(\mathrm{tr}\ (\mathrm{P}^q))/(1-q)$ with $q = 2$ [18]. Further, the gradient system realizing the ALEH in the case of $C = -2I$ turns out to be equivalent to a quantum inference problem dealt with Braunstein [17].

The future problems would be thought of as follows. One is on convergence of trajectories. In the present study, we cannot clearify the convergence of trajectories of gradient systems on the QIS except for the case of $C = -2I$. The other is to return discrete-time versions of the fKarmarkar flow and the ALEH on the QIS or on a certain Hilbert space, which is however seems to be *big*. Several approach could be thought of to this problem. For example, a 'lift' those systems to systems on a certain Hilbert space is worth applied (see [9] for example). After the lift, discretization in time could be applied. Alternatively, the order of lift and discritization could be switched.

Acknowlegement The authors would like to thank Professor Toshihiro Iwai at Kyoto University for his valuable comments on the present work. This work is partly supported by Special Research Budget **B14** of Future University Hakodate, 2009.

References

[1] N.Karmarkar, A new polynomial-time algorithm for linear programming, *Combinatorica*, Vol. 4, 373, 1984.

[2] N.Karmarkar, Riemannian geometry underlying interior-point methods for linear programming, *Contemporary Mathematics*, Vol. 114, 51, 1990.

[3] E.Oja, A Simplified Neuron Model as a Principal Component Analyzer, *Journal of mathematical biology*, Vol. 15, 267, 1982.

[4] Y.Nakamura, A new nonlinear dynamical system that leads to eigenvalues, *Japan Journal of Industrial and Applied Mathematics*, Vol. 9, 133, 1992.

[5] Y.Nakamura, Completely integrable gradient systems on the manifolds of gaussian and multinomial distributions, *Japan Journal of Industrial and Applied Mathematics*, Vol. 10, 179, 1993.

[6] Y.Nakamura, Lax pair and fixed point analysis of karmarkar's projective scaling trajectory for linear programming, *Japan Journal of Industrial and Applied Mathematics*, Vol. 11, 1, 1994.

[7] Y.Nakamura, Neurodynamics and nonlinear integrable systems of lax type, *Japan Journal of Industrial and Applied Mathematics*, Vol. 11, 11, 1994.

[8] Y.Uwano, Lax-type equation unifying gradient systems in quantum and classical statistical models, *Czechoslovak Journal of Physics*, Vol. 56, 1311, 2006.

[9] Y.Uwano, H.Hino and Y.Ishiwatari, Certain integrable system on a space associated with a quantum search algorithm, *Physics of Atomic Nuclei*, Vol. 70, 784, 2007.

[10] Y.Uwano and H.Yuya, A gradient system on the quantum information space realizing the karmarkar flow for linear programming – a clue to effective algorithms, *to appear in Electronic Journal of Theoretical Physics*, 2010.

[11] P.W.Shor, Algorithms for quantum computation: Discrete logarithms and factoring, In *Proceedings of 35th Annual Symposium on Foundations of Computer Science*, 124, 1994.

[12] L.K.Grover, A fast quantum mechanical algorithm for database search, In *Proceedings of 28th Annual ACM Symposium on the Theory of Computing*, 212, 1996.

[13] S.Amari and H.Nagaoka, *Methods of Information Geometry (Translations of Mathematical Monographs)*, chapter 7.3, Amer Mathematical Society, 2000.

[14] S.Kobayashi and K.Nomizu, *Foundations of Differential Geometry vol.2*, 337, Wiley, 1969.

[15] Y.Uwano and H.Yuya, A gradient system on the quantum information space that realizes the karmarkar flow for linear programming, *arXive:0807.4053v1[math.DS]*, 2008.

[16] Y.Uwano and H.Yuya, A gradient system on the quantum information space realizing the averaged learning equation of hebb type for the principal component analyzer, *arXive:0912.3328[math.DS]*, 2009.

[17] S.L.Braunstein, Geometry of quantum inference, *Physics Letter A*, Vol. 219, 169, 1996.

[18] I.Bengtsson and K.Zyczkowski, *Geometry of Quantum States : An Introduction to Quantum Entanglement*, 286, Cambridge University Press, 2006.

[19] D.O.Hebb, *The Organization of Behavior : a neuropsychological theory*, John Wiley and Sons, Chapman and Hall, 1949.

[20] L.Faybusovich, On a matrix generalization of affine-scaling vector fields, *Society for Industrial and Applied Mathematics Journal of Matrix Analysis and Applications 16(3)*, 886, 1995.

[21] L.Faybusovich, A Hamiltonian structure for generalized affine-scaling vector fields, *Journal of Nonlinear Science*, Vol. 5, 11, 1995.