数理解析研究所講究録
第 1671 巻 2009 年 20-36

Wasserstein geometry of non-linear Fokker–Planck type equations

東北大学大学院理学研究科 高津飛鳥 (Takatsu Asuka)
Mathematical Institute, Tohoku University

1. Introduction

This note is a survey of the author's preprint [17], which concerns the geometric structure of the q-Gaussian measures in terms of L^2-Wasserstein geometry and solutions to porous medium equations. We give an explicit expression of the solution to the porous medium equation when the initial data is a q-Gaussian measure.

Otto's remarkable paper [14] studied a formal Riemannian structure of the L^2-Wasserstein space, and gave applications to the study of porous medium equations. He showed that non-linear Fokker–Planck type equations can be considered to be gradient flows on the space of probability measures, equipped with the formal Riemannian manifold structure whose arc length distance coincides with the L^2-Wasserstein distance W_2. (The definition of W_2, which is given in the next section, has its roots in the Monge–Kantorovich transport theory. Note that the convergence in the sense of L^2-Wasserstein distance is somewhat stronger than weak convergence, see [21, Theorem 6.9].) Precisely speaking, the gradient flow of the Tsallis entropy E_m is the porous medium equation

$$\frac{\partial}{\partial t}\rho = \mathrm{grad}E_{m|\rho} = \Delta(\rho^m) \qquad \mathrm{PME}_m$$

for $m > d/(d+2)$, $m \geq (d-1)/d$ and $m \neq 1$. Here we identify a probability measure μ with its density. The Tsallis entropy E_m and its free energy density e_m are given by

$$E_m(\mu) = -\int_{\mathbb{R}^d} e_m\left(\frac{d\mu}{dx}\right)dx,$$

$$e_m(x) = \frac{x^m - x}{m-1}.$$

(See [18] for further discussions about Tsallis entropy.) When m converges to 1, the Boltzmann entropy is recovered, that is

$$E_m(\mu) \xrightarrow{m\to 1} E(\mu) = -\int_{\mathbb{R}^d} e\left(\frac{d\mu}{dx}\right)dx,$$

$$e_m(x) \xrightarrow{m\to 1} e(x) = x\ln x.$$

E-mail: sa6m21@math.tohoku.ac.jp.

Otto also demonstrated that the gradient flow of the Boltzmann entropy E is the heat equation

$$\frac{\partial}{\partial t}\rho = \mathrm{grad}E_{|\rho} = \Delta\rho. \tag{HE}$$

In what follow, E_1 stands for the Boltzmann entropy E and PME_1 stands for the heat equation HE. The Boltzmann entropy is obtained when all the component particles of a thermodynamic system as statistically independent. Note that the Boltzmann entropy is a Lyapunov functional, that is, monotonically increasing functional, under the heat equation.

For a solution ρ to PME_m, we define $\widetilde{\rho}$ by

$$\rho(t,x) = \frac{1}{t^{d\alpha}}\widetilde{\rho}\left(\ln t, \frac{x}{t^\alpha}\right), \quad \alpha = \frac{1}{d(m-1)+2}. \tag{1}$$

Then $\widetilde{\rho}$ is a solution to a non-linear Fokker–Planck type equation

$$\frac{\partial}{\partial t}\widetilde{\rho} = \Delta\widetilde{\rho}^m + \alpha \ \mathrm{div}\left[\widetilde{\rho}\left(\nabla\frac{1}{2}|x|^2\right)\right], \tag{NFPE_m}$$

where div stands for the adjoint operator of the gradient ∇. This non-linear Fokker–Planck equation NFPE_m has a stationary solution $\widetilde{\rho}^*_m$ given by

$$\widetilde{\rho}^*_m(y) = \begin{cases} [A - B|y|^2]^{\frac{1}{m-1}}, & \text{if } A - B|y|^2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $B = (m-1)\alpha/(2m)$. The other constant A is defined by the total mass of the solution. In our case, we normalized the total mass and a precise value of A is defined in (8). (A more detailed treatment can be found for instance in [14, Subsection 3.4].)

Otto moreover verified that NFPE_m can be regarded as a gradient flow of a functional F_m, given by

$$F_m(\mu) = E_m(\mu) + \frac{\alpha}{2}\int_{\mathbb{R}^d}|y|^2 d\mu(y).$$

This gradient structure derives the following asymptotic behaviors of the solutions $\widetilde{\rho}$ to NFPE_m:

$$\frac{d}{d\tau}\left[\exp(2\alpha\tau)|\mathrm{grad}F_{m|\widetilde{\rho}}|^2\right] \le 0, \tag{2}$$

$$\frac{d}{d\tau}\left[\exp(2\alpha\tau)(F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m))\right] \le 0, \tag{3}$$

$$\frac{d}{d\tau}\left[\exp(2\alpha\tau)W_2(\widetilde{\rho}, \widetilde{\rho}^*_m)^2\right] \le 0, \tag{4}$$

where $|\mathrm{grad}F_{m|\widetilde{\rho}}|^2$ is identified with a functional:

$$|\mathrm{grad}F_{m|\widetilde{\rho}}|^2 = \int_{\mathbb{R}^d}\left|\nabla\left(e'_m(\widetilde{\rho}) + \frac{\alpha}{2}|y|^2\right)\right|^2\widetilde{\rho}(y)dy.$$

As m tends to 1, $F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m)$ tends to the relative entropy $H(\widetilde{\rho}|\widetilde{\rho}^*_m)$ between $\widetilde{\rho}$ and $\widetilde{\rho}^*_m$, similarly, $|\mathrm{grad} F_{m|\widetilde{\rho}}|^2$ tends to the relative Fisher information $I(\widetilde{\rho}|\widetilde{\rho}^*_m)$ between $\widetilde{\rho}$ and $\widetilde{\rho}^*_m$. Here the relative entropy H and the relative Fisher information I are given by

$$H(\mu|\nu) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} \frac{d\mu}{d\nu} \ln \frac{d\mu}{d\nu} d\nu, & \text{if } \mu \text{ is absolutely continuous w.r.t } \nu \\ +\infty, & \text{otherwise} \end{cases}$$

$$I(\mu|\nu) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} \left|\nabla \ln \frac{d\mu}{d\nu}\right|^2 d\mu, & \text{if } \mu \text{ is absolutely continuous w.r.t. } \nu \\ +\infty, & \text{otherwise.} \end{cases}$$

When μ is absolutely continuous with respect to ν, the relative entropy H and the relative Fisher information I are expressed in terms of the free energy density e of the Boltzmann entropy as follows:

$$H(\mu|\nu) = \int_{\mathbb{R}^d} \left[e\left(\frac{d\mu}{dx}\right) - e\left(\frac{d\nu}{dx}\right) - e'\left(\frac{d\nu}{dx}\right)\left(\frac{d\mu}{dx} - \frac{d\nu}{dx}\right)\right] dx,$$

$$I(\mu|\nu) = \int_{\mathbb{R}^d} \left|\nabla \left[e'\left(\frac{d\mu}{dx}\right) - e'\left(\frac{d\nu}{dx}\right)\right]\right|^2 \frac{d\mu}{dx} dx.$$

From this point of view, it is natural to define functionals associated with the Tsallis entropy, called m-relative entropy H_m and m-relative Fisher information I_m as follows:

$$H_m(\mu|\nu) = \int_{\mathbb{R}^d} \left[e_m\left(\frac{d\mu}{dx}\right) - e_m\left(\frac{d\nu}{dx}\right) - e'_m\left(\frac{d\nu}{dx}\right)\left(\frac{d\mu}{dx} - \frac{d\nu}{dx}\right)\right] dx,$$

$$I_m(\mu|\nu) = \int_{\mathbb{R}^d} \left|\nabla \left[e'_m\left(\frac{d\mu}{dx}\right) - e'_m\left(\frac{d\nu}{dx}\right)\right]\right|^2 \frac{d\mu}{dx} dx.$$

Throughout the paper, we use the convention that $\infty \cdot 0 = 0$. Otto [14] showed a relation between $F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m)$ and $H_m(\widetilde{\rho}|\widetilde{\rho}^*_m)$:

$$F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m) \begin{cases} \geq H_m(\widetilde{\rho}|\widetilde{\rho}^*_m), & \text{if } m > 1 \\ = H_m(\widetilde{\rho}|\widetilde{\rho}^*_m), & \text{if } m \leq 1. \end{cases}$$

The functional $|\mathrm{grad} F_{m|\widetilde{\rho}}|^2$ coincides with $I_m(\widetilde{\rho}|\widetilde{\rho}^*_m)$:

$$|\mathrm{grad} F_{m|\widetilde{\rho}}|^2 = I_m(\widetilde{\rho}|\widetilde{\rho}^*_m).$$

The key ingredient for proving the asymptotic results is some "convexity" of E_m. This concept is called displacement convexity, introduced by McCann [10]. Otto derived the following inequalities which play crucial roles in the proof of asymptotic results from the convexity of E_m:

$$F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m) \leq \frac{1}{2\alpha} I_m(\widetilde{\rho}|\widetilde{\rho}^*_m),$$

$$W_2(\widetilde{\rho}, \widetilde{\rho}^*_m)^2 \leq \frac{2}{\alpha}\left(F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m)\right),$$

$$F_m(\widetilde{\rho}) - F_m(\widetilde{\rho}^*_m) \leq I_m(\widetilde{\rho}|\widetilde{\rho}^*_m) W_2(\widetilde{\rho}, \widetilde{\rho}^*_m).$$

More generally, the displacement convexity of E_m brings out the following inequalities:

$$H_m(\mu|\nu) \leq \frac{1}{2\lambda} I_m(\mu|\nu), \qquad \mathrm{LS}_m(\lambda)$$

$$W_2(\mu,\nu) \leq \sqrt{\frac{2}{\lambda} H_m(\mu|\nu)}, \qquad \mathrm{T}_m(\lambda)$$

$$H_m(\mu|\nu) \leq I_m(\mu|\nu) W_2(\mu,\nu) - \frac{2}{K} W_2(\mu,\nu)^2. \qquad \mathrm{HWI}_m(K)$$

Here K may take any values, however we assume that λ is positive. If we have equalities in $\mathrm{LS}_m(\lambda)$, $\mathrm{T}_m(\lambda)$ and $\mathrm{HWI}_m(K)$, then ν must be a q-Gaussian measure (see [1],[6],[7],[8],[17]).

When $m=1$, the inequality $\mathrm{LS}_m(\lambda)$ is called logarithmic Sobolev inequality and the inequality $\mathrm{T}_m(\lambda)$ is called Talagrand inequality. Equalities in the logarithmic Sobolev inequality and the Talagrand inequality hold if and only if μ is a translation of ν and ν is a Gaussian measure whose covariance matrix is a scalar matrix (see [4],[9] and also [15]).

A probability measure with mean v and covariance matrix V is a q-Gaussian measures $N_q(v,V)$ if it maximizes the Tsallis entropy E_q. The q-Gaussian measures are characterized by the q-exponential function $\exp_q$, which is given by

$$\exp_q(t) = \begin{cases} [1+(1-q)t]^{\frac{1}{1-q}}, & \text{if } 1+(1-q)t > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This function converges to the general exponential function exp when q converges to 1. For example, $\widetilde{\rho}_m^* dx$ is one of the q-Gaussian measures when $m+q=2$. When q tends to 1, $N_q(v,V)$ tends to the Gaussian measure $N(v,V)$ with mean v and covariance matrix V. (Note that Gaussian measures, which are characterized by the exponential function, maximize the Boltzmann entropy E.) We only treat the case of the parameter m and q satisfying that

$$m > d/(d+2), \quad m \geq (d-1)/d, \quad m < 2 \quad \text{and} \quad q = 2-m.$$

Ohara–Wada [13] showed that the space $\mathcal{N}(q,d)$ of q-Gaussian measures on $\mathbb{R}^d$ is invariant under PME_m for $1 < m < 2$. This fact implies that the solution to PME_m can be explicitly solved ([13, Remark 2]). We give an explicit expression of the solution to PME_m:

Theorem. *We assume that $m+q=2$ and $0 < q < (d+4)/(d+2)$. Let C be a positive constant defined in* (9). *For any $N_q(v, C\Theta)$ in $\mathcal{N}(q,d)$, we set a time dependent matrix Θ_t as*

$$\begin{cases} \Theta_t = \Theta + \sigma(t) I_d, \\ \dfrac{d}{dt}\sigma(t) = 2\alpha \left(\det \Theta_t\right)^{-\frac{1-q}{2}} \end{cases}$$

Then the density of $N_q(v, C\Theta_t)$ *is a solution to the porous medium equation* PME_m.

Here I_d is the unit matrix of size d. In the case of $m = 1$, this theorem corresponds to the well-known fact that a solution to the heat equation is obtained by a convolution of an initial data with the heat kernel:

$$N(v, \Theta_t) = N(v, \Theta + 2tI_d) = N(v, \Theta) * N(0, 2tI_d).$$

Due to the rescaling in (1), we also have an explicit expression of a solution to NFPE_m when an initial data is a q-Gaussian measure. Indeed, if a time dependent matrix Ξ_τ satisfies

$$\begin{cases} \Xi_1 = \Xi : \text{symmetric positive definite matrix}, \\ \dfrac{d}{d\tau}\Xi_\tau = -2\alpha\Xi_\tau + 2\alpha(\det \Xi_\tau)^{-\frac{1-q}{2}} I_d, \end{cases}$$

then the density of $N_q(0, C\Xi_\tau)$ is a solution to NFPE_m. In particular, the density of $N_q(0, CI_d)$ is a stationary solution to NFPE_m, that is,

$$N_q(0, CI_d) = \widetilde{\rho}^*_m dx.$$

This explicit expression of the solutions to NFPE_m helps us to understand the asymptotic behavior of the solutions to NFPE_m. In Section 3, we consider correspondences to the results in Otto [14].

Finally, we state the properties of H_m and I_m. These functionals are non-negative and they are equal to 0 if and only if $\mu = \nu$. Note that H_m is a β-divergence up to a multiplicative constant depending on m when $\beta = m-1$. (See [13] for properties of the β-divergence and references therein.) The β-divergence satisfies the Pythagorean relation. Namely, for an absolutely continuous measure μ, let ν_* be a minimizing q-Gaussian measure for the variational problem

$$\min_{\nu \in \mathcal{N}(q,d)} H_m(\mu|\nu).$$

Then the following Pythagorean relation holds for all ν in $\mathcal{N}(q, d)$:

$$H_m(\mu|\nu) = H_m(\mu|\nu_*) + H_m(\nu_*|\nu). \tag{5}$$

(See the books of Amari [2] and Amari–Nagaoka [3] for more information.) It means that the β-divergence is a generalized square of the distance function. Thus the inequality $\mathrm{T}_m(\lambda)$ means the comparison between the two "distance functions", the L^2-Wasserstein distance W_2 and the square root of the m-relative entropy H_m, not H_m itself. Speaking in broad terms, $-I_m(\mu_t|\nu)$ is a differential of $H_m(\mu_t|\nu)$ and $\mathrm{LS}_m(\lambda)$ means the convexity of H_m.

The organization of this paper is as follows. We review some preliminary materials in Section 2. We first introduce the generalized logarithmic function and the generalized exponential function, then we define the q-Gaussian measures. After reviewing the L^2-Wasserstein geometry, we discuss the q-Gaussian measures as the solutions to PME_m.

(The details can be found in [17].) Section 3 is devoted to the asymptotic behavior of the solution with the initial data in $\mathcal{N}(q,d)$ to NFPE$_m$. Especially, we show Otto's result (2)–(4) with elementally calculations.

Acknowledgement

The author would like to thank Sumio Yamada for his suggestions and encouragement. This work was partially supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

2. Preliminary

2.1. Generalized logarithmic function and Generalized exponential function. We first introduce generalized logarithmic functions and generalized exponential functions. See [11] and [12] for further information.

We fix a positive, strictly increasing function φ on $[0,\infty)$. We define a generalized logarithmic function $\ln_\varphi$ by

$$\ln_\varphi(t) = \int_1^t \frac{1}{\varphi(s)} ds.$$

Since $\ln_\varphi$ is a strictly increasing function, $\ln_\varphi$ has an inverse function, called generalized exponential function $\exp_\varphi$.

If μ is an absolutely continuous measure with respect to the Lebesgue measure dx, there exists a nonnegative Borel function f on $\mathbb{R}^d$ such that

$$\mu[A] = \int_A f d\nu$$

for all Borel sets A in $\mathbb{R}^d$. The function f is called density of μ and denoted by $d\mu/dx$.

For two absolutely continuous probability measures μ and ν in $\mathcal{P}^{\mathrm{ac}}$, we define a Bregman divergence by

$$D_\varphi(\mu|\nu) = \int_{\mathbb{R}^d} \left[F_\varphi\left(\frac{d\mu}{dx}\right) - F_\varphi\left(\frac{d\nu}{dx}\right) - \ln_\varphi\left(\frac{d\nu}{dx}\right)\left(\frac{d\mu}{dx} - \frac{d\nu}{dx}\right)\right] dx,$$

where the function F_φ on $[0,\infty)$ is given by

$$F_\varphi(\tau) = \int_1^\tau \ln_\varphi(t)dt.$$

We further assume that

$$F_\varphi(0) = \lim_{\tau \searrow 0} F_\varphi(\tau) < +\infty$$

in choosing φ. We note that the Bregman divergence satisfies the Pythagorean relation (5).

In the special case of $\varphi(u) = u^q$ $(0 < q < 2,\ q \neq 1)$, $\ln_\varphi$ and $\exp_\varphi$ are particularly called q-logarithmic function and q-exponential function, denoted by $\ln_q$ and $\exp_q$, respectively:

$$\ln_q(t) = \frac{t^{1-q} - 1}{1-q},$$
$$\exp_q(t) = \begin{cases} [1 + (1-q)t]^{\frac{1}{1-q}}, & \text{if } 1 + (1-q)t > 0 \\ 0, & \text{otherwise.} \end{cases}$$

They converge to the natural logarithmic function and the natural exponential function when q converges to 1. In this case, the corresponding Bregman divergence is called the β-divergence.

2.2. **q-Gaussian.** Let us summarize the definition of q-Gaussian measures. Background information on Tsallis entropy and q-Gaussian measure is in Tsallis' book [19].

Let $\mathrm{Sym}^+(d, \mathbb{R})$ be the set of symmetric positive definite matrices of size d. The maximum entropy principle for the Tsallis entropy E_q under the mean constraint v in $\mathbb{R}^d$ and the covariance constraint V in $\mathrm{Sym}^+(d, \mathbb{R})$

$$\begin{cases} \displaystyle\int_{\mathbb{R}^d} v\rho(x)dx = v \\ \displaystyle\int_{\mathbb{R}^d} (x-v)^{\mathsf{T}}(x-v)\rho(x)dx = V \end{cases}$$

yields the q-Gaussian measure $N_q(v, V)$

$$\frac{dN_q(v,V)}{dx}(x) = C_0 (\det V)^{-\frac{1}{2}} \exp_q \left[-\frac{1}{2} C_1 \left\langle x - v, V^{-1}(x-v) \right\rangle \right].$$

Here vectors in $\mathbb{R}^d$ are column and ${}^{\mathsf{T}}x$ stands for the transpose of x. Moreover, C_0 and C_1 are the positive constants given by

$$C_0 = C_0(q,d) = \begin{cases} \dfrac{\Gamma(\frac{1}{q-1})}{\Gamma(\frac{1}{q-1} - \frac{d}{2})} \left(\dfrac{(q-1)C_1}{2\pi} \right)^{\frac{d}{2}}, & \text{if } q > 1 \\ \dfrac{\Gamma(\frac{2-q}{1-q} + \frac{d}{2})}{\Gamma(\frac{2-q}{1-q})} \left(\dfrac{(1-q)C_1}{2\pi} \right)^{\frac{d}{2}}, & \text{if } 0 < q < 1 \end{cases}$$
$$C_1 = C_1(q,d) = \frac{2}{2 + (d+2)(1-q)}$$

and $\Gamma(\cdot)$ is the Γ-function. For $0 < q < (d+4)/(d+2)$ and $q \neq 1$, the q-Gaussian measure is well-defined. As q tends to 1, $N_q(v, V)$ tends to the Gaussian measure $N(v, V)$. We denote the densities $dN_q(v, V)/dx$ and $dN(v, V)/dx$ by $N_q(v, V)(\cdot)$ and $N(v, V)(\cdot)$, respectively.

2.3. **L^2-Wasserstein space.** We discuss the L^2-Wasserstein geometry. It is a pair of the subset of probability measures on a complete, separable metric space and a distance function W_2 derived from the Monge–Kantorovich transport problem. The convergence in the sense of W_2 is somewhat stronger than the weak convergence. For simplicity, we consider only the case that the underlying metric space is the standard Euclidean normed space $(\mathbb{R}^d, |\cdot|)$. See [20] and [21] for the general theory.

The set of all Borel probability measures μ on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x) < \infty$$

will be denoted by $\mathcal{P}_2$. A transport plan π between μ and ν in $\mathcal{P}_2$ is a Borel probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν, that is,

$$\pi[M \times \mathbb{R}^d] = \mu[M], \quad \pi[M \times \mathbb{R}^d] = \nu[M]$$

for all measurable sets M in $\mathbb{R}^d$. The L^2-Wasserstein distance between μ and ν in $\mathcal{P}_2$ is defined by

$$W_2(\mu, \nu) = \left(\inf_{\pi} \int_{M\times M} |x - y|^2 d\pi(x, y) \right)^{\frac{1}{2}},$$

where the infimum is taken over all the transport plans π between μ and ν. Then W_2 is a distance function on $\mathcal{P}_2$. We call the pair $(\mathcal{P}_2, W_2)$ L^2-Wasserstein space.

For a symmetric positive definite matrix X, we define a symmetric positive definite matrix $X^{1/2} = \sqrt{X}$ so that $X^{1/2} \cdot X^{1/2} = X$. The author [17] showed that the L^2-Wasserstein distance between $N_q(v, V)$ and $N_q(u, U)$ is given by

$$W_2(N_q(v, V), N_q(u, U))^2 = |v - u|^2 + \mathrm{tr}V + \mathrm{tr}U - 2\mathrm{tr}\sqrt{U^{\frac{1}{2}} V U^{\frac{1}{2}}}. \quad (6)$$

2.4. **q-Gaussian measure as solution to porous medium equation.** It is well-known that the porous medium equation PME_m allows for a self-similar solution of form

$$\rho_m(x, t) = [At^{-d\alpha(m-1)} - B|x|^2 t^{-1}]_+^{\frac{1}{m-1}} = [A - B|x|^2 t^{-2\alpha}]_+^{\frac{1}{m-1}} t^{-d\alpha},$$

where the constant α and B are given by

$$\alpha = \alpha(m, d) = \frac{1}{d(m-1)+2}, \quad B = B(m, d) = \frac{(m-1)\alpha}{2m}. \quad (7)$$

The other constant $A = A(m, d)$ is defined by the total mass of the solution and we normalized it such that

$$\int_{\mathbb{R}^d} \rho_m(x, t) dx = 1.$$

Precisely, A is given by

$$A^{\frac{1}{2\alpha(m-1)}} = \begin{cases} \dfrac{\Gamma(\frac{m}{m-1}+\frac{d}{2})}{\Gamma(\frac{m}{m-1})}\left(\dfrac{B}{\pi}\right)^{\frac{d}{2}}, & \text{if } m > 1 \\ \dfrac{\Gamma(\frac{1}{1-m})}{\Gamma(\frac{1}{1-m}-\frac{d}{2})}\left(\dfrac{-B}{\pi}\right)^{\frac{d}{2}}, & \text{if } m < 1. \end{cases} \tag{8}$$

This solution was discovered by Barenblatt [5], Pattle [16] and called Barenblatt–Pattle solution. When m tends to 1, we have the following behaviors:

$$A \longrightarrow 1, \quad B \longrightarrow 0, \quad \alpha \longrightarrow 1/2,$$
$$\rho_m(x,t) \longrightarrow (4\pi t)^{-\frac{d}{2}} \exp\left(-\frac{|x|^2}{4t}\right) = N(0, 2tI_d)(x).$$

Namely, the Barenblatt–Pattle solution approaches the heat kernel.

The rest of this section is devoted to study the relation between the q-Gaussian measures and the Barenblatt–Pattle solution. For simplicity, we use the following notations for V in $\mathrm{Sym}^+(d,\mathbb{R})$ and t in $\mathbb{R}$:

$$|x|_V = \sqrt{\langle x, V^{-1}x\rangle}, \quad \Theta(V) = (\det V)^{-\alpha(1-q)}V, \quad [t]_+ = \max\{t, 0\}.$$

Let $\mathcal{M}(q,d)$ be the subset of $\mathcal{P}^{\mathrm{ac}}$ defined by

$$\left\{ \begin{array}{l} \rho_m(v,V)(x) = \left[A(\det V)^{-\alpha(1-q)} - B|x-v|_V^2\right]_+^{\frac{1}{1-q}} \\ \quad = \left[A - B|x-v|_{\Theta(V)}^2\right]_+^{\frac{1}{1-q}}(\det \Theta(V))^{-\frac{1}{2}} \end{array} \middle| \begin{array}{l} v \in \mathbb{R}^d, \\ V \in \mathrm{Sym}^+(d,\mathbb{R}) \end{array} \right\},$$

where the constants A, B, α are defined in (7),(8) and $m+q=2$. Then $\rho_m(0, tI_d)(\cdot)$ coincides with the Barenblatt–Pattle solution $\rho_m(\cdot, t)$. The assumption that

$$m > 1 - \frac{2}{d+2} \quad \text{or equivalently} \quad q < \frac{d+4}{d+2}$$

guarantees that $\mathcal{M}(q,d)$ is embedded into in $\mathcal{P}_2$. Actually, we obtain

$$\rho_m(v,V)(x) = N_q(v, C\Theta(V))(x),$$

where the constant $C = C(q,d)$ is given by

$$C = \frac{2(2-q)A}{1+2\alpha(1-q)}. \tag{9}$$

Therefore the set $\mathcal{M}(q,d)$ can be identified with the set $\mathcal{N}(q,d)$ for any q satisfying $0 < q < (d+4)/(d+2)$.

One of the remarkable properties of the Gaussian measures is that $\mathcal{N}(d)$ is invariant under the heat equation. Namely, the solution to the heat equation with the initial data $N(v,V)$ stays in $\mathcal{N}(d)$ for all future time. Because a solution to the heat equation is obtained by a convolution of an initial data with the heat kernel. When the initial data is a Gaussian measure, we additionally have the explicit expression of the solution to the heat equation:

$$N(v,V) * N(0,2tI_d) = N(v, V+2tI_d) \in \mathcal{N}(d).$$

Ohara–Wada [13, Proposition 5] demonstrated an analogy for the porous medium equation PME_m, which states that a solution to PME_m with an initial data in $\mathcal{M}(q,d)$ belongs to $\mathcal{M}(q,d)$ for $t > 0$. This fact implies that the solution to PME_m can be explicitly solved ([13, Remark 2]). We give the explicit expression of the solution to PME_m.

Theorem. *We assume that $m+q=2$ and $0<q<(d+4)/(d+2)$. For any $\rho_m(v,V)$ in $\mathcal{M}(q,d)$, we set a time dependent matrix V_t as*

$$\begin{cases} \Theta(V_t) = \Theta(V) + \sigma(t) I_d, \\ \dfrac{d}{dt}\sigma(t) = 2\alpha \left(\det \Theta(V_t)\right)^{\frac{1-m}{2}}. \end{cases} \tag{10}$$

Then $\rho_m(v,V_t)$ is a solution to the porous medium equation PME_m.

Remark . Note that $\Theta(V_t)$ is regarded as Θ_t in the introduction.

Proof. Since V is a symmetric positive definite matrix, so are V_t and $\Theta(V_t)$ for all time $t>0$. We set

$$\Theta_t = \Theta(V_t), \quad F(t,x) = \left[A - B|x-v|^2_{\Theta_t}\right]_+, \quad D(t) = \det \Theta_t,$$

then $\rho_m(v,V_t)$ is expressed by

$$\rho_m(v,V_t)(x) = F(t,x)^{\frac{1}{m-1}} D(t)^{-\frac{1}{2}}.$$

Note that $D(t)$ is positive for all $t>0$. In the case of $m>1$, $F(t,x)$ is positive for all x in $\mathbb{R}^d$ and $t>0$. Thus any power of F is well-defined. In the case of $m<1$, $F(t,x)$ may become zero for some x in $\mathbb{R}^d$ and $t>0$. However, all parameters which appear in the exponents of F as below are positive. Therefore we can justify the following calculations.

We first consider the differential of $D(t)$. For any time dependent invertible matrix X_t, we know the following result:

$$\frac{d}{dt}\det(X_t) = (\det X_t)\operatorname{tr}\left(X_t^{-1}\frac{d}{dt}X_t\right).$$

Combining this fact with the assumption

$$\frac{d}{dt}\Theta_t = 2\alpha D(t)^{\frac{1-m}{2}} I_d,$$

we obtain

$$\frac{d}{dt}D(t)^{-\frac{1}{2}} = -\frac{1}{2}D(t)^{-\frac{3}{2}}\frac{d}{dt}D(t) = -\alpha D^{-\frac{m}{2}}(t)\operatorname{tr}\left(\Theta_t^{-1}\right).$$

We next compute the differential of $F(t,x)$ with respect to t. The following result concerning a time dependent invertible matrix X_t

$$\frac{d}{dt}X_t^{-1} = -X_t^{-1}\left(\frac{d}{dt}X_t\right)X_t^{-1}$$

yields

$$\frac{\partial}{\partial t}F(t,x) = -B\left\langle x-v, \left(\frac{d}{dt}\Theta_t^{-1}\right)(x-v)\right\rangle = 2\alpha B D^{\frac{1-m}{2}}(t)|x-v|^2_{\Theta_t^2}.$$

Then we acquire

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\rho_m(v,V_t)(x)\right)\\
&= \frac{\partial}{\partial t}\left(F(t,x)^{\frac{1}{m-1}}D(t)^{-\frac{1}{2}}\right)\\
&= \left(\frac{\partial}{\partial t}F(t,x)^{\frac{1}{m-1}}\right)D(t)^{-\frac{1}{2}} + F(t,x)^{\frac{1}{m-1}}\frac{\partial}{\partial t}D(t)^{-\frac{1}{2}}\\
&= \frac{1}{m-1}\left(2\alpha B D^{\frac{1-m}{2}}(t)|x-v|_{\Theta_t^2}\right)F(t,x)^{\frac{1}{m-1}-1}D(t)^{-\frac{1}{2}}\\
&\quad + F(t,x)^{\frac{1}{m-1}}\left(-\alpha D^{-\frac{m}{2}}(t)\mathrm{tr}\left(\Theta_t^{-1}\right)\right)\\
&= \alpha F^{\frac{1}{m-1}-1}(t,x)D(t)^{-\frac{m}{2}}\left(\frac{2B}{m-1}|x-v|^2_{\Theta_t^2} - F(t,x)\mathrm{tr}\left(\Theta_t^{-1}\right)\right).
\end{aligned}$$

By the direct computation, we have the gradient of F:

$$\nabla F(t,x) = -2B\Theta_t^{-1}(x-v).$$

Therefore, the gradient of $\rho_m(v,V_t)^m$ is as follows:

$$\begin{aligned}
\nabla\left(\rho_m(v,V_t)^m(x)\right) &= \nabla\left(F(t,x)^{\frac{m}{m-1}}D(t)^{-\frac{m}{2}}\right)\\
&= \frac{m}{m-1}F(t,x)^{\frac{m}{m-1}-1}D(t)^{-\frac{m}{2}}\nabla F(t,x).
\end{aligned}$$

The Laplacian of $\rho_m(v,V_t)^m$ is obtained by taking the divergence of the above equation:

$$\begin{aligned}
&\Delta\left(\rho_m(v,V_t)^m(x)\right)\\
&= \mathrm{div}\left(\frac{m}{m-1}F(t,x)^{\frac{m}{m-1}-1}D(t)^{-\frac{m}{2}}\nabla F(t,x)\right)\\
&= \left\langle\nabla\left(\frac{m}{m-1}F(t,x)^{\frac{m}{m-1}-1}\right), D(t)^{-\frac{m}{2}}\nabla F(t,x)\right\rangle\\
&\quad + \frac{m}{m-1}F(t,x)^{\frac{m}{m-1}-1}D(t)^{-\frac{m}{2}}\,\mathrm{div}\nabla F(t,x)\\
&= \frac{m}{m-1}\frac{1}{m-1}F(t,x)^{\frac{m}{m-1}-2}D(t)^{-\frac{m}{2}}\,|\nabla F(t,x)|^2\\
&\quad - 2B\frac{m}{m-1}F(t,x)^{\frac{m}{m-1}-1}D(t)^{-\frac{m}{2}}\mathrm{tr}(\Theta_t^{-1})\\
&= \alpha F(t,x)^{\frac{1}{m-1}-1}D(t)^{-\frac{m}{2}}\left(\frac{2B}{m-1}|x-v|^2_{\Theta_t^2} - F(t,x)\mathrm{tr}(\Theta_t^{-1})\right).
\end{aligned}$$

Hence we have

$$\frac{\partial}{\partial t}\rho_m(v,V_t)(x) = \Delta\left(\rho_m(v,V_t)^m(x)\right),$$

proving that $\rho_m(v, V_t)$ is the solution to PME_m. $\square$

3. Otto's calculations

Since the previous theorem ensures that mean of the solution to PME_m dose not depend on the time t, we fix mean $v = 0$ and denote $\rho_m(0, X)$ by $\rho_m(X)$ and $N_q(0, Y)$ by $N_q(Y)$. We define a time dependent matrix V_t as in (10), that is, $\rho_m(V_t)$ is a solution to PME_m. Due to the rescaling given by (1),

$$\widetilde{\rho}\left(\ln t, \frac{x}{t^\alpha}\right) = t^{d\alpha}\rho_m(V_t)(x) = \left[A - B|x|^2_{\Theta_t}\right]_+^{\frac{1}{m-1}} (\det \Theta_t)^{-\frac{1}{2}}$$
$$= \left[A - B\left|\frac{x}{t^\alpha}\right|^2_{t^{-2\alpha}\Theta_t}\right]_+^{\frac{1}{m-1}} (\det t^{-2\alpha}\Theta_t)^{-\frac{1}{2}}$$

is a solution to NFPE_m. Setting

$$U_\tau = \frac{1}{t}V_t = e^{-\tau}V_{e^\tau}, \quad \Xi_\tau = \Theta(U_\tau) = \Theta\left(\frac{1}{t}V_t\right) = t^{-2\alpha}\Theta_t = e^{-2\alpha\tau}\Theta_t,$$

the solution to NFPE_m is expressed by

$$\widetilde{\rho}(\tau, y) = \left[A - B\,|y|^2_{\Xi_\tau}\right]_+^{\frac{1}{m-1}} (\det \Xi_\tau)^{-\frac{1}{2}} = \rho_m(U_\tau)(y).$$

Relations of determinants and traces between Ξ_τ and U_τ are as follows:

$$\det \Xi_\tau = (\det U_\tau)^\alpha, \quad \mathrm{tr}\Xi_\tau = (\det U_\tau)^{\alpha(1-m)}\mathrm{tr}U_\tau.$$

Moreover, if $\rho_m(U_\tau)$, that is the density of $N_q(\Xi_\tau)$, is the solution to NFPE_m, then the time dependent matrix U_τ satisfies

$$\begin{cases} \Xi_1 = \Xi : \text{ symmetric positive definite matrix,} \\ \dfrac{d}{d\tau}\Xi_\tau = -2\alpha\Xi_\tau + 2\alpha(\det \Xi_\tau)^{-\frac{1-q}{2}} I_d. \end{cases}$$

In what follow, let X be a symmetric positive matrix and $Y = \Theta(X)$. When $m \geq 1$, the support of $N_q(CY)$ is $\mathbb{R}^d$ and relations between Otto's notations (right hand sides) and ours (left hand sides) are given as follows:

$$\begin{aligned} N_q(CE) &= \widetilde{\rho}^*_m, \\ H_m(N_q(CY)|N_q(CE)) &= H(N_q(CY)|N_q(CE)), \\ &= F(N_q(CY)) - F(N_q(CE)), \\ I_m(N_q(CY)|N_q(CE)) &= |\mathrm{grad}F_{|N_q(CY)}|^2. \end{aligned}$$

Otto [14] proved inequalities including $\mathrm{LS}_m(\lambda)$, $\mathrm{T}_m(\lambda)$ and the weakened version of $\mathrm{HWI}_m(K)$ for solutions to NFPE_m and then showed the asymptotic results (2)–(4) using these inequalities. However, we can show the asymptotic results without $\mathrm{T}_m(\lambda)$, $\mathrm{HWI}_m(K)$ when the initial data of the solution to NFPE_m is the q-Gaussian measure. Moreover, we prove these inequalities using only linear algebra when μ, ν are q-Gaussian measures. In the rest of this paper, we assume $m > 1$.

For the solution $N_q(C\Xi_\tau)$ to NFPE_m, we set

$$\begin{aligned}
W_2(\tau) &= W_2(N_q(C\Xi_\tau), N_q(CI_d)) = W_2(\rho_m(U_\tau), \rho_m(I_d)),\\
I_m(\tau) &= I_m(N_q(C\Xi_\tau)|N_q(CI_d)) = I_m(\rho_m(U_\tau)|\rho_m(I_d)),\\
H_m(\tau) &= H_m(N_q(C\Xi_\tau)|N_q(CI_d)) = H_m(\rho_m(U_\tau)|\rho_m(I_d)).
\end{aligned}$$

Applying (6), we get the value of $W_2(\tau)$:

$$\begin{aligned}
W_2(\tau)^2 &= C\cdot \mathrm{tr}\left[(\Xi_\tau^{\frac{1}{2}} - I_d)^2\right] = C\cdot \mathrm{tr}(\Xi_\tau - 2\Xi_\tau^{\frac{1}{2}} + I_d)\\
&= C\cdot \mathrm{tr}[(\det U_\tau)^{\alpha(1-m)}U_\tau - 2(\det U_\tau)^{\frac{\alpha(1-m)}{2}}U_\tau^{\frac{1}{2}} + I_d].
\end{aligned}$$

The definition of covariance matrix asserts that

$$\int_{\mathbb{R}^d} \langle Z_1x, Z_2x\rangle \, N_q(CY)(dx) = C\mathrm{tr}(Z_1Y^\top Z_2)$$

for any matrices Z_1 and Z_2. By a direct computation, we have

$$\nabla e_m'(\rho_m(X)(x)) = \nabla\frac{m}{m-1}(\rho_m^{m-1}(X)(x) - 1) = -\alpha(\det Y)^{\frac{1-m}{2}}Y^{-1}x.$$

Therefore we acquire the value of $I_m(\tau)$:

$$\begin{aligned}
I_m(\tau) &= \int_{\mathbb{R}^d} |\nabla[e_m'(\rho_m(U_\tau)(x)) - e_m'(\rho_m(I_d)(x))]|^2\, \rho_m(U_\tau)(x)dx\\
&= \int_{\mathbb{R}^d} \alpha^2 \left|(\det\Xi_\tau)^{\frac{1-m}{2}}\Xi_\tau^{-1}x - x\right|^2 N_q(C\Xi_\tau)(dx)\\
&= C\alpha^2\left[(\det\Xi_\tau)^{1-m}\mathrm{tr}\Xi_\tau^{-1} + \mathrm{tr}\Xi_\tau - 2d(\det\Xi_\tau)^{\frac{1-m}{2}}\right]\\
&= C\alpha^2(\det U_\tau)^{\alpha(1-m)}\left(\mathrm{tr}U_\tau + \mathrm{tr}U_\tau^{-1} - 2d\right).
\end{aligned}$$

We next consider $H_m(\tau)$. By straightforward calculations, we get

$$\begin{aligned}
\int_{\mathbb{R}^d}\rho_m(X)(x)dx &= \int_{\mathbb{R}^d}[A - B|x|_Y^2]^{\frac{m}{m-1}}(\det Y)^{-\frac{m}{2}}dx\\
&= (\det Y)^{\frac{1-m}{2}}\int_{\mathbb{R}^d} A^{\frac{m}{m-1}}\left[1 - \frac{B}{A}|x|^2\right]^{\frac{m}{m-1}}dx\\
&= (\det Y)^{\frac{1-m}{2}}A^{\frac{m}{m-1}}\left(\frac{A}{-B}\right)^{\frac{d}{2}}\int_0^\infty(1+r^2)^{\frac{m}{m-1}}r^{d-1}dr\frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}\\
&= (\det Y)^{\frac{1-m}{2}}A^{\frac{1}{2\alpha(m-1)}+1}|-B|^{\frac{d}{2}}\frac{\Gamma\left(\frac{d}{2}\right)\Gamma\left(\frac{m}{1-m}-\frac{d}{2}\right)}{2\Gamma\left(\frac{m}{1-m}\right)}\frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}\\
&= (\det Y)^{\frac{1-m}{2}}C\alpha.
\end{aligned}$$

Hence we have $H_m(\tau)$:

$$\begin{aligned}
H_m(\tau) &= \int_{\mathbb{R}^d} \left[e_m(\rho_m(U_\tau)) - e_m(\rho_m(I_d)) - e'_m(\rho_m(I_d))(\rho_m(U_\tau) - \rho_m(I_d))\right] dx \\
&= \frac{1}{m-1}\int_{\mathbb{R}^d} \left[\rho_m(U_\tau)^m - \rho_m(I_d)^m + mB|x|^2 \left(N_q(C\Xi_\tau) - N_q(CI_d)\right)\right] dx \\
&= \frac{C\alpha}{2}\left[\frac{2}{m-1}(\det \Xi_\tau)^{\frac{1-m}{2}} + \operatorname{tr}\Xi_\tau - \left(\frac{2}{m-1} + d\right)\right] \\
&= \frac{C\alpha}{2}\left[(\det U_\tau)^{\alpha(1-m)}\left(\frac{2}{m-1} + \operatorname{tr}U_\tau\right) - \left(\frac{2}{m-1} + d\right)\right].
\end{aligned}$$

We give a brief sketch of the above calculations:

$$\begin{aligned}
H_m(\tau) &= \frac{C\alpha}{2}\left[\frac{2}{m-1}(\det \Xi_\tau)^{\frac{1-m}{2}} + \operatorname{tr}\Xi_\tau - \left(\frac{2}{m-1} + d\right)\right], \\
&= \frac{C\alpha}{2}\left[(\det U_\tau)^{\alpha(1-m)}\left(\frac{2}{m-1} + \operatorname{tr}U_\tau\right) - \left(\frac{2}{m-1} + d\right)\right], \\
I_m(\tau) &= C\alpha^2\left[(\det \Xi_\tau)^{1-m}\operatorname{tr}\Xi_\tau^{-1} + \operatorname{tr}\Xi_\tau - 2d(\det \Xi_\tau)^{\frac{1-m}{2}}\right], \\
&= C\alpha^2(\det U_\tau)^{\alpha(1-m)}\left(\operatorname{tr}U_\tau + \operatorname{tr}U_\tau^{-1} - 2d\right), \\
W_2(\tau)^2 &= C(\operatorname{tr}\Xi_\tau + d - 2\operatorname{tr}\Xi_\tau^{\frac{1}{2}}), \\
&= C[(\det U_\tau)^{\alpha(1-m)}\operatorname{tr}U_\tau + d - 2(\det U_\tau)^{\frac{\alpha(1-m)}{2}}\operatorname{tr}U_\tau^{\frac{1}{2}}].
\end{aligned}$$

We first prove that the inequalities $\mathrm{LS}_m(\lambda)$, $\mathrm{T}_m(\lambda)$ and $\mathrm{HWI}_m(K)$ hold when $\nu = N_q(CI_d)$, $\lambda = K = \alpha > 0$ and μ is a solution to NFPE_m. In the proof, we use the characteristic of the q-Gaussian measures, not the solutions to NFPE_m. Hence we extend the inequalities to the case that μ is a q-Gaussian measure.

Define a function φ on $[0,\infty)$ by

$$\varphi(t) = \frac{2}{m-1}t^{-\alpha(1-m)} - \left(\frac{2}{m-1} + d\right) - t^{-\alpha(1-m)}d\left(t^{\frac{1}{d}} - 2\right).$$

Then the assumption that $m \le (d-1)/d$ guarantees that φ takes a maximum value $\varphi(1) = 0$ at $t = 1$. It implies that

$$\begin{aligned}
H_m(\tau) &= \frac{C\alpha}{2}\left[(\det U_\tau)^{\alpha(1-m)}\left(\frac{2}{m-1} + \operatorname{tr}U_\tau\right) - \left(\frac{2}{m-1} + d\right)\right] \\
&\le \frac{C\alpha}{2}(\det U_\tau)^{\alpha(1-m)}\left(d(\det U_\tau)^{-\frac{1}{d}} - 2d + \operatorname{tr}U_\tau\right) \\
&\le \frac{C\alpha}{2}(\det U_\tau)^{\alpha(1-m)}\left(\operatorname{tr}U_\tau^{-1} + \operatorname{tr}U_\tau - 2d\right) \\
&= \frac{1}{2\alpha}I_m(\tau),
\end{aligned}$$

proving $\mathrm{LS}_m(\lambda)$. Here the last inequality follows from the arithmetic geometric mean inequality. That is to say, for a symmetric positive definite matrix X of size d, we obtain

$$d(\det X)^{1/d} \leq \mathrm{tr} X.$$

Applying the arithmetic geometric mean inequality again, we get

$$\begin{aligned} W_2(\tau)^2 &= C[(\det U_\tau)^{\alpha(1-m)}\mathrm{tr}U_\tau + d - 2(\det U_\tau)^{\frac{\alpha(1-m)}{2}}\mathrm{tr}U_\tau^{\frac{1}{2}}] \\ &\leq C[(\det U_\tau)^{\alpha(1-m)}\mathrm{tr}U_\tau + d - 2d(\det U_\tau)^{\frac{\alpha(1-m)}{2}+\frac{1}{2d}}] \\ &= C[(\det U_\tau)^{\alpha(1-m)}\mathrm{tr}U_\tau + 2d(1-(\det U_\tau)^{\frac{\alpha}{d}}) - d]. \end{aligned} \tag{11}$$

For any positive number a, a function $\psi(t) = a^t$ is convex. Then the assumption that $0 < (1-m) \leq 1/d$ guarantees that

$$\frac{\psi(1-m)-1}{1-m} = \frac{\psi(1-m)-\psi(0)}{1-m} \leq \frac{\psi(1/d)-\psi(0)}{1/d} = d(\psi(1/d)-1).$$

Setting $a = (\det U_\tau)^\alpha$ and substituting it into (11), we obtain

$$\begin{aligned} W_2(\tau)^2 &\leq C\left[(\det U_\tau)^{\alpha(1-m)}\mathrm{tr}U_\tau + \frac{2}{1-m}(1-(\det U_\tau)^{\alpha(1-m)}) - d\right] \\ &= \frac{2}{\alpha}H_m(\tau). \end{aligned}$$

Thus we conclude $\mathrm{T}_m(\lambda)$.

We now prove $\mathrm{HWI}_m(K)$. Setting symmetric matrices W and I as

$$W = \Xi_\tau^{\frac{1}{2}} - I_d, \quad I = \Xi_\tau^{\frac{1}{2}} - (\det \Xi_\tau)^{\frac{1-m}{2}}\Xi_\tau^{-\frac{1}{2}},$$

we acquire the following relations:

$$W_2(\tau)^2 = \frac{C\alpha}{2}\mathrm{tr}(W^{\mathsf{T}}W), \quad I_m(\tau) = C\mathrm{tr}(I^{\mathsf{T}}I).$$

Define G by

$$G(Z_1, Z_2) = \mathrm{tr}(Z_1^{\mathsf{T}}Z_2)$$

for all square matrices Z_1 and Z_2 of size d, then G is an inner product on the space of all square matrices of size d. Then we obtain

$$\begin{aligned} \frac{1}{C\alpha}\sqrt{I_m}W_2 &= \sqrt{G(I,I)G(W,W)} \\ &\geq G(I,W) \\ &= \mathrm{tr}\Xi_\tau - \mathrm{tr}\Xi_\tau^{\frac{1}{2}} + (\det\Xi_\tau)^{\frac{1-m}{2}}(\mathrm{tr}\Xi_\tau^{-\frac{1}{2}} - d) \\ &\geq \mathrm{tr}\Xi_\tau - \mathrm{tr}\Xi_\tau^{\frac{1}{2}} + (\det\Xi_\tau)^{\frac{1-m}{2}}d((\det\Xi_\tau)^{-\frac{1}{2d}} - 1) \\ &\geq \mathrm{tr}\Xi_\tau - \mathrm{tr}\Xi_\tau^{\frac{1}{2}} + \frac{(\det\Xi_\tau)^{\frac{1-m}{2}}}{1-m}((\det\Xi_\tau)^{-\frac{1-m}{2}} - 1) \\ &= \frac{1}{C\alpha}\left(H_m(\tau) + \frac{\alpha}{2}W_2(\tau)^2\right), \end{aligned}$$

proving $\mathrm{HWI}_m(K)$. We apply the Cauchy–Schwarz inequality in the first inequality, and apply the arithmetic geometric mean inequality in the second inequality. The third inequality follows from the convexity of the function $\psi(t) = a^t$.

We are now going to prove the asymptotic results (2)–(4). We have the following inequalities, which prove (2).

$$
\begin{aligned}
&\frac{(\det \Xi_\tau)^{\frac{m-1}{2}}}{2C\alpha^3}\left[\frac{d}{d\tau}I_m(\tau) + 2\alpha I_m(\tau)\right] \\
&= (d(1-m)-1)\mathrm{tr}\left[(\det \Xi_\tau)^{\frac{1-m}{2}}\Xi_\tau^{-1} - I_d\right]^2 \\
&\quad + (1-m)(\det \Xi_\tau)^{1-m}\left[(\mathrm{tr}\Xi_\tau^{-1})^2 - d\mathrm{tr}\Xi_\tau^{-2}\right] \\
&\le (d(1-m)-1)\mathrm{tr}\left[(\det \Xi_\tau)^{\frac{1-m}{2}}\Xi_\tau^{-1} - I_d\right]^2 \\
&\le 0,
\end{aligned}
$$

where we apply the Cauchy–Schwarz inequality in the first inequality. The second inequality follows from the assumption $m \ge (d-1)/d$ and the positivity of the inner product G.

We show (3) using $\mathrm{LS}_m(\lambda)$:

$$
\begin{aligned}
\frac{d}{d\tau}H_m(\tau) &= \frac{C\alpha}{2}\left[-(\det \Xi_\tau)^{\frac{1-m}{2}}\mathrm{tr}\left(\Xi_\tau^{-1}\frac{d}{d\tau}\Xi_\tau\right) + \mathrm{tr}\left(\frac{d}{d\tau}\Xi_\tau\right)\right] \\
&= -I_m(\tau) \\
&\le -2\alpha H_m(\tau).
\end{aligned}
$$

Finally, we prove (4). We have the following inequalities:

$$
\begin{aligned}
&\frac{1}{2C\alpha}\left[\frac{d}{d\tau}W_2(\tau)^2 + 2\alpha W_2(\tau)^2\right] \\
&= \mathrm{tr}\left(\Xi_\tau^{\frac{1}{2}} - I_d\right)\left[(\det \Xi_\tau)^{\frac{1-m}{2}}\Xi_\tau^{-\frac{1}{2}} - I_d\right] \\
&= d(\det \Xi_\tau)^{\frac{1-m}{2}} + d - \mathrm{tr}\Xi_\tau^{\frac{1}{2}} - (\det \Xi_\tau)^{\frac{1-m}{2}}\mathrm{tr}\Xi_\tau^{-\frac{1}{2}} \\
&\le d\left[(\det \Xi_\tau)^{\frac{1-m}{2}}\left(1 - (\det \Xi_\tau)^{-\frac{1}{2d}}\right) + \left(1 - (\det \Xi_\tau)^{\frac{1}{2d}}\right)\right] \\
&= d\left(1 - (\det \Xi_\tau)^{-\frac{1}{2d}}\right)\left((\det \Xi_\tau)^{\frac{1-m}{2}} - (\det \Xi_\tau)^{\frac{1}{2d}}\right).
\end{aligned}
$$

The inequality follows from the arithmetic geometric mean inequality. The assumption $-1/d < 0 < (1-m) \le 1/d$ implies that

$$\left(1 - a^{-\frac{1}{2d}}\right)\left(a^{\frac{1-m}{2}} - a^{\frac{1}{2d}}\right)$$

is non-negative for a positive number a. Setting $a = \det \Xi_\tau$, we have (4).

References

1. M. Agueh, N. Ghoussoub, and X. Kang, *Geometric inequalities via a general comparison principle for interacting gases*, Geom. Funct. Anal. **14** (2004), no. 1, 215–244.

2. S. Amari, *Differential-geometrical methods in statistics*, Lecture Notes in Statistics, vol. 28, Springer-Verlag, New York, 1985.
3. S. Amari and H. Nagaoka, *Methods of information geometry*, Translations of Mathematical Monographs, vol. 191.
4. D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206.
5. G. I. Barenblatt, *On self-similar motions of a compressible fluid in a porous medium*, Akad. Nauk SSSR. Prikl. Mat. Meh. **16** (1952), 679–698.
6. J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani, and A. Unterreiter, *Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities*, Monatsh. Math. **133** (2001), no. 1, 1–82.
7. J. A. Carrillo, R. J. McCann, and C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoamericana **19** (2003), no. 3, 971–1018.
8. D. Cordero-Erausquin, W. Gangbo, and C. Houdré, *Inequalities for generalized entropy and optimal transportation*, Recent advances in the theory and applications of mass transport, Contemp. Math., vol. 353, Amer. Math. Soc., Providence, RI, 2004, pp. 73–94.
9. L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083.
10. R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179.
11. J. Naudts, *Deformed exponentials and logarithms in generalized thermostatistics*, Phys. A **316** (2002), no. 1-4, 323–334.
12. ———, *Generalized thermostatistics based on deformed exponential and logarithmic functions*, Phys. A **340** (2004), no. 1-3, 32–40.
13. A. Ohara and T. Wada, *q-Gaussian Densities and Behaviors of solutions to Related Diffusion Equations*, arXiv:0814.0624.
14. F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174.
15. F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400.
16. R. E. Pattle, *Diffusion from an instantaneous point source with a concentration-dependent coefficient*, Quart. J. Mech. Appl. Math. **12** (1959), 407–409.
17. A. Takatsu, *Wasserstein geometry of non-linear Fokker–Planck type equations*, in preparation.
18. C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*, J. Statist. Phys. **52** (1988), no. 1-2, 479–487.
19. ———, *Introduction to Nonextensive Statistical Mechanics -Approaching a Complex World*, Springer-Verlag, Berlin, 2009.
20. C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
21. ———, *Optimal transport, old and new*, Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer-Verlag, Berlin, 2009.