Asymptotic equivalence of the discrete variational functional and a rate-large-deviation-like functional in the Wasserstein gradient flow of the porous medium equation

Citation for published version (APA):

Document status and date:
Published: 01/01/2013

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Asymptotic equivalence of the discrete variational functional and a rate-large-deviation-like functional in the Wasserstein gradient flow of the porous medium equation

Manh Hong Duong

Department of Mathematics and Computer Sciences, Technische Universiteit
Email: m.h.duong@tue.nl

July 22, 2013

Abstract

In this paper, we study the Wasserstein gradient flow structure of the porous medium equation. We prove that, for the case of $q$-Gaussians on the real line, the functional derived by the JKO-discretization scheme is asymptotically equivalent to a rate-large-deviation-like functional. The result explains why the Wasserstein metric as well as the combination of it with the Tsallis-entropy play an important role.

Key words. Porous medium equation, Gamma-convergence, Wasserstein gradient flow, variational methods

1 Introduction

1.1 The porous medium equation

In a seminal paper, Otto [Ott01] shows that the porous medium equation

$$\partial_t \rho(t, x) = \Delta \rho^{2-q}(t, x), \quad \text{for } (t, x) \in (0, \infty) \times \mathbb{R}^d \quad \text{and } \rho(0, x) = \rho_0(x), \quad (1)$$

can be interpreted as a gradient flow of the $q$-Tsallis entropy functional$^1$

$$E_q(\rho) = \begin{cases} \frac{1}{q} \int_{\mathbb{R}^d} \rho(x) \left[ \rho(x)^{1-q} - 1 \right] dx & \text{if } q \neq 1, \\ \int_{\mathbb{R}^d} \rho(x) \log \rho(x) dx & \text{if } q = 1, \end{cases} \quad (2)$$

with respect to the Wasserstein distance $W_2$ on the space of probability measures with finite second moment $\mathcal{P}_2(\mathbb{R}^d)$,

$$W_2(\mu, \nu) = \left\{ \inf_{P \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 P(dx dy) \right\}^{1/2}, \quad (3)$$

where $\Gamma(\mu, \nu)$ is the set of all probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals $\mu$ and $\nu$.

This statement can be understood in a variety of different ways. In [Ott01], Otto shows this from a differential geometry point of view by considering $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ as an infinite dimensional Riemannian manifold. However, for the purpose of this paper, the most useful way is that the solution $t \mapsto \rho(t, x)$ can be approximated by the JKO-discretization scheme. Let $T > 0$ be given

$^1E_1$ is the Boltzmann-Shannon entropy
and $h > 0$ be a time step. The discrete approximated solution $\tilde{\rho}^{(n)} = \rho(nh); n = 0, \ldots, \left\lfloor \frac{T}{h} \right\rfloor$ is defined recursively by (see also [JKO98 WW10] and [AGS08] for an exposition of this subject)

$$\tilde{\rho}^{(0)} = \rho_0,$$

$$\tilde{\rho}^{(n)} \in \arg\min_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} K_h(\rho, \tilde{\rho}^{(n-1)}), \quad K_h(\rho, \tilde{\rho}^{(n-1)}) = \frac{1}{4h} W_2^2(\rho, \tilde{\rho}^{(n-1)}) + \frac{1}{2}(E_0(\rho) - E_q(\rho^{(n-1)})).$$

(4)

The theory of Wasserstein gradient flow has been developed tremendously since [JKO98 Ott01]. At least two different approaches have been investigated. The first one is to explore a larger class of evolution equations that are Wasserstein (generalized/modified) gradient flows. Many equations are now known to belong to this class [Gla03 AGS08 GST09 MMS09 Lis09 FG10 Mie11 LMS12]. Recently some attempts have been made to extend the theory to discrete setting [Maa11 Mie11] or to systems that contain also conservative behavior [Hua00 HJ00 DPZ13].

The second approach is to understand why the Wasserstein metric and the combination of it with entropy functional appear in the setting. This direction recently has been received a lot of attention. In [ADPZ11 DLZ12 DLR12 PRV11 Leo12], the authors address this question, for the case of the linear diffusion equation, i.e., for $q = 1$, by establishing an intriguing connection between a microscopic many-particle model and the macroscopic gradient flow structure of the diffusion equation. They show that the functional $K_h$ in (1) is asymptotically equivalent as $h \to 0$ to a discrete rate functional $\tilde{J}_h$ that comes from the large deviation principle of the microscopic model. We now briefly recall the results in these papers in more details. The rate functional $\tilde{J}_h: \mathcal{P}_2(\mathbb{R}^d) \to [0, +\infty]$ is defined by

$$\tilde{J}_h(\rho|\rho_0) = \inf_{Q \in \Gamma(\rho_0, \rho)} H(Q||\tilde{\Phi}_{0\to h}),$$

(5)

where

$$\tilde{\Phi}_{0\to h}(dx, dy) = p_h(x, y)\rho_0(dx)dy; \quad p_h(x, y) = \frac{1}{(4\pi h)^{d/2}} e^{-\frac{|x-y|^2}{4h}},$$

and $H(Q||\tilde{\Phi}_{0\to h})$ is the relative entropy of $Q$ with respect to $\tilde{\Phi}_{0\to h}$

$$H(Q||\tilde{\Phi}_{0\to h}) = \int_{\mathbb{R}^{2d}} \log \left( \frac{dQ}{d\tilde{\Phi}_{0\to h}} \right) dQ, \quad \text{if} \quad \frac{dQ}{d\tilde{\Phi}_{0\to h}} \text{ exists and otherwise being } +\infty.$$

Note that in the case of $d = 1$ and $\rho_0(x) = N(\mu_0, \sigma_0^2)$, then $\tilde{\Phi}_{0\to h}$ is a bivariate Gaussian $N(\mu_0, \sigma_0, \mu_0, \sigma_0^2, \theta_h)$, where $\sigma_h^2 = \sigma_0^2 + 2h$, and $\theta_h = \frac{\mu_h}{\sigma_h^2}$.

It is proven in [Leo12] that

$$4h \tilde{J}_h(\cdot, \rho_0) \overset{\Gamma}{\to} W_2^2(\cdot, \rho_0) \quad \text{as} \quad h \to 0,$$

(6)

where $\Gamma$ is denoted for Gamma-convergence, which is introduced in Section 2. This result is improved to the next order in [ADPZ11 DLZ12 DLR12 PRV11] to give

$$\tilde{J}_h(\cdot|\rho_0) = \frac{1}{4h} W_2^2(\cdot, \rho_0) \overset{\Gamma}{\to} \frac{1}{2}(E_1(\cdot) - E_1(\rho_0)) \quad \text{as} \quad h \to 0.$$ 

(7)

(6) and (7) indicate the asymptotic Gamma development of the functional $\tilde{J}_h(\cdot|\rho_0)$

$$\tilde{J}_h(\rho|\rho_0) \approx \frac{1}{4h} W_2^2(\rho, \rho_0) + \frac{1}{2}(E_1(\rho) - E_1(\rho_0)) + o(1) \quad \text{as} \quad h \to 0.$$ 

(8)

It is worth pointing out that the right hand side of (8) is exactly the functional $K_h$ in the variational scheme (4). It is this asymptotic Gamma development that not only provides the
link between the microscopic model and the Wasserstein gradient flow formulation of the linear diffusion equation but also explains why the Wasserstein metric as well as the combination of it with the entropy play a role in the setting.

The aim of this paper is to generalise (9) - (14) to the nonlinear porous medium equation for the class of \( q \)-Gaussian measures in 1D. As we see in section 2 this class plays an important role because it is invariant under the semigroup of the porous-medium equation and is isometric to the space of Gaussian measures with respect to the Wasserstein metric. We now describe our result in the next Section.

1.2 Main result

Before introducing the main result of the present paper, we recall corresponding concepts for the \( q \)-Gaussian measures.

The \( q \)-exponential function and its inverse, the \( q \)-logarithmic function, are defined respectively by

\[
\exp_q(t) = [1 + (1 - q)t]^{\frac{1}{1-q}},
\]

where \([x]_+ = \max\{0, x\}\) and by convention \(0^{\nu} := \infty\); and

\[
\log_q(t) = \frac{t^{1-q} - 1}{1-q} \quad \text{for} \ t > 0.
\]

Given \( m \), that is specified later on, the \( m \)-relative entropy between \( Q(dx) = f(x)dx \) and \( P(dx) = g(x)dx \)

\[
H_m(Q\|P) = \frac{1}{2 - m} \int [f \log_m f - g \log_m g - (2 - m) \log_m g(f - g)] \, dx
\]

\[
= \frac{1}{2 - m} \int [f \log_m f + (1 - m)g \log_m g - (2 - m)f \log_m g] \, dx.
\]

For \( v \in \mathbb{R}^d \) and \( V \in \text{Sym}^+(d, \mathbb{R}) \), which is the set of all symmetric positive definite matrices of size \( d \), the \( q \)-Gaussian measure with mean \( v \) and covariance matrix \( V \) is

\[
\mathcal{N}_q(v, V) = C_0(q, d)(\det V)^\frac{1}{2} \exp_q \left[ -\frac{1}{2}C_1(q, d)(x - v, V^{-1}(x - v)) \right] \mathcal{L}^d,
\]

where \( \mathcal{L}^d \) is the Lesbesgue measure on \( \mathbb{R}^d \) and \( C_0(q, d), C_1(q, d) \) are explicit positive constants depending only on \( d \) and \( q \) and are given in section 2.

In particular, the \( q \)-Gaussian measure in 1D has density

\[
\mathcal{N}_q(\mu, \sigma^2) = \frac{C_0(q, 1)}{\sigma} \exp_q \left( -\frac{1}{2}C_1(q, 1)\frac{(x - \mu)^2}{\sigma^2} \right).
\]

While in 2D, we call \( q \)-bivariate Gaussian \( \mathcal{N}_q(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \theta) \) with density

\[
\mathcal{N}_q(\mu, \sigma^2) = \frac{C_0(q, 2)}{\sigma_1\sigma_2\sqrt{1 - \theta^2}} \exp_q \left\{ -\frac{1}{2}C_1(1, q) \frac{1}{1 - \theta^2} \left[ \frac{(x - \mu_1)^2}{\sigma_1^2} + \frac{(y - \mu_2)^2}{\sigma_2^2} - \frac{2\theta(x - \mu_1)(y - \mu_2)}{\sigma_1\sigma_2} \right] \right\}
\]

which corresponds to the mean vector \( \mu \) and covariance matrix \( \Sigma \)

\[
\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \theta \sigma_1 \sigma_2 \\ \theta \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}.
\]

It is shown in section 2 that, if the initial data is \( \rho_0(x) = N_q(\mu_0, C\sigma_0^2) \), then the solution to the porous medium equation at time \( t \) is \( \rho(t, x) = N_q(\mu(t), C(t + \sigma_0^{-2})^{\frac{1}{1-q}}) \), where \( C \) is an explicit constant given in section 2.

We are now in the position to introduce the main result.
Theorem 1.1. Let \( d = 1 \), \( q \in \mathbb{Q}_d \equiv (0,1) \cup \left( 1, \frac{d+1}{d+2} \right) \) and \( N_0 = N_q(\mu_0, C\sigma_0^2) \) be given.

We set \( m = 3 - \frac{2}{q} \), \( \sigma_h^2 = (h + \sigma_0^{3-q})^{\frac{2}{3-q}} \) and

\[
Q_{0 \to h} = N_m(\mu_0, C\sigma_0^2, \mu_0, C\sigma_2^2, \frac{\sigma_0}{\sigma_h}),
\]

(16)

For \( N_q = N_q(\mu, C\sigma^2) \), we define the functional \( J_h(N_q, N_0) \)

\[
J_h(N_q|N_0) := \inf_{\mathcal{Q} \in \mathcal{Q}} H_m(Q||Q_{0 \to h}),
\]

(17)

where

\[
\mathcal{Q} := \{N_m(\mu_0, C\sigma_0^2, \mu, C\sigma^2, \theta) \mid \theta \in \mathbb{R} \}.
\]

Then there exist explicit constants \( a = a(\sigma_0, q) \) and \( b = b(\sigma_0, q) \), which are given respectively in (64) and (65), such that the following statements hold

1. \( a(\sigma_h^2 - \sigma_0^2)^{\frac{1}{2}} J_h(\cdot|N_0) \overset{\Gamma}{\to} W_2^2(\cdot, N_0) \) as \( h \to 0 \).

2. \( ab(\sigma_h^2 - \sigma_0^2)^{\frac{1}{2}} J_h(\cdot|N_0) - \frac{b}{(\sigma_h^2 - \sigma_0^2)} W_2^2(\cdot, N_0) \overset{\Gamma}{\to} E_q(\cdot) - E_q(N_0) \) as \( h \to 0 \).

3. If \( 0 < q < 1 \), then \( ab(\sigma_h^2 - \sigma_0^2)^{\frac{1}{2}} J_h(\cdot|N_0) - \frac{1}{2h} W_2^2(\cdot, N_0) \overset{\Gamma-\lim \inf}{\to} E_q(\cdot) - E_q(N_0) \) as \( h \to 0 \).

When \( q \to 1 \), then \( a \to 4, b \to \frac{1}{2} \), \( \sigma_h^2 - \sigma_0^2 \to h \) and we recover (6)–(7) for the diffusion equation.

1.3 Discussion

Given \( N_q \) and \( N_0 \), both \( H_m(N_q||N_0) \) and \( J_h(N_q|N_0) \) are always non-negative and are equal to 0 if and only if \( N_q = N_0 \). By its definition, \( J_h(N_q|N_0) \) measures the deviation of \( N_q \) from the solution of the porous medium equation at time \( h \) given the initial data \( N_0 \). Hence minimizing \( J_h \) means to find the best approximation of the solution at a time \( h \). As a consequence of the Gamma convergence, minimizers of the functionals converge to minimizer of the Gamma-limit functional. Thus the main theorem explains why the Wasserstein metric is involved and why we should minimize the combination of it with the Tsallis entropy.

For the linear diffusion equation, by Sanov’s theorem the relative entropy is the static rate functional of the empirical process of many i.i.d particles. While the functional \( J_h \) in (5) is the rate functional after time \( h \) of the empirical process of many Brownian motions. Hence it has a clear microscopic interpretation. The \( m \)-relative entropy \( H_m \) and the functional \( J_h \) are defined and have similar properties as \( H \) and \( J_h \). However, it is unclear whether \( J_h \) can be proven to be the rate functional of some microscopic stochastic process. In a recent paper [BLMV13], the authors introduce a stochastic particle system, the Ginzburg-Landau dynamics, and show that the porous medium equation is the hydrodynamic limit of the system. Moreover, the \( m \)-relative entropy is the static rate functional of the invariant measures. It would be interesting to study whether the functional \( J_h \) is actually (or related to) the rate functional at time \( h \) of the Ginzburg-Landau dynamics. Another question for future research is to extend the main theorem to a larger class of measures.

1.4 Organisation of the paper

The rest of the paper is organized as follows. In section 2, we first recall the definition of Gamma convergence and relevant properties of the \( q \)-Gaussians. Next we compute the functional \( J_h \) in section 3. Finally, the proof of the main theorem is given in section 4.


2 Preliminaries

We first recall the definition of Gamma convergence for the reader’s convenience.

**Definition.** [Bra02] Let $X$ be a metric space. We say that a sequence $f_n : X \to \mathbb{R}$ $\Gamma$-converges in $X$ to $f : X \to \mathbb{R}$, denoted by $f_n \Gamma \rightharpoonup f$, if for all $x \in X$ we have

1. (lower bound part) For every sequence $x_n$ converging to $x$, \[ \liminf_{n \to \infty} f_n(x_n) \geq f(x), \] \[ (18) \]
2. (upper bound part) There exists a sequence $x_n$ converging to $x$ such that \[ \lim_{n \to \infty} f_n(x_n) = f(x). \] \[ (19) \]

If $f_n$ satisfy the lower (or upper, respectively) bound part then we write $f_n \Gamma\text{-lim inf} f$ (or $f_n \Gamma\text{-lim sup} f$ respectively).

2.1 Properties of $q$-Gaussian measures

The $q$-exponential function and $q$-logarithmic function satisfy the following properties

$$
\exp_q(x) = \log_q(\exp_q x) = x,
$$

and

$$
\log_q(xy) = \log_q x + \log_q y + (1 - q) \log_q(x) \log_q y = \log_q x + x^{1-q} \log_q y = \log_q y + y^{1-q} \log_q x.
$$

The constants $C_0(q,d)$ and $C_1(q,d)$ in (13) are given by

$$
C_1(q,d) = \frac{2}{2 + (d + 2)(1 - q)},
$$

and

$$
C_0(q,d) = \begin{cases} 
\frac{\Gamma\left(\frac{1}{2} + \frac{1}{q}\right)}{\Gamma\left(\frac{1}{2}\right)} \left(\frac{1-q}{2\pi}\right)^\frac{d}{2}C_1(q,d) \right)^\frac{d}{2} & \text{if } 0 < q < 1, \\
\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2} - \frac{1}{q}\right)} \left(\frac{q-1}{2\pi}\right)^\frac{d}{2}C_1(q,d) \right)^\frac{d}{2} & \text{if } 1 < q < \frac{d+4}{d+2}.
\end{cases}
$$

2.2 $q$-Gaussian measures and solutions of the porous medium equation

The porous medium equation (1) has a self-similar solution, which is called the Barenblatt-Pattle solution, of the form

$$
\rho_q(x,t) := \left[A t^{-\alpha(1-q)} - B |x|^{2\alpha} t^{-1}\right]_+^{\frac{1}{1-q}},
$$

$$
= \left[A - B |x|^{2\alpha} t^{-1}\right]_+^{\frac{1}{1-q}} t^{-\alpha},
$$

where

$$
\alpha = \alpha(q,d) := \frac{1}{d(1-q) + 2}, B = B(q,d) := \frac{(1-q)\alpha}{2(2-q)}.
$$

(24)
and \( A \) is a normalization constant

\[
\int_{\mathbb{R}}^{d} \rho_q(x, t) \mathcal{L}^d(dx) = 1.
\]

More precise,

\[
A := C_q 2^{α(1−q)} \left[ \frac{α}{(2−q)C_1} \right]^{dα(1−q)}.
\]  \hspace{1cm} (26)

It is straightforward to see that

\[
\rho_q(x, t) = N_q(0, C\Theta(V_t^d)),
\]

where

\[
C := \frac{(2−q)C_1}{α} A. \hspace{1cm} (27)
\]

It is well known that a solution to the diffusion equation is obtained by a convolution of an initial data with the diffusion kernel. Hence if the initial data is a Gaussian measure \( N(v, V) \) then the solution at time \( t \) is again a Gaussian \( N(v, V_t) \), which is given by

\[
N(v, V_t) = N(v, V) * N(0, 2tI_d) = N(v, V + 2tI_d).
\]

In \cite{Tak12, OW10}, the authors show that a similar statement holds for the porous medium equation on the space of \( q \)-Gaussian measures. That is, a solution to the porous medium equation with an initial data being a \( q \)-Gaussian measure is again a \( q \)-Gaussian for all time. Moreover, to find a solution at time \( t > 0 \), it reduces to solving an ordinary differential equation for the covariance matrix. Let \( Θ \) be a map on \( \text{Sym}^+(d, \mathbb{R}) \) defined by

\[
Θ(V) := (\det V)^{-α(1−q)/3} V. \hspace{1cm} (28)
\]

We note that \( ρ_q(x, t) = N_q(0, CΘ(tI_d)) \).

**Proposition 2.1** \cite{Tak12}. For any \( q \in Q_d \) and \( V \in \text{Sym}^+(d, \mathbb{R}) \), we set the time-dependent matrix \( V_t \) as

\[
Θ(V_t) = Θ(V) + σ(t)I_d, \quad \frac{d}{dt}σ(t) = 2α(\det Θ(V_t))^{-\frac{1}{3}}. \hspace{1cm} (29)
\]

Then \( N_q(v, CΘ(V_t)) \) is a solution to the porous medium equation.

**Remark 2.2.** The assertion also holds true for \( q = 1 \).

We will work out this theorem in 1D in more detail. We calculate relevant variables.

\[
α = \frac{1}{3−q}; \quad Θ(V) = (\det V)^{-α(1−q)} V = V^{1−α(1−q)} = V^{\frac{3}{3−q}}; \quad \det Θ(V_t) = V_t. \hspace{1cm} (30)
\]

The ODE becomes

\[
Θ(V_t) = Θ(V) + σ(t)I_d, \quad \frac{d}{dt}σ(t) = 2α(\Theta(V_t))^{-\frac{1}{3}}. \hspace{1cm} (31)
\]

Solving this ODE we get

\[
Θ(V_t) = \left[ α(3−q)t + Θ(V)^{\frac{3}{3−q}} \right]^{\frac{3−q}{3}} = (t + V)^{\frac{3−q}{3}}.
\]

So if \( ρ_0(x) = N_q(v, CV^{\frac{3}{3−q}}) \) then \( ρ(t, x) = N_q(v, C(t + V)^{\frac{3}{3−q}}) \) for all \( t > 0 \). In other words, if \( ρ_0(x) = N_q(μ_0, Cσ_0^2) \) then \( ρ(t, x) = N_q(μ_0, C(t + σ_0^{3−q})^{\frac{3}{3−q}}) \) for all \( t > 0 \).
2.3 q-Gaussian and the Wasserstein metric

The q-Gaussian measures have another important property stated in the following Proposition.

Proposition 2.3. For any $q \in Q_d$, the space of q-Gaussian measures is convex and isometric to the space of Gaussian measures with respect to the Wasserstein metric.

Hence

$$W_2(N_q(\mu, \Sigma), N_q(\nu, V))^2 = W_2(N(\mu, \Sigma), N(\nu, V))^2 = |\mu - \nu|^2 + \text{tr} \Sigma + \text{tr} V - 2\text{tr} \sqrt{\Sigma^* V^*}. $$

In particular

$$W_2(N_q(\mu_1, \sigma_1^2), N_q(\mu_2, \sigma_2^2))^2 = (\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2. $$

(31)

3 Computing the functional $J_h$

In this section, we compute the functional $J_h$ explicitly. The following Proposition is true for any dimensional $d$. For clarity, we use a boldface font for vectors.

Proposition 3.1. Let $P(dx) = g(x)dx$ be absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$. Let $Q$ be the set of all Borel measures $Q = f(x)dx$ in $\mathbb{R}^d$ satisfying

$$\int r_i(x) \cdot f(x)dx = a_i, \; i \in \{1, 2, \ldots N\}, \quad (32)$$

where $r_i(x): \mathbb{R}^d \to \mathbb{R}$ are given functions and $a_i \in \mathbb{R}$.

Assume that there is a measure $Q^* \in Q$ that has a density satisfying the equation

$$\log_m f^*(x) = \log_m g(x) + \sum_{i=1}^{N} \lambda_i r_i(x), \quad (33)$$

for some $\lambda_i \in \mathbb{R}$. Then there holds

$$H_m(Q\|P) = H_m(Q^*\|P) + H_m(Q\|Q^*) \quad \text{for all } Q \in Q. \quad (34)$$

and as a consequence, $Q^*$ is the unique minimiser of $H_m(Q\|P)$ over all $Q \in Q$.

Proof. We have

$$(2 - m)H_m(Q\|P) = \int f \log_m f + (1 - m)g \log_m g - (2 - m)f \log_m g$$

$$= \int f \log_m f + (1 - m)f^* \log_m f^* - (2 - m)f \log_m f^*$$

$$+ \int (2 - m)f(\log_m f^* - \log_m g) - (1 - m)(f^* \log_m f^* - g \log_m g)$$

$$= (2 - m)H_m(Q^*\|Q) + \int (2 - m)f(\log_m f^* - \log_m g) - (1 - m)(f^* \log_m f^* - g \log_m g)$$

(35)
We rewrite the second line in (35) using (32) and (33)

\[
\begin{align*}
(2 - m) \int f(\log m f^* - \log m g) - (1 - m)(f^* \log m f^* - g \log m g) \\
= (2 - m) \int f \sum_{i=1}^{N} \lambda_i r_i - (1 - m) \int (f^* \log m f^* - g \log m g) \\
= (2 - m) \int f^* \sum_{i=1}^{N} \lambda_i r_i - (1 - m) \int (f^* \log m f^* - g \log m g) \\
= (2 - m) \int f^*(\log m f^* - \log m g) - (1 - m) \int (f^* \log m f^* - g \log m g) \\
= \int f^* \log m f^* - (2 - m)f^* \log m g + (1 - m)g \log m g
\end{align*}
\]

From (35) and (36) we obtain (34).

Remark 3.2. The property that the relative entropy and the \(m\)-relative entropy satisfy a generalized Pythagorean relation is well-known in the literature, see for instance [Csi75] and [OW10] for similar relationship.

We now apply this Proposition to the \(q\)-bivariate measures.

**Proposition 3.3.** Let \(P = N_m(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \theta)\) be given. We define

\[
Q := \left\{ N_m(\nu_1, \xi_1, \nu_2, \xi_2, \tilde{\theta}) \mid \tilde{\theta} \in \mathbb{R} \right\}
\]

The minimizer of the minimization problem

\[
\min_{Q \in \mathcal{Q}} H_m(Q \parallel P),
\]

is given by

\[
Q^* = N_m(\nu_1, \xi_1, \nu_2, \xi_2, \eta),
\]

where

\[
\frac{\eta}{(1 - \eta^2)^{2 - m}} = \frac{\theta}{(1 - \theta^2)^{2 - m}} \left( \frac{\xi_1 \xi_2}{\sigma_1 \sigma_2} \right)^{2 - m}.
\]

**Proof.** We actually prove a stronger statement, namely that \(Q^*\) is the minimiser of \(H(\cdot \parallel P)\) over \(\mathcal{Q}\), which is defined as the set of all \(Q(dx) = f(x)dx\) satisfying

\[
\begin{align*}
\int f(x, y)dxdy &= 1, \\
\int xf(x, y)dxdy &= \nu_1, \\
\int x^2 f(x, y)dxdy &= \xi_1^2 + \nu_1^2, \\
\int yf(x, y)dxdy &= \nu_2, \\
\int y^2 f(x, y)dxdy &= \xi_2^2 + \nu_2^2.
\end{align*}
\]
Since $Q^* \in Q \subset \tilde{Q}$, $Q^*$ is also the unique minimizer on $Q$. Let $g(x)$ and $f^*(x)$ respectively denote the densities of $P$ and $Q^*$. By Proposition 3.3, $f^*$ satisfies the following equation

$$\log_m f^*(x) = \log_m g(x) + \sum_{i=1}^{5} \lambda_i r_i(x),$$

where $r_1 = 1$, $r_2 = x$, $r_3 = y$, $r_4 = x^2$, $r_5 = y^2$. To get the equation (40), we simply equalise the coefficients of $xy$ in two sides of (41) and obtain

$$\frac{\theta}{\sigma_1\sigma_2(1-\theta^2)} \left[ \frac{C_0(m,d)}{\sigma_1\sigma_2\sqrt{1-\theta^2}} \right]^{1-m} \frac{\eta}{\xi_1\xi_2(1-\eta^2)} \left[ \frac{C_0(m,d)}{\sigma_1\sigma_2\sqrt{1-\eta^2}} \right]^{1-m},$$

which is equivalent to

$$\frac{\eta}{(1-\eta^2)^{\frac{1-m}{2}}} = \frac{\theta}{(1-\theta^2)^{\frac{1-m}{2}}} \left( \frac{\xi_1\xi_2}{\sigma_1\sigma_2} \right)^{2-m}.$$

\[\square\]

We now compute the $m$-entropy of a $m$-Gaussian and the $m$-relative entropy between two $m$-Gaussians.

**Proposition 3.4.** We have

$$E_m(N_m(\mu, \Sigma)) - E_m(N_m(\mu, V)) = (2-m)C_1(m, d) \left( \frac{C_0(m,d)}{\det V} \right)^{\frac{1-m}{2}} \log_m \left( \frac{\det V}{\det \Sigma} \right)^{\frac{1}{2}}.$$  

and

$$H_m(N_m(\mu, \Sigma), N_m(\nu, V)) = \frac{1}{2} C_1(m, d) \left( \frac{C_0(m,d)}{\det V} \right)^{\frac{1-m}{2}} \left[ \text{tr}(V^{-1}\Sigma) + \langle \mu - \nu, V^{-1}(\mu - \nu) \rangle + 2 \log_m \left( \frac{\det \Sigma}{\det V} \right)^{\frac{1}{2}} - d \right].$$

**Remark 3.5.** When $m = 1$, we recover the corresponding formula for the Gaussian measures.

**Proof.** We denote by $f(x)$ and $g(x)$ respectively the densities of $N_m(\mu, \Sigma)$ and $N_m(\nu, V)$. Using the explicit formula of $f$ and $g$ as in [18] and by a straightforward calculation we get

$$\int f \log_m f \, dx = \log_m \left[ \frac{C_0(m,d)}{\det \Sigma} \exp_m \left( -\frac{d}{2} C_1(m,d) \right) \right]\tag{41}$$

$$\int g \log_m g \, dx = \log_m \left[ \frac{C_0(m,d)}{\det V} \exp_m \left( -\frac{d}{2} C_1(m,d) \right) \right]\tag{42}$$

$$\int f \log_m g \, dx = \log_m \left[ \frac{C_0(m,d)}{\det V} \exp_m \left( -\frac{d}{2} C_1(m,d) \left( \text{tr}(V^{-1}\Sigma) + \langle \mu - \nu, V^{-1}(\mu - \nu) \rangle \right) \right) \right].$$

We now compute the $m$-entropy of a $m$-Gaussian and the $m$-relative entropy between two $m$-Gaussians.
Hence
\[
\int (f \log_m f - g \log_m g) \, dx = \left( 1 - (1 - m) \frac{d}{2} C_1(m, d) \right) \left[ \log_m \frac{C_0(m, d)}{(\det \Sigma)^{\frac{d}{2}}} - \log_m \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right] 
\]
\[
= \left( 1 - (1 - m) \frac{d}{2} C_1(m, d) \right) \left( \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right)^{1-m} \log_m \frac{C_0(m, d)}{(\det \Sigma)^{\frac{d}{2}}}. \tag{49}
\]
and
\[
\int (g - f) \log_m g \, dx = \frac{1}{2} C_1(m, d) \left( \text{tr}(V^{-1} \Sigma) + (\mu - \nu, V^{-1}(\mu - \nu)) - d \right) \left[ 1 + (1 - m) \log_m \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right] 
\]
\[
= \frac{1}{2} C_1(m, d) \left( \text{tr}(V^{-1} \Sigma) + (\mu - \nu, V^{-1}(\mu - \nu)) - d \right) \left( \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right)^{1-m}. \tag{50}
\]
Since
\[
1 - (1 - m) \frac{d}{2} C_1(m, d) = (2 - m) C_1(m, d),
\]
We get
\[
E_m(N_m(\mu, \Sigma)) - E_m(N_m(\mu, V)) = \int (f \log_m f - g \log_m g) \, dx 
\]
\[
= (2 - m) \left( \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right)^{1-m} \log_m \frac{C_0(m, d)}{(\det \Sigma)^{\frac{d}{2}}}. \tag{52}
\]
and
\[
H_m(N_m(\mu, \Sigma), N_m(\nu, V)) = \frac{1}{2 - m} \int [f \log_m f - g \log_m g - (2 - m) \log_m g(f - g)] \, dx 
\]
\[
= \frac{1}{2} C_1(m, d) \left( \frac{C_0(m, d)}{(\det V)^{\frac{d}{2}}} \right)^{1-m} \left[ \text{tr}(V^{-1} \Sigma) + (\mu - \nu, V^{-1}(\mu - \nu)) + 2 \log_m \frac{(\det V)^{\frac{d}{2}}}{(\det \Sigma)^{\frac{d}{2}}} - d \right]. \tag{53}
\]

**Proposition 3.6.** Let \( N_q(\mu_0, C\sigma^2_0), N_q(\mu, C\sigma^2) \) be given. Set \( \sigma^2_n = (h + \sigma_0^{3-q}) \frac{t}{\alpha_0} \) and let \( N_q(\mu_0, C\sigma^2_h) \) be the solution at time \( h \). Let
\[
Q_{0 \to h} = N_m(\mu_0, C\sigma^2_0, \mu_0, C\sigma^2_h, \theta_h), \quad \text{where} \quad \theta_h = \frac{\sigma_0}{\sigma_h},
\]
be the \( m \)-bivariate with mean vector
\[
\mu_{0 \to h} = \begin{pmatrix} \mu_0 \\ \mu_0 \end{pmatrix}, \quad \Sigma_{0 \to h} = \begin{pmatrix} C\sigma^2_0 & C\theta_h \sigma_0 \sigma_h \\ C\theta_h \sigma_0 \sigma_h & C\sigma^2_h \end{pmatrix} = \begin{pmatrix} C\sigma^2_0 & C\sigma^2_h \\ C\sigma^2_0 & C\sigma^2_h \end{pmatrix}.
\]
Define
\[
Q = \left\{ N_m(\mu_0, C\sigma^2_0, \mu, C\sigma^2, \theta) \mid \theta \in \mathbb{R} \right\}
\]
and
\[
Q^* = \arg\min_{Q \in Q} H_m(Q \| Q_{0 \to h})
\]
Then
\[
Q^* = N_m(\mu_0, C\sigma^2_0, \mu, C\sigma^2, \eta_h),
\]
which is a \( m \)-bivariate with the mean vector and covariance matrix
\[
\mu^* = \begin{pmatrix} \mu_0 \\ \mu_0 \end{pmatrix}, \quad \Sigma^* = \begin{pmatrix} C\sigma^2_0 & C\eta_h \sigma_0 \sigma \sigma \end{pmatrix}.
\]
where \( \eta_h \) satisfies the equation

\[
\frac{\eta_h}{(1 - \eta_h^2)^{\frac{1}{2-m}}} = \frac{\theta_h}{(1 - \theta_h^2)^{\frac{1}{2-m}}} \left( \frac{\sigma}{\sigma_h} \right)^{2-m}.
\]  

(60)

Moreover, the \( m \)-relative entropy \( H_m(Q^* || Q_{0 \rightarrow h}) \) is

\[
H_m(Q^* || Q_{0 \rightarrow h}) = \frac{1}{2} C_1(m, 2) \left[ \frac{C_0(m, 2)}{(\det \Sigma_{0 \rightarrow h})^{\frac{1}{2}}} \right]^{1-m} \left[ \frac{\det \Sigma_{0 \rightarrow h}^{\frac{1}{2}}}{\det \Sigma^*^{\frac{1}{2}}} - 2 \right].
\]

(61)

The difference \( E_q(N_q(\mu, C\sigma^2)) - E_q(N_q(\mu_0, C\sigma_0^2)) \) is

\[
E_q(N_q(\mu, C\sigma^2)) - E_q(N_q(\mu_0, C\sigma_0^2)) = (2 - q) C_1(q, 1) \left( \frac{C_0(q, 1)}{\sigma_0 \sqrt{C}} \right)^{1-q} \log \left( \frac{\sigma_0}{\sigma} \right),
\]

(62)

and the Wasserstein distance \( W_2^2(N_q(\mu, C\sigma^2), N_q(\mu_0, C\sigma_0^2)) \) is

\[
W_2(N_q(\mu, C\sigma^2), N_q(\mu_0, C\sigma_0^2))^2 = C(\sigma - \sigma_0)^2 + (\mu - \mu_0)^2.
\]

(63)

Proof. By Proposition 3.4

We now calculate each term in the above formula explicitly.

\[
\det \Sigma_{0 \rightarrow h} = C^2 \sigma_0^2 (\sigma_h^2 - \sigma_0^2) = C^2 \sigma_0^2 \sigma_h^2 (1 - \theta_h^2), \quad \Sigma_{0 \rightarrow h}^{-1} = \frac{1}{C(\sigma_h^2 - \sigma_0^2)} \begin{pmatrix} \sigma_h^2 & -1 \\ -1 & 1 \end{pmatrix}.
\]

\[
\det \Sigma^* = C^2 \sigma_0^2 \sigma_0^2 (1 - \eta^2).
\]

\[
\text{tr} \left( \Sigma_{0 \rightarrow h}^{-1} \Sigma^* \right) = \frac{\sigma_h^2 + \sigma_0^2 - 2 \eta \sigma_0 \sigma}{\sigma_h^2 - \sigma_0^2} = \frac{(\sigma - \sigma_0)^2}{\sigma_h^2 - \sigma_0^2} + \frac{2 \sigma_0 \sigma (1 - \eta)}{\sigma_h^2 - \sigma_0^2} + 1.
\]

\[
(\mu^* - \mu_{0 \rightarrow h})^T \Sigma_{0 \rightarrow h}^{-1} (\mu^* - \mu_{0 \rightarrow h}) = \frac{(\mu - \mu_0)^2}{C(\sigma_h^2 - \sigma_0^2)}.
\]

\[
\frac{\det \Sigma_{0 \rightarrow h}}{\det \Sigma^*} = \frac{\sigma_h^2}{\sigma^2} \frac{1 - \theta_h^2}{1 - \eta^2}.
\]

By (60)

\[
\frac{\eta_h}{(1 - \eta_h^2)^{\frac{1}{2-m}}} = \frac{\theta_h}{(1 - \theta_h^2)^{\frac{1}{2-m}}} \left( \frac{\sigma}{\sigma_h} \right)^{2-m}.
\]

Hence

\[
\frac{1 - \theta_h^2}{1 - \eta_h} = \left( \frac{\theta_h}{\eta_h} \right)^{\frac{1}{2-m}} \left( \frac{\sigma}{\sigma_h} \right)^{\frac{2(2-m)}{2-m}}.
\]

and

\[
\frac{\sigma_h^2}{\sigma^2} \frac{1 - \theta_h^2}{1 - \eta_h} = \left( \frac{\sigma_h \theta_h}{\sigma \eta_h} \right)^{\frac{1}{2-m}} = \left( \frac{\sigma_0}{\sigma \eta_h} \right)^{\frac{2}{2-m}}.
\]
Therefore
\[
\frac{(\det \Sigma_{0 \to h})^{\frac{1}{2}}}{(\det \Sigma^*)^{\frac{1}{2}}} = \left( \frac{\sigma_0}{\sigma\eta_h} \right)^{\frac{1}{1-m}},
\]
and (61) follows.

(62) is a direct consequence of the first equality in Proposition 3.4 and (63) has been shown in (61).

\[\square\]

\section{Proof of the main theorem}

In this section, we bring all ingredients together to prove the main theorem. Suppose that the assumption of the main theorem is true. We set
\[
a(q, \sigma_0) = \frac{2C}{C_1(m, 2)} \left( \frac{C_0(m, 2)}{C\sigma_0} \right)^{m-1} = \frac{2C^{2-m}}{C_1(m, 2)} \left( \frac{C_0(m, 2)}{\sigma_0} \right)^{m-1}. \tag{64}
\]
\[
b(\sigma_0, q) = \frac{(2-\theta)C_1(q, 1)}{C} \left( \frac{C_0(q, 1)}{\sigma_0\sqrt{C}} \right)^{1-q} = \frac{2-\theta}{C} \left( \frac{C_0(q, 1)}{\sigma_0} \right)^{1-q}. \tag{65}
\]

Let \(Q^*\) be the minimizer in (17). By Proposition 3.6, we have
\[Q^* = \mathcal{N}_m(\mu_0, \sigma_0^2, \mu, \sigma^2, \eta_h),\]
where \(\eta_h = \eta(h, \sigma)\) satisfies the following equation
\[
\frac{\eta_h}{1-\eta_h^{2-q}} = \frac{\theta h}{1-\eta_h} \left( \frac{\sigma}{\sigma_0} \right)^{2-q} = \frac{\theta h}{1-\eta_h} \left( \frac{\sigma}{\sigma_0} \right)^{2-q}. \tag{66}
\]

Since \(|\eta_h| \leq 1\) and the right hand side of (66) is positive, it holds that \(0 < \eta_h < 1\). Using the relationship \(m = 3 - \frac{2}{q}\), we can rewrite the above equation as follows
\[
\frac{\eta_h^q}{1-\eta_h^q} = \frac{\sigma_0^q - \sigma^q}{\sigma_h^q - \sigma_0^q}. \tag{67}
\]

We now use the following statement whose proof is straightforward.

Assume that \(x_0\) is given. For all \(\epsilon > 0\), and for all \(h > 0\) sufficiently small, there exists a constant \(C = C(\epsilon, x_0)\) such that
\[\left( h + x_0 \right)^\epsilon - x_0^\epsilon \leq Ch. \tag{68}\]

In particular, if \(\epsilon < 1\) then \(C = \epsilon x_0^{\epsilon-1}\).

Using (63) for \(\epsilon = \frac{2}{3-q}, x_0 = \sigma^{3-q}\) and from (67), we get
\[
1 - \eta_h = \frac{\eta_h^q}{1 + \eta_h} \frac{\sigma_0^q - \sigma^q}{\sigma_h^q - \sigma_0^q} = \frac{\eta_h^q}{1 + \eta_h} \frac{(h + \sigma_0^{3-q})^{\epsilon-1} - \sigma_0^{\epsilon-1}}{\sigma_h^{\epsilon-1} - \sigma_0^{\epsilon-1}} \leq \frac{Ch}{\sigma^{(\epsilon-1)}}. \tag{69}
\]

where \(C > 0\) is a constant depending only on \(\sigma_0\) and \(q\). This implies that for fixed \(\sigma\), \(\lim_{h \to 0} \eta_h = \lim_{h \to 0} \eta(h, \sigma) = 1\) and as a sequence of functions \(\eta(h, \cdot) \to 1\) locally uniform.

1. We now prove the first statement of the main theorem. We need to prove
\[
a(\sigma_h^2 - \sigma_0^2) \overset{\text{d}}{\to} J_h(\cdot | \mathcal{N}_0) \overset{\text{d}}{\to} W_2^2(\cdot, \mathcal{N}_0) \quad \text{as} \quad h \to 0. \tag{70}
\]

Let \(\mathcal{N}_q(\mu, C\sigma^2)\) be given and we denote it by \(\mathcal{N}_q\) for short. By Proposition 3.6 we have
\[
a(\sigma_h^2 - \sigma_0^2) \overset{\text{d}}{\to} J_h(\mathcal{N}_q|\mathcal{N}_0) = C(\sigma - \sigma_0)^2 + (\mu - \mu_0)^2 + 2C\sigma_0\sigma(1-\eta_h) + C(\sigma_h^2 - \sigma_0^2) \left[ 2 \log_m \left( \frac{\sigma_0}{\sigma\eta_h} \right)^{\frac{1}{1-m}} - 1 \right]. \tag{71}
\]
2. We now prove the second statement of the main theorem. We need to prove

\[ W_2^2 (\mathcal{N}_q, \mathcal{N}_0) = C (\sigma - \sigma_0)^2 + (\mu - \mu_0)^2. \]  

(72)

For the lower bound part: Assume that \( N_q^h = \mathcal{N}_q (\eta_h, C \xi_h^2) \to \mathcal{N}_q \). This means that \((\nu_h - \mu)^2 + (\xi_h - \sigma)^2 \to 0\). Hence we can assume that \( 0 < \frac{\sigma}{2} \leq \sup_h \xi_h \leq \frac{3}{2} \sigma \).

Let \( \eta_h = \eta(h, \xi_h) \) be the solution of (69) where \( \sigma \) is replaced by \( \xi_h \). By (69) we have \( \lim_{h \to 0} \eta_h = 1 \), so we can assume that \( \xi_h \eta_h \) is uniformly bounded above and away from 0.

We now have

\[
a(\sigma_h^2 - \sigma_0^2) \frac{\partial}{\partial h} J_h (N_q^h | \mathcal{N}_0) \\
= C (\xi_h - \sigma_0)^2 + (\nu_h - \mu_0)^2 + 2 C \sigma_0 \xi_h (1 - \eta_h) + C (\sigma_h^2 - \sigma_0^2) \left[ 2 \log_m \left( \frac{\sigma_0}{\xi_h \eta_h} \right) \right] - 1.
\]

\[
\geq C (\xi_h - \sigma_0)^2 + (\nu_h - \mu_0)^2 + C (\sigma_h^2 - \sigma_0^2) \left[ 2 \log_m \left( \frac{\sigma_0}{\xi_h \eta_h} \right) \right] - 1.
\]

Hence

\[
\liminf_{h \to 0} a(\sigma_h^2 - \sigma_0^2) \frac{\partial}{\partial h} J_h (N_q^h | \mathcal{N}_0)
\geq \liminf_{h \to 0} \left\{ C (\xi_h - \sigma_0)^2 + (\nu_h - \mu_0)^2 + C (\sigma_h^2 - \sigma_0^2) \left[ 2 \log_m \left( \frac{\sigma_0}{\xi_h \eta_h} \right) \right] - 1 \right\}
= C (\sigma - \sigma_0)^2 + (\mu - \mu_0)^2 = W_2^2 (\mathcal{N}_q, \mathcal{N}_0).
\]

For the upper bound part: as a recovery sequence, we just simply take the fixed sequence \( N_q^h = \mathcal{N}_q \).

2. We now prove the second statement of the main theorem. We need to prove:

\[
ab(\sigma_h^2 - \sigma_0^2) \frac{\partial}{\partial h} J_h (\mathcal{N}_q | \mathcal{N}_0) = \frac{b}{\sigma_h^2 - \sigma_0^2} W_2^2 (\cdot, \mathcal{N}_0) \xrightarrow{\Gamma} E_q (\cdot) - E_q (\mathcal{N}_0) \quad \text{as } h \to 0.
\]

(73)

Let \( \mathcal{N}_q = \mathcal{N}_q (\mu, \sigma^2) \) be given. Let \( \eta_h \) be the solution of (66). We have

\[
ab(\sigma_h^2 - \sigma_0^2) \frac{\partial}{\partial h} J_h (\mathcal{N}_q | \mathcal{N}_0) = b C \left\{ \frac{(\sigma - \sigma_0)^2}{\sigma_h^2 - \sigma_0^2} + \frac{(\mu - \mu_0)^2}{C (\sigma_h^2 - \sigma_0^2)} + 2 \frac{\sigma_0 \sigma (1 - \eta_h)}{\sigma_h^2 - \sigma_0^2} + \left[ 2 \log_m \left( \frac{\sigma_0}{\sigma \eta_h} \right) \right] - 1 \right\},
\]

\[
= b C \left\{ \frac{(\sigma - \sigma_0)^2}{\sigma_h^2 - \sigma_0^2} + \frac{(\mu - \mu_0)^2}{C (\sigma_h^2 - \sigma_0^2)} \right\},
\]

\[
E_q (\mathcal{N}_q) - E_q (\mathcal{N}_0) = b C \log \left( \frac{\sigma_0}{\sigma} \right).
\]

Define

\[
F_h (\mathcal{N}_q | \mathcal{N}_0) := \frac{2 \sigma_0 \sigma (1 - \eta_h)}{\sigma_h^2 - \sigma_0^2} + 2 \log_m \left( \frac{\sigma_0}{\sigma \eta_h} \right) \frac{1}{\tau_m} - 1,
\]

\[
F (\mathcal{N}_q | \mathcal{N}_0) := \log \left( \frac{\sigma_0}{\sigma} \right).
\]

We need to prove

\[
F_h (\cdot | \mathcal{N}_0) \xrightarrow{\Gamma} F (\cdot | \mathcal{N}_0).
\]

(77)

We first prove that \( F_h (\cdot, \mathcal{N}_0) \to F (\cdot, \mathcal{N}_0) \) locally uniform. We now rewrite the RHS of \( F_h \) using the relationship between \( m \) and \( q \).
Since for any \( t > 0 \)
\[
\log_m t = \frac{t^{\frac{1}{m}} - 1}{1 - m} = \frac{t^{1-q} - 1}{2(1-q)} = \frac{q}{2} \log_q t;
\]
Hence
\[
2 \log_m \left( \frac{\sigma_0}{\sigma_\eta h} \right) = q \log_q \left( \frac{\sigma_0}{\sigma_\eta h} \right) . \tag{78}
\]
From (79), we get
\[
\frac{\eta_h}{1 + \eta_h} = \left( \frac{\sigma_0 \sigma (1 - \eta_h)}{\sigma_\eta h - \sigma_0^2} \right)^{\frac{1}{1-q}} \left( \frac{\sigma}{\sigma_0} \right)^{\frac{1}{1-m}} . \tag{79}
\]
This implies that
\[
\frac{\sigma_0 \sigma (1 - \eta_h)}{\sigma_\eta h - \sigma_0^2} = \frac{\eta_h}{1 + \eta_h} \left( \frac{\sigma_0}{\sigma} \right)^{1-q} = \frac{\eta_h}{1 + \eta_h} \left( 1 - q \right) \log_q \left( \frac{\sigma_0}{\sigma} \right) + 1 . \tag{80}
\]
From (78) and (80) we obtain
\[
F_h(\mathcal{N}_q; \mathcal{N}_0) = 2 \frac{\eta_h}{1 + \eta_h} \left( \frac{\sigma_0}{\sigma} \right)^{1-q} + q \log_q \left( \frac{\sigma_0}{\sigma_\eta h} \right) - 1 .
\]
Now we have the following estimate
\[
\left| F_h - F \right| = \left| 2 \frac{\eta_h}{1 + \eta_h} \left( \frac{\sigma_0}{\sigma} \right)^{1-q} + q \log_q \left( \frac{\sigma_0}{\sigma_\eta h} \right) - \log_q \left( \frac{\sigma_0}{\sigma} \right) - 1 \right|
\]
\[
= \left| 2 \frac{\eta_h}{1 + \eta_h} \left( \frac{\sigma_0}{\sigma} \right)^{1-q} - 1 - (1 - q) \log_q \left( \frac{\sigma_0}{\sigma} \right) + q \log_q \left( \frac{\sigma_0}{\sigma_\eta h} \right) - q \log_q \left( \frac{\sigma_0}{\sigma} \right) \right|
\]
\[
= \left( \frac{\sigma_0}{\sigma} \right)^{1-q} \left| 2 \frac{\eta_h}{1 + \eta_h} - 1 + \frac{q}{1 - q} \left( \eta_h(q-1) - 1 \right) \right|
\]
\[
\leq \left( \frac{\sigma_0}{\sigma} \right)^{1-q} \left| \left[ 2 \frac{\eta_h}{1 + \eta_h} - 1 \right] + \frac{q}{1 - q} \left( \eta_h(q-1) - 1 \right) \right|
\]
\[
\leq \frac{1}{\sigma_0^{1-q}} C \left( 1 - \eta_h \right)
\leq \frac{Ch}{\sigma_0^{3-2q}} . \tag{81}
\]
where \( C \) is a constant depending only on \( \sigma_0 \) and \( q \).

The locally uniform convergence of \( F_h \) to \( F \) thus follows from the estimate (81).

For the \( \Gamma \)-convergence, we get the lower bound part by the local uniform convergence and the continuity of the entropy. Indeed, let \( \mathcal{N}_q := \mathcal{N}_q(\mu, C\xi^2) \) be given and assume that \( \mathcal{N}_q^h := \mathcal{N}_q(\mu_h, C\xi_h^2) \to \mathcal{N}_q \). Then \( \mu_h \to \mu \) and \( \xi_h \to \sigma \). Hence we can assume without loss of generality that \( \frac{\xi_h}{\sigma} \leq \sup \xi_h \leq \frac{3\xi}{\sigma} \). We have the following estimate
\[
\left| F_h(\mathcal{N}_q^h; \mathcal{N}_q^0) - F(\mathcal{N}_q^h; \mathcal{N}_q^0) \right| \leq \left| F_h(\mathcal{N}_q^h; \mathcal{N}_0) - F(\mathcal{N}_q^h; \mathcal{N}_0^0) \right| + \left| F(\mathcal{N}_q^h; \mathcal{N}_0) - F(\mathcal{N}_q^h; \mathcal{N}_q^0) \right|
\]
\[
\leq \frac{Ch}{\xi_h^{3-2q}} + \left| \log_q \left( \frac{\sigma_0}{\xi_h} \right) - \log_q \left( \frac{\sigma_0}{\xi_h} \right) \right|
\]
\[
\leq \frac{Ch}{\sigma_0^{3-2q}} + \left| \log_q \left( \frac{\sigma_0}{\xi_h} \right) - \log_q \left( \frac{\sigma_0}{\xi_h} \right) \right| \to 0 .
\]
Therefore
\[
\lim_{h \to 0} F_h(\mathcal{N}_q^h; \mathcal{N}_q^0) = F(\mathcal{N}_q^h; \mathcal{N}_q^0) .
\]
For the upper part, as a recovery sequence, we can choose the fixed sequence \( \mathcal{N}_q^h = \mathcal{N}_q \).
3. We now prove the third statement of the main theorem. Assume that \( 0 < q < 1 \). We need to prove

\[
abla \frac{\sigma^2_f - \sigma^2_0}{1 - q} \Gamma h(\cdot|N_0) - \frac{1}{2h} W^2_2(\cdot, N_0) \xrightarrow{\Gamma \text{-lim inf}} E_q(\cdot) - E_q(N_0) \quad \text{as } h \rightarrow 0.
\]

We have

\[
ab(\sigma^2_f - \sigma^2_0) \frac{1}{1 - q} J_h(\cdot|N_0) - \frac{1}{2h} W^2_2(\cdot, N_0)
= ab(\sigma^2_f - \sigma^2_0) \frac{1}{1 - q} J_h(\cdot|N_0) - \frac{b}{\sigma^2_f - \sigma^2_0} W^2_2(\cdot, N_0) + \left( \frac{b}{\sigma^2_f - \sigma^2_0} - \frac{1}{2h} \right) W^2_2(\cdot, N_0).
\]

Since \( 0 < q < 1, 0 < \frac{2}{3 - q} < 1 \), using (61) for \( \epsilon = \frac{2}{3 - q}, \ x_0 = \sigma^3_{0 - q} \), we have

\[
\sigma^2_f - \sigma^2_0 = (h + \sigma^{3 - q}) \frac{2}{(3 - q)\sigma^{1 - q}_0} \leq \frac{2}{(3 - q)\sigma^{1 - q}_0} .
\]

Therefore

\[
\frac{b}{\sigma^2_f - \sigma^2_0} \geq \frac{(3 - q)bx_0^{1 - q}}{2h} = \frac{1}{2h}.
\]

It implies that

\[
ab(\sigma^2_f - \sigma^2_0) \frac{1}{1 - q} J_h(\cdot|N_0) - \frac{1}{2h} W^2_2(\cdot, N_0) \geq ab(\sigma^2_f - \sigma^2_0) \frac{1}{1 - q} J_h(\cdot|N_0) - \frac{b}{\sigma^2_f - \sigma^2_0} W^2_2(\cdot, N_0),
\]

and the third statement thus follows from the second one.

This completes the proof of the main theorem.

Acknowledgement

The research of the paper has received funding from the ITN “FIRST” of the Seventh Framework Programme of the European Community (grant agreement number 238702). This project was done partially when the author was in the University of Bath as part of the ITN program. The author would like to thank Vaios Laschos for his helpful discussion.

References


