Statistics and Probability Letters 137 (2018) 54–62

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Convergence in total variation distance for (in)homogeneous Markov processes

Yong-Hua Mao [a], Liping Xu [b,*], Ming Zhang [c], Yu-Hui Zhang [a]

[a] School of Mathematical Sciences, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing Normal University, Beijing, 100875, PR China

[b] Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris cedex 05, France

[c] Department of Science and Technology, China University of Political Science and Law, Beijing, 102249, PR China

ARTICLE INFO

Article history:
Received 6 January 2017
Received in revised form 6 January 2018
Accepted 11 January 2018
Available online 2 February 2018

MSC:
26D10
47D07
60J25
60J60

Keywords:
(In)homogeneous Markov processes
Functional inequalities
Rate of convergence
Total variation distance

ABSTRACT

In this paper, we study the rate of convergence in total variation distance for time continuous Markov processes, by using some I_ψ and $I_{\psi,t}$-inequalities. For homogeneous reversible process, we use some homogeneous inequalities, including the Poincaré and relative entropy inequalities. For the time-inhomogeneous diffusion process, we use some inhomogeneous inequalities, including the time-dependent Poincaré and Log-Sobolev inequalities. This extends some results for the time-homogeneous diffusion processes in Cattiaux and Guillin (2009).

1. Introduction

In the recent work (Cattiaux and Guillin, 2009), Cattiaux and Guillin obtained, for homogeneous continuous Markov processes, the rate of convergence to the invariant measure in total variation distance, by using functional inequalities. The authors considered the ergodic *diffusion processes*, which means that the infinitesimal generator L satisfies the hypotheses: for $\Psi \in C^2$ and smooth f,

$$L\Psi(f) = \frac{\partial \Psi}{\partial x}(f)Lf + \frac{1}{2}\frac{\partial^2 \Psi}{\partial x^2}(f)\Gamma(f), \quad \text{and} \quad \Gamma(\Psi(f)) = (\Psi'(f))^2\Gamma(f),$$

where Γ is the corresponding *carré du champ* operator associated to L, see Definition 2.2.

In the present paper, we extend their results in two directions. Firstly, we consider the general but *reversible* Markov processes. Secondly, we consider the *inhomogeneous* diffusion process. The inhomogeneity brings more flexibility and complex behaviors might appear consequently, see Saloff-Coste and Zúñiga (2007, 2009, 2011) for inhomogeneous finite Markov chains, and Arnaudon et al. (2008), Cheng and Zhang (2016) for inhomogeneous diffusion. In both cases, we give the

* Corresponding author.
E-mail addresses: maoyh@bnu.edu.cn (Y.-H. Mao), lipingxu@mail.bnu.edu.cn (L. Xu), zhangming0408@mail.bnu.edu.cn (M. Zhang), zhangyh@bnu.edu.cn (Y.-H. Zhang).

https://doi.org/10.1016/j.spl.2018.01.011
0167-7152/

rates of convergence in the total variation distance by using corresponding inhomogeneous functional inequalities. Indeed, to get the convergence rates, we will also use two elementary lemma, Lemmas 3.2 and 3.3 for the time homogeneous and time-inhomogeneous case, respectively. There is no doubt that we follow closely the idea of Cattiaux and Guillin (2009). However, in the inhomogeneous case, we keep the same assumptions as in Cattiaux and Guillin (2009). We deal with the diffusion processes, without the reversible assumption. Since the generators are time-dependent, the reversible probability measures should be time-dependent, which is not helpful for the convergence. See Saloff-Coste and Zúñiga (2007) for example of inhomogeneous random walks, which admit the reversible probability measures time by time. We use the reversibility instead of the property of diffusion, while in the time-inhomogeneous case what we get is new. Finally, let us mention (Collet and Malrieu, 2008), where a commutation relation in the time-inhomogeneous case was given to guarantee the time-dependent Φ-Sobolev inequalities.

The paper is organized as follows. In Section 2, we present the main results of our paper on the time-homogeneous reversible Markov process and time-inhomogeneous diffusion process. In Section 3, we give the proofs of main theorems.

2. Main results

First, let us introduce some basic notations and definitions, used through this paper. We use C or C' for a positive constant, which may vary from line to line, and will indicate (in the subscript) the dependent parameters. We denote the entropy and oscillation of a (probability density) function f, respectively by

$$\mathrm{Ent}_\mu(f) = \int f \log f \, d\mu \quad \text{and} \quad \mathrm{Osc}(f) = \sup_{x,y} |f(x) - f(y)| \, .$$

We always let ψ be a fixed convex function in $C^2(\mathbb{R}^+)$, which satisfies

- ψ is locally uniform convex, i.e. for any $A > 0$, $\inf_{[0,A]} \psi'' > 0$.
- $\psi(1) = 0$ and $\lim_{x\to\infty} \psi(x)/x = \infty$.

Definition 2.1. Given a measurable space $(E, \mathcal{E})$, let $\mathscr{P}(E)$ be the total of probability measures. For any $\mu_1, \mu_2 \in \mathscr{P}(E)$, we define the total variation distance between μ_1 and μ_2 by

$$\|\mu_1 - \mu_2\|_{TV} = \sup_{A\in\mathcal{E}} \{ |\mu_1(A) - \mu_2(A)| \}.$$

Our results below differ from that in Cattiaux and Guillin (2009) in two folds. Firstly, in Section 2.1, we give the convergence rate in the total variation distance for time-homogeneous *reversible* Markov processes, in contract the diffusion processes in Cattiaux and Guillin (2009). Secondly, in Section 2.2, we give the merging in the total variation distance for *time-inhomogeneous* diffusion processes. Here the merging means the convergence from different initial laws, and we refer to Saloff-Coste and Zúñiga (2007, 2009, 2011) for the case of finite Markov chains

2.1. Time-homogeneous reversible Markov processes

Assume the time-homogeneous Markov process $(X_t, \mathbb{P}_x)$ is reversible with respect to its stationary distribution μ, for more details, see Wang (2004). Let $(P_t)_{t\geq 0}$ be the associated semi-group with the infinitesimal generator $(L, D(L))$ in $L^2(\pi)$, that is

$$\int_E f P_t g d\mu = \int_E g P_t f d\mu, \quad f, g \in L^2(\mu) \quad \text{and} \quad \int_E f L g d\mu = \int_E g L f d\mu, \quad f, g \in D(L).$$

For any probabilistic density h w.r.t. μ, the law of X_t with initial distribution $hd\mu$ is given by $P_t h d\mu$. We also use notations $I_\psi(t, h) = \int \psi(P_t h) d\mu$ and $I_\psi(h) = I_\psi(0, h) = \int \psi(h) d\mu$. The typical examples of reversible Markov processes include the ergodic reversible diffusion processes in Wang (2004), reversible Markov jump processes in Chen (2004). We also refer to these references for the functional inequalities.

Definition 2.2. We define a bilinear form

$$\Gamma(f, g) = \frac{1}{2}(L(fg) - gLf - fLg), \quad f, g \in D(L),$$

the homogeneous carré du champ operator associated to the semigroup or the infinitesimal generator. We use abbreviation $\Gamma(f) = \Gamma(f, f)$.

The carré du champ operator plays a crucial role in the analysis and geometry of Markov processes. The simplest example is that the Euclidean Laplacian Δ gives rise to the standard carré du champ operator $\Gamma(f, g) = \nabla f \cdot \nabla g$, the usual scalar product of the gradients. For further information, we refer to the original paper (Bakry and Émery, 1984) and a recent monograph (Bakry et al., 2014).

Now, we can state our first main theorem for the time-homogeneous reversible Markov processes.

Theorem 2.3. *The following two statements are equivalent.*

(I) *For all probabilistic density h with $I_\psi(h) < \infty$ and all time t, we have exponential decay*

$$I_\psi(t,h) \le e^{-t/C_\psi} I_\psi(h).$$

(II) *For all "nice" probabilistic density h, in the sense of Cattiaux et al. (2007) and Bakry et al. (2014), the following I_ψ-inequality holds,*

$$\int \psi(h) d\mu \le C_\psi \int \Gamma(\psi'(h), h) d\mu. \tag{1}$$

Moreover, if any of the conditions is satisfied, we have

$$\|P_t h\mu - \mu\|_{TV} \le C'_\psi\, e^{-t/2C_\psi} \sqrt{I_\psi(h)}. \tag{2}$$

In order to get the convergence for fixed h, condition (II) can be relaxed to hold for all $P_t h$, not necessary for all probability densities; The I_ψ-inequality is just the Poincaré inequality when $\psi(x) = (x-1)^2$, and is the entropy inequality studied in Zhang and Mao (2001) when $\psi(x) = x\log x$. These special cases of Theorem 2.3 are then the generalization of Cattiaux and Guillin (2009, Theorem 2.1 and 2.7).

We also consider the weak I_ψ-inequality condition as in Röckner and Wang (2001) and Cattiaux and Guillin (2009), by using the truncation argument. Then the total variation can be controlled by two terms, with the first one controlled by $\sqrt{I_\psi(t, h\wedge K)}$ for $K > 0$, and the second one controlled from the De La Vallée-Poussin theorem. Finally we have the following theorem.

Theorem 2.4. *If there exist a non-increasing positive function β_ψ and a positive functional N with $N(P_t^* h) \le N(h)$, such that the following weak I_ψ-inequality*

$$\int \psi(h)\, d\mu \le \beta_\psi(s) \int \Gamma(\psi'(h), h)\, d\mu + sN(h) \tag{3}$$

holds for all probability density h and $s > 0$, then we have

$$I_\psi(t,h) \le \xi(t)(I_\psi(h) + N(h)),$$

where

$$\xi(t) := \inf\{s > 0,\ \beta_\psi(s)\log(\frac{1}{s}) \le t\}. \tag{4}$$

Moreover, we have the total variation estimate

$$\|P_t h\mu - \mu\|_{TV} \le C_\psi \sqrt{\xi(t)\left(I_\psi(h\wedge K) + N(h\wedge K)\right)} + \frac{2\int h\varphi(h)\, d\mu}{\varphi(K)}, \quad \forall K > 0, \tag{5}$$

for some nonnegative φ growing to infinity.

Similarly, when $\psi(x) = (x-1)^2$, the weak I_ψ-inequality becomes the weak Poincaré inequality, while the case $\psi(x) = x\log x$ corresponds to the weak relative entropy inequality. In the later case with $N(h) = \mathrm{Osc}^2(\sqrt{h})$, we also give the ε-decay of entropy from the weak relative entropy inequality.

Theorem 2.5. *If there exists a non-increasing positive β such that the weak relative entropy inequality*

$$\mathrm{Ent}_\mu(h) \le \beta(s) \int \Gamma(\log h, h)\, d\mu + s\, Osc^2(\sqrt{h})$$

holds for all probability density h and $s > 0$, then we have ε-decay of entropy

$$\mathrm{Ent}_\mu(P_t h) \le (\frac{1}{e} + \varepsilon)\, \xi(\varepsilon, t)\, Osc^2(\sqrt{h}), \quad \forall \varepsilon > 0, \tag{6}$$

where

$$\xi(\varepsilon, t) := \inf\{s > 0,\ \beta(s)\log(\frac{\varepsilon}{s}) \le 2t\}.$$

Moreover, we have the total variation estimate

$$\|P_t h\mu - \mu\|_{TV} \le \frac{4\int h\varphi(h)\, d\mu}{(\varphi\circ\tilde{\varphi}^{-1})\left(\frac{\sqrt{2}\int h\varphi(h)\, d\mu}{(\frac{1}{e}+\varepsilon)^{\frac{1}{2}}\sqrt{\xi(\varepsilon,t)}}\right)}, \quad \text{with } \tilde{\varphi}(x) = \sqrt{x}\varphi(x), \tag{7}$$

for some nonnegative φ growing to infinity.

2.2. Time-inhomogeneous diffusion processes

We study in this subsection the merging in the total variation distance for inhomogeneous diffusion processes with different initial laws. Following Collet and Malrieu (2008), on $\mathbb{R}^d$, we consider the following in-homogeneous SDE:

$$X_{r,t}^x = x + \int_r^t b(s, X_{r,s}^x)ds + \sqrt{2}\int_r^t \sigma(s, X_{r,s}^x)dB_s, \quad t \geq r \geq 0,$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^d$. Let L_t be the in-homogeneous generator:

$$L_t f(x) = \sum_{i,j+1}^{d} a_{ij}(t,x)\partial_{ij} f(x) + \sum_{i=1}^{d} b_i(t,x)\partial f(x),$$

where $a = \sigma\sigma^T$ and b are smooth functions of (t,x), and let $(P_{s,t})_{0\leq s\leq t}$ be the associated semigroup, which satisfies the semigroup property: for any $s \leq r \leq t$, $P_{s,t} = P_{s,r}P_{r,t}$, $P_{t,t} = I$, and the Kolmogorov equations:

$$\partial_s P_{s,t} f = -L_s P_{s,t} f, \quad \partial_t P_{s,t} f = P_{s,t} L_t f \,.$$

We will always denote by $\mu_t = \mu_0 P_{0,t} = \mu_s P_{s,t}$, the law of X_t.

In Collet and Malrieu (2008), the log-Sobolev inequality is studied for the in-homogeneous diffusion processes, and we also refer to Cheng (2017) for a recent study on diffusions on manifolds with time-depending metrics. For more examples of in-homogeneous diffusion processes see Collet and Malrieu (2008), Cheng (2017) and references therein.

Definition 2.6 (*Collet and Malrieu, 2008*)**.** Define a bilinear form

$$\Gamma_t(f,g) = \frac{1}{2}\big(L_t(fg) - fL_t(g) - gL_t(f)\big), \quad f, g \in D(L_t),$$

the inhomogeneous carré du champ operator associated to the semigroup or infinitesimal generator. We use abbreviation $\Gamma_t(f) = \Gamma_t(f,f)$.

Assume that L_t is a diffusion operator, that is, for any $\Phi \in C^2$,

$$L_t(\Phi(f)) = \Phi'(f)\, L_t(f) + \Phi''(f)\, \Gamma_t(f), \quad f \in D(L_t) \tag{8}$$

and

$$\Gamma_t(\Phi(f), g) = \Phi'(f)\Gamma_t(f,g) \quad f, g \in D(L_t). \tag{9}$$

Given two different initial laws μ and ν, we assume μ and μ_t are absolutely continuous w.r.t. ν and ν_t, with Radon–Nikodym derivatives

$$f_0 = \frac{d\mu}{d\nu}, \quad \text{and} \quad f_t = \frac{d\mu_t}{d\nu_t}.$$

Note that if μ is absolutely continuous w.r.t ν, then μP_t is absolutely continuous w.r.t νP_t. Indeed, let $P_t(x, dy)$ be the probability transition function. By assuming that $\nu P_t(A) = \int \nu(dx)P_t(x,A) = 0$ for some measurable set A, we see $P_t(x,A) = 0$, ν-a.s., so that $\mu P_t(A) = \int_E f_0(x)\nu(dx)P_t(x,A) = 0$. We also use notations

$$I_{\psi,t} = \int \psi(f_t)\, d\nu_t, \quad \text{and} \quad I_\psi = I_{\psi,0}\,.$$

Now, we can state our second main result for continuous inhomogeneous diffusion Markov processes.

Theorem 2.7. *Fix $t > 0$. If the following $I_{\psi,s}$-inequality*

$$\int \psi(f_s)\, d\nu_s \leq C(s)\int \psi''(f_s)\Gamma_s(f_s)\, d\nu_s \tag{10}$$

holds for $0 \leq s \leq t$ and μ_s and ν_s, then we have exponential-like decay

$$I_{\psi,s} \leq e^{-\int_0^s \frac{dr}{C(r)}}\, I_\psi$$

holds for $0 \leq s \leq t$. Moreover, we have the total variation estimate

$$\|\mu_s - \nu_s\|_{TV} \leq C'_\psi\, e^{-\frac{1}{2}\int_0^s \frac{dr}{C(r)}}\, \sqrt{I_\psi}\,. \tag{11}$$

The constant $C(s)$ is dependent of s and ψ. Similar to the homogeneous case in Cattiaux and Guillin (2009), when $\psi(x) = (x-1)^2$, the $I_{\psi,t}$-inequality is just the Poincaré inequality while the case that $\psi(x) = x\log x$ corresponds to the Log-Sobolev inequality. We give these two special cases in the following two corollaries without proof.

Corollary 2.8. *Fix $t > 0$. If Poincaré inequality*

$$\nu_s(f_s^2) - (\nu_s f)^2 \le C(s)\, \nu_s(\Gamma_s(f_s))$$

holds for $0 \le s \le t$ and μ_s and ν_s, then

$$\nu_s(f_s^2) - (\nu_s f)^2 \le e^{-\int_0^s \frac{2dr}{C(r)}} (\nu(f_0^2) - 1), \quad s \le t.$$

Moreover, we have the total variation estimate

$$\|\mu_s - \nu_s\|_{TV} \le e^{-\int_0^s \frac{dr}{C(r)}} \sqrt{\nu(f_0^2) - 1}.$$

Corollary 2.9. *Fix $t > 0$. If Log-Sobolev inequality*

$$\mathrm{Ent}_{\nu_s}(f_s) \le C(s)\, \nu_s\left(\frac{\Gamma_s(f_s)}{f_s}\right)$$

holds for $0 \le s \le t$ and μ_s and ν_s, then

$$\mathrm{Ent}_{\nu_s}(f_s) \le e^{-\int_0^s \frac{dr}{C(r)}} \mathrm{Ent}_\nu(d\nu/d\mu), \quad s \le t.$$

Moreover, we have the total variation estimate

$$\|\mu_s - \nu_s\|_{TV} \le \sqrt{2}\, e^{-\frac{1}{2}\int_0^s \frac{dr}{C(r)}} \sqrt{\mathrm{Ent}_\nu(d\nu/d\mu)}\,.$$

Similar to the time-homogeneous case (Theorem 2.3), we can also rewrite Theorem 2.7 in the following equivalence form.

Theorem 2.10. *Given $t > 0$, the following are equivalent.*

(I) *For all $0 \le s_0 \le s \le t$, we have exponential decay*

$$I_{\psi,s} \le e^{-\int_{s_0}^s \frac{dr}{C(r)}} I_{\psi,s_0}. \tag{12}$$

(II) *For all $0 \le s \le t$, $I_{\psi,s}$-inequality holds for μ_s and ν_s, i.e.*

$$\int \psi(f_s)\, d\nu_s \le C(s) \int \psi''(f_s)\Gamma_s(f_s)\, d\nu_s. \tag{13}$$

Moreover, if any of the above two conditions is satisfied, then we have the total variation estimate

$$\|\mu_s - \nu_s\|_{TV} \le C'_\psi\, e^{-\frac{1}{2}\int_{s_0}^s \frac{dr}{C(r)}} \sqrt{I_{\psi,s_0}}\,, \quad \forall\, 0 \le s_0 \le s \le t\,.$$

3. Proofs of the main results

In this section, we will give the proofs of our main results in Section 2. In Section 3.1, we consider the time-homogeneous case, and in Section 3.2, we consider the time-inhomogeneous case. We recall the function ψ being a fixed convex function in $C^2(\mathbb{R}^+)$, which satisfies

- ψ is locally uniform convex, i.e. for any $A > 0$, $\inf_{[0,A]} \psi'' > 0$.
- $\psi(1) = 0$ and $\lim_{x\to\infty} \psi(x)/x = \infty$.

3.1. Time-homogeneous case

Lemma 3.1 (*Pinsker Inequality (Pinsker, 1964; Cattiaux and Guillin, 2009)*)**.** *Given a pair of probability measures $(\mathbb{P}, \mathbb{Q})$, there exists $C_\psi > 0$, such that*

$$\|\mathbb{P} - \mathbb{Q}\|_{TV} \le C_\psi \sqrt{\int \psi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) d\mathbb{P}}.$$

Lemma 3.2.

$$\frac{d}{dt} I_\psi(t,h) = -\int \Gamma(\psi'(P_t h), P_t h)\, d\mu.$$

Proof. By definition, we have

$$\frac{d}{dt}I_\psi(t,h)=\int \psi'(P_th)\frac{d}{dt}P_th\,d\mu=\int \psi'(P_th)LP_th\,d\mu.$$

Using the symmetry of L, we obtain

$$\begin{aligned}\frac{d}{dt}I_\psi(t,h)&=\int \big(L(\psi'(P_th)P_th)-P_thL(\psi'(P_th))-2\Gamma(\psi'(P_th),P_th)\big)\,d\mu\\&=\int \psi'(P_th)P_thL^*1\,d\mu-\int \psi'(P_th)L^*(P_th)\,d\mu-2\int \Gamma(\psi'(P_th),P_th)\,d\mu\\&=\int \psi'(P_th)P_thL1\,d\mu-\int \psi'(P_th)L(P_th)\,d\mu-2\int \Gamma(\psi'(P_th),P_th)\,d\mu\\&=-\int \psi'(P_th)L(P_th)\,d\mu-2\int \Gamma(\psi'(P_th),P_th)\,d\mu\\&=-\frac{d}{dt}I_\psi(t,h)-2\int \Gamma(\psi'(P_th),P_th)\,d\mu,\end{aligned}$$

which gives the conclusion. □

Now, we can proceed to prove the main theorems in Section 2.1.

Proof of Theorem 2.3. The total variation estimate (2) comes immediately from the Pinsker inequality (Lemma 3.1) and condition (I). Then it suffices to prove the equivalence of the two conditions.

(I) $\Longrightarrow$ (II): If (I) holds, then

$$\frac{I_\psi(t,h)-I_\psi(h)}{t}\le\frac{e^{-t/C_\psi}-1}{t}I_\psi(h).$$

Taking $t\to 0$ yields

$$\frac{d}{dt}\Big|_{t=0}I_\psi(t,h)\le-\frac{I_\psi(h)}{C_\psi},$$

which gives by Lemma 3.2 that

$$\int \psi(h)d\mu\le C_\psi\int \Gamma(\psi'(h),h)d\mu.$$

(II) $\Longrightarrow$ (I): If (II) holds, then replacing h by P_th, we have

$$I_\psi(t,h)=\int \psi(P_th)d\mu\le C_\psi\int \Gamma(\psi'(P_th),P_th)d\mu,$$

which gives by Lemma 3.2 that

$$I_\psi(t,h)\le-C_\psi\frac{d}{dt}I_\psi(t,h).$$

By using the Gronwall inequality,

$$I_\psi(t,h)\le e^{-t/C_\psi}I_\psi(h).\quad □$$

Proof of Theorem 2.4. If the weak I_ψ-inequality (3) holds, then replacing h by P_th, we have

$$\begin{aligned}I_\psi(t,h)&\le\beta_\psi(s)\int \Gamma(\psi'(P_th),P_th)d\mu+sN(P_th)\\&=-\beta_\psi(s)\frac{d}{dt}I_\psi(t,h)+sN(P_th)\\&\le-\beta_\psi(s)\frac{d}{dt}I_\psi(t,h)+sN(h),\end{aligned}$$

and therefore

$$\frac{1}{\beta_\psi(s)}I_\psi(t,h)+\frac{d}{dt}I_\psi(t,h)\le\frac{s}{\beta_\psi(s)}N(h).$$

Multiplying both sides with $e^{t/\beta_\psi(s)}$ and integrating on t over $[0,t]$, we have

$$e^{t/\beta_\psi(s)}I_\psi(t,h)-I_\psi(h)\le sN(h)(e^{t/\beta_\psi(s)}-1).$$

That is,

$$I_\psi(t,h) \le e^{-t/\beta_\psi(s)} I_\psi(h) + sN(h), \quad \forall s > 0.$$

By the definition (4) of $\xi(t)$ and minimizing on the right-hand side for $e^{-t/\beta_\psi(s)} \le s$, we get the decay estimate

$$I_\psi(t,h) \le \xi(t)(I_\psi(h) + N(h)).$$

By the truncation argument, Pinsker inequality (Lemma 3.1) and De La Vallée-Poussin theorem, we finally get the total variation estimate (5):

$$\begin{aligned}
\|P_t h\mu - \mu\|_{TV} &= \|P_t h - 1\|_{L^1_\mu} \\
&= \int \left|P_t(h\wedge K) - \int (h\wedge K)d\mu\right| d\mu + 2\int (h-K)\mathbf{1}_{h\ge K} d\mu \\
&\le C_\psi \sqrt{I_\psi(t, h\wedge K)} + \frac{2}{\varphi(K)} \int h\varphi(h)d\mu \\
&\le C_\psi \sqrt{\xi(t)\big(I_\psi(h\wedge K) + N(h\wedge K)\big)} + \frac{2}{\varphi(K)} \int h\varphi(h)d\mu,
\end{aligned}$$

for all $K > 0$ and some non-negative φ growing to infinity. □

Proof of Theorem 2.5. The proof of the first part is similar to that of Cattiaux et al. (2007, Proposition 4.1). So, it suffices to prove the total variation estimate (7). Again by using the truncation argument, Pinsker inequality (Lemma 3.1), De La Vallée-Poussin theorem, and the entropy decay (6) in the first part, we get

$$\begin{aligned}
\|P_t h\mu - \mu\|_{TV} &= \|P_t h - 1\|_{L^1_\mu} \\
&= \int \left|P_t(h\wedge K) - \int (h\wedge K)d\mu\right| d\mu + 2\int (h-K)\mathbf{1}_{h\ge K} d\mu \\
&\le \sqrt{2}\int (h\wedge K)d\mu \sqrt{\int P_t\Big(\frac{h\wedge K}{\mu(h\wedge K)}\Big)\log\Big(P_t\Big(\frac{h\wedge K}{\mu(h\wedge K)}\Big)\Big)d\mu} \\
&\quad + 2\int (h-K)\mathbf{1}_{h\ge K} d\mu \\
&\le \sqrt{2}\Big(\frac{1}{e}+\varepsilon\Big)^{\frac12}\sqrt{\xi(\varepsilon,t)}\int (h\wedge K)d\mu \,\mathrm{Osc}\Big(\sqrt{P_t\Big(\frac{h\wedge K}{\mu(h\wedge K)}\Big)}\Big) \\
&\quad + 2\,\frac{\int h\varphi(h)d\mu}{\varphi(K)} \\
&\le \sqrt{2}\Big(\frac{1}{e}+\varepsilon\Big)^{\frac12}\sqrt{\xi(\varepsilon,t)}\sqrt{K} + 2\,\frac{\int h\varphi(h)d\mu}{\varphi(K)}.
\end{aligned}$$

On the right-hand side by letting the two terms equal, i.e. taking

$$K = \tilde{\varphi}^{-1}\left(\frac{\sqrt{2}\int h\varphi(h)\,d\mu}{(\frac{1}{e}+\varepsilon)^{\frac12}\sqrt{\xi(\varepsilon,t)}}\right),$$

we finally get the desired total variation estimate (7). □

3.2. Time-inhomogeneous case

In this section we will give the proof of our second main result in this paper, for the inhomogeneous case in Section 2.2. We focus on the proof of Theorem 2.7, and at first we give the following lemma concerning the time derivative.

Lemma 3.3. *For any measurable function h,*

$$\nu_t\Big(h\frac{d}{dt}f_t\Big) = \nu_t\big(f_t L_t h - L_t(hf_t)\big).$$

Proof. By Radon–Nikodym theorem and differentiation on time, we have that

$$\frac{d}{dt}\int h d\mu_t = \frac{d}{dt}\int P_{0,t} h d\mu = \int P_{0,t} L_t h d\mu = \int L_t h d\mu_t = \int f_t L_t h d\nu_t$$

and

$$\frac{d}{dt}\int hf_t d\nu_t = \frac{d}{dt}\int P_{0,t}(hf_t)d\nu = \int\big(P_{0,t}L_t(hf_t) + P_{0,t}(h\frac{d}{dt}f_t)\big)d\nu = \int\big(L_t(hf_t) + h\frac{d}{dt}f_t\big)d\nu_t,$$

which gives finally

$$\int f_t L_t h d\nu_t = \int\big(L_t(hf_t) + h\frac{d}{dt}f_t\big)\, d\nu_t\,.$$

The formula is then proved by definition. □

This lemma is elementary, but critical to handle the inhomogeneous case. So far, we have not seen such kind of lemma used before to study the inhomogeneous asymptotic behavior. We remark that in the homogeneous case, Lemma 3.3 is reduced into the trivial one:

$$\nu(h\frac{d}{dt}f_t) = \nu(f_t Lh).$$

This is always true by letting $\nu_t \equiv \nu$ and $L_t = L$ by homogeneity. On the other hand, even if L_t admits a reversible probability measure, say $\tilde{\nu}_t$, what we need is an equation above concerning ν_t, rather than $\tilde{\nu}_t$.

Proof of Theorem 2.7. By differentiation, we have

$$\begin{aligned}\frac{d}{dt}\int \psi(f_t)d\nu_t &= \frac{d}{dt}\int P_{0,t}\psi(f_t)d\nu \\ &= \int\big(P_{0,t}L_t\psi(f_t) + P_{0,t}\psi'(f_t)\frac{d}{dt}f_t\big)d\nu \\ &= \nu_t\big(L_t\psi(f_t) + \psi'(f_t)\frac{d}{dt}f_t\big).\end{aligned} \tag{14}$$

Taking $h = \psi'(f_t)$ in Lemma 3.3, we have

$$\nu_t(\psi'(f_t)\frac{d}{dt}f_t) = \nu_t(f_t L_t\psi'(f_t) - L_t(\psi'(f_t)f_t)).$$

After putting this into (14) and using the diffusion condition (8), (9) and the $I_{\psi,t}$-inequality (10), we have

$$\begin{aligned}\frac{d}{dt}\int \psi(f_t)d\nu_t &= \nu_t\big(L_t\psi(f_t) + f_tL_t\psi'(f_t) - L_t(\psi'(f_t)f_t)\big) \\ &= \nu_t\big(\psi'(f_t)L_tf_t + f_tL_t\psi'(f_t) - L_t(\psi'(f_t)f_t) + \psi''(f_t)\Gamma_t(f_t)\big) \\ &= \nu_t(-2\Gamma_t(\psi'(f_t), f_t) + \psi''(f_t)\Gamma_t(f_t)) \\ &= -\,\nu_t\big(\psi''(f_t)\Gamma_t(f_t)\big) \\ &\le -\,\frac{1}{C(t)}\int \psi(f_t)d\nu_t.\end{aligned}$$

By using the Gronwall inequality, we finally get the exponential decay

$$\int \psi(f_s)\, d\nu_s \le e^{-\int_0^s \frac{dr}{C(r)}}\int \psi(f_0)\, d\nu.$$

The total variation control (11) is again an obvious application of the Pinsker inequality (Lemma 3.1). □

The proof of the equivalent form, Theorem 2.10, is similar. We make the differential in the exponential decay to get $I_{\psi,s}$-inequality.

Proof of Theorem 2.10. Similar to the proof of Theorem 2.3, we now prove the other side: (I) $\Longrightarrow$ (II). In the previous proof of Theorem 2.7, we know that

$$\frac{d}{dt}I_{\psi,t} = -\nu_t\big(\psi''(f_t)\,\Gamma_t(f_t)\big)\,. \tag{15}$$

From the decay inequality (12), we have

$$\frac{I_{\psi,s} - I_{\psi,s_0}}{s - s_0} \le \frac{e^{-\int_{s_0}^s \frac{dr}{C(r)}} - 1}{s - s_0}\, I_{\psi,s_0}$$

and after taking $s \to s_0$ and using (15),

$$\int \psi''(f_{s_0})\Gamma_{s_0}(f_{s_0})\, d\nu_{s_0} = -\frac{d}{ds}\Big|_{s=s_0} I_{\psi,s} \ge \frac{1}{C(s_0)}\, I_{\psi,s_0}\,.$$

This gives the inequality (13). □

Acknowledgments

Research was supported in part by NSFC (No. 11131003, 11571043, 11626245, 11771047). We thank a lot for a referee's constructive suggestions and valuable comments. The second author is also very grateful to the School of Mathematical Sciences, Beijing Normal University, for providing so nice work environment during this project.

References

Arnaudon, M., Coulibaly, K.A., Thalmaier, A., 2008. Brownian motion with respect to a metric depending on time: definition, existence and applications to Ricci flow. C. R. Math. Acad. Sci. Paris 346 (13–14), 773–778.

Bakry, D.,Émery, M., 1984. Hypercontractivité de semi-groupes de diffusion. C. R. Acad. Sci. Paris Sér. I Math. 299 (15), 775–778.

Bakry, D., Gentil, I., Ledoux, M., 2014. Analysis and Geometry of Markov Diffusion Operators. In: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, p. xx+552.

Cattiaux, P., Gentil, I., Guillin, A., 2007. Weak logarithmic Sobolev inequalities and entropic convergence. Probab. Theory Related Fields 139 (3–4), 563–603.

Cattiaux, P., Guillin, A., 2009. Trends to equilibrium in total variation distance. Ann. Inst. H. Poincaré Probab. Stat. 45 (1), 117–145.

Chen, M.-F., 2004. From Markov Chains to Non-Equilibrium Particle Systems. World Scientific.

Cheng, L.-J., 2017. Diffusion semigroup on manifolds with time-dependent metrics. Forum Math. 29 (4), 775–798.

Cheng, L.-J., Zhang, K., 2016. Reflecting diffusion processes on manifolds carrying geometric flow. J. Theoret. Probab. 1–35.

Collet, J.-F., Malrieu, F., 2008. Logarithmic Sobolev inequalities for inhomogeneous Markov semigroups. ESAIM Probab. Stat. 12, 492–504.

Pinsker, M.S., 1964. Information and Information Stability of Random Variables and Processes. Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, p. xii+243. Translated and edited by Amiel Feinstein.

Röckner, M., Wang, F.-Y., 2001. Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. J. Funct. Anal. 185 (2), 564–603.

Saloff-Coste, L., Zúñiga, J., 2007. Convergence of some time inhomogeneous Markov chains via spectral techniques. Stochastic Process. Appl. 117 (8), 961–979.

Saloff-Coste, L., Zúñiga, J., 2009. Merging for time inhomogeneous finite Markov chains. I. Singular values and stability. Electron. J. Probab. 14, 1456–1494.

Saloff-Coste, L., Zúñiga, J., 2011. Merging for inhomogeneous finite Markov chains, Part II: Nash and log-Sobolev inequalities. Ann. Probab. 39 (3), 1161–1203.

Wang, F.-Y., 2004. Functional Inequalities, Markov Processes and Spectral Theory. Sci. Press, Beijing.

Zhang, S., Mao, Y., 2001. Exponential convergence rate in Boltzmann-Shannon entropy. Sci. China Ser. A 44 (3), 280–285.