Concentration of Additive Functionals for Markov Processes and Applications to Interacting Particle Systems

Frank Redig* Florian Völlering†

February 26, 2010

We consider additive functionals of Markov processes in continuous time with general (metric) state spaces. We derive concentration bounds for their exponential moments and moments of finite order. Applications include diffusions, interacting particle systems and random walks. In particular, for the symmetric exclusion process we generalize large deviation bounds for occupation times to general local functions. The method is based on coupling estimates and not spectral theory, hence reversibility is not needed. We bound the exponential moments(or the moments of finite order) in terms of a so-called coupled function difference, which in turn is estimated using the generalized coupling time. Along the way we prove a general relation between the contractivity of the semigroup and bounds on the generalized coupling time.

Keywords: Markov processes, Polish state space, additive functionals, coupling, generalized coupling time, concentration estimates, exclusion process
AMS classification: 60J25, 60J55, 60F10

1 Introduction

The study of concentration properties of additive functionals of Markov processes is the subject of many recent publications, see e.g. [9], [4]. This subject is strongly connected to functional inequalities such as the Poincaré and log-Sobolev inequality,as well as to the concentration of measure phenomenon [6]. In the present paper we

*University of Nijmegen, IMAPP, Heyendaalse weg 135, 6525 AJ Nijmegen, The Netherlands
redig@math.leidenuniv.nl
†Leiden University, Mathematical Institute, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
fvolleri@math.leidenuniv.nl

consider concentration properties of a general class of additive functionals of the form $\int_0^T f_t(X)_t \, dt$ in the context of continuous-time Markov processes on a Polish space. The simplest and classical case is where $f_t = f$ does not depend on time.

Our approach is based on coupling ideas. More precisely, we estimate exponential moments or k-th order moments using the so-called coupled function difference which is estimated in terms of a so-called generalized coupling time, a generalization of the concept used in [3]. Our method covers several cases such as diffusion processes, jump processes, random walks and interacting particle systems. The example of random walk shows that for unbounded state spaces, the concentration inequalities depend on which space the functions f_t belong to. The simple example of independent Poissonian random variables shows that for unbounded state space with continuous-time, we cannot expect in general Gaussian concentration bounds for the exponential moment. The main application to the exclusion process, which has slow relaxation to equilibrium and therefore does not satisfy any functional inequality such as e.g. log-Sobolev (in infinite volume), shows the full power of the method. Besides, we give a one-to-one correspondence between the exponential contraction of the semigroup and the fact that the generalized coupling time is bounded by the metric. For discrete state spaces, this means that the semigroup is exponentially contracting if and only if the generalized coupling time is bounded.

Our paper is organized as follows: in section 2 we prove our concentration inequalities in the general context of a continuous-time Markov process on a metric space. We derive estimated for exponential moments and moments of finite order. In section 3 we study the generalized coupling time and its relation to contractivity of the semigroup. Section 4 is devoted to examples. Section 5 deals with the symmetric exclusion process.

2 General Notation

Let $\mathbb{X} = (X_t)_{t\geq 0}$ be a Markov process in the polish state space E. Denote by $\mathbb{P}_x$ its associated measure on the path space of cadlag trajectories $D_{[0,\infty[}(E)$ started in $x \in E$ and with

$$\mathfrak{F}_t := \sigma\left\{X_s; 0 \leq s \leq t\right\}, \quad t \geq 0,$$

the canonical filtration. Its infinitesimal generator we denote by A and the corresponding semigroup with $(S_t)_{t\geq 0}$. We denote by $\mathbb{E}_x^{\mathbb{X}}$ the expectation with respect to the measure $\mathbb{P}_x$. For ν a probability measure on E, we define $\mathbb{E}_\nu^{\mathbb{X}} := \int \mathbb{E}_x^{\mathbb{X}} \, \nu(dx)$, i.e. expectation in the process starting from ν.

3 Concentration inequalities

The content of this section is to derive concentration inequalities on functionals of the form

$$F(\mathbb{X}) := \int_0^\infty f_t(X_t)\,dt, \quad f_t : E \to \mathbb{R}, \tag{1}$$

with respect to the Markov process $\mathbb{X}$. The most familiar case is when F is of the form

$$\int_0^T f(X_t)\,dt,$$

i.e. $f_t \equiv f$ for $t \le T$ and $f_t \equiv 0$ for $t > T$. We now formulate the conditions on the family of functions f_t we need in what follows.

Definition 3.1. *We say the family of functions $\{f_t, t \ge 0\}$ is k-regular for $k \in \mathbb{N}$, if:*

a) The f_t are Borel measurable and the integral F exists for all paths;

b) $\mathbb{E}_x^{\mathbb{X}}\mathbb{E}_x^{\mathbb{Y}} \sup\limits_{0\le s\le\epsilon} |\, f_{t+s}(X_s) - f_{t+s}(Y_s)\,|^k < \infty$ for $t > 0$ and $x \in E$ arbitrary and $\epsilon > 0$ small enough;

c) There exists a function $r : E \to \mathbb{R}$ with $\mathbb{E}_x^{\mathbb{X}} r(X_\epsilon)^k < \infty$ so that

$$\sup_{t\ge 0} \mathbb{E}_x^{\mathbb{X}} \int_0^\infty |\, f_{t+\epsilon+u}(X_u) - f_{t+u}(X_u)\,|\, du \le \epsilon r(x).$$

Remark If $F(\mathbb{X}) = \int_0^T f(X_t)\,dt$, then $\mathbb{E}_x^{\mathbb{X}} \sup\limits_{0\le t\le T+\epsilon_0} |\, f(X_t)\,|^k < \infty$ for some $\epsilon_0 > 0$ implies conditions b) and c) of the k-regularity.

The technique to obtain concentration inequalities for functionals of the form (1) is to use a telescoping approach where one conditions on $\mathfrak{F}_t$, i.e., where we average $F(\mathbb{X})$ under the knowledge of the path of the Markov process $\mathbb{X}$ up to time t.

Definition 3.2. *For $0 \le s \le t$, define the increments*

$$\Delta_{s,t} := \mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_t] - \mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_s]$$

and the initial increment

$$\Delta_{\star,0} := \mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_0] - \mathbb{E}_\nu[F(\mathbb{X})],$$

which depends on the initial distribution ν.

The basic property of the increments is the decomposition $\Delta_{s,u} = \Delta_{s,t} + \Delta_{t,u}$ for $s < t < u$. Also, we have

$$\mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_T] - \mathbb{E}_{\nu}[F(\mathbb{X})] = \Delta_{\star,0} + \Delta_{0,T},$$

where we have to use $\Delta_{\star,0}$ to accommodate for the initial distribution ν. To better work with the increment $\Delta_{s,t}$, we will rewrite it in a more complicated but also more useful way.

Definition 3.3. *Given the family of functions $\{f_t : t \geq 0\}$, the coupled function difference is defined as*

$$\Phi_t(x,y) := \int_0^\infty \mathbb{E}_x^{\mathbb{X}} f_{t+u}(X_u) - \mathbb{E}_y^{\mathbb{Y}} f_{t+u}(Y_u)\,du.$$

Remark We call Φ_t the coupled function difference because later we will see that we need estimates on $|\,\Phi_t\,|$, and for a coupling $\widehat{\mathbb{E}}$ of $\mathbb{X}$ starting in x and y we have the estimate

$$\Phi_t(x,y) \leq \int_0^\infty \widehat{\mathbb{E}}_{x,y}^{\mathbb{X},\mathbb{Y}}\,|\,f_{t+u}(X_u) - f_{t+u}(Y_u)\,|\,du.$$

In the next lemma we express the increments $\Delta_{s,t}$ in terms of the coupled function difference Φ_t.

Lemma 3.4. *Let $\mathbb{E}_x^{\mathbb{Z}.+s}$ denote the expectation under $\mathbb{P}_x(d\mathbb{Z})$ when $\mathbb{Z}$ is started in x at time s, i.e. $\mathbb{Z} = (Z_t)_{t\geq s}$ and $Z_s = x$ a.s. With this notation,*

$$\Delta_{s,t} = \mathbb{E}_{X_s}^{\mathbb{Z}.+s}\left(\int_s^t f_u(X_u) - f_u(Z_u)\,du\right) + \mathbb{E}_{X_s}^{\mathbb{Z}.+s}\Phi_t(X_t, Z_t).$$

Proof First, we note that

$$\mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_t] = \int_0^t f_u(X_u)\,du + \mathbb{E}_{X_t}^{\mathbb{Y}.+t}\int_t^\infty f_u(Y_u)\,du,$$

and by the Markov property

$$\mathbb{E}\,[F(\mathbb{X})|\mathfrak{F}_s] = \int_0^s f_u(X_u)\,du + \mathbb{E}_{X_s}^{\mathbb{Z}.+s}\left(\int_s^t f_u(Z_u)\,du + \mathbb{E}_{Z_t}^{\mathbb{Y}.+t}\int_t^\infty f_u(Y_u)\,du\right).$$

Hence,

$$\begin{aligned}
\Delta_{s,t} &= \int_0^t f_u(X_u)\,du + \mathbb{E}_{X_t}^{\mathbb{Y}^{(1)}_{\cdot+t}} \int_t^\infty f_u(Y^{(1)}_u)\,du \\
&\quad - \int_0^s f_u(X_u)\,du - \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \int_s^t f_u(Z_u)\,du - \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \mathbb{E}_{Z_t}^{\mathbb{Y}^{(2)}_{\cdot+t}} \int_t^\infty f_u(Y^{(2)}_u)\,du \\
&= \int_s^t f_u(X_u)\,du - \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \int_s^t f_u(Z_u)\,du \\
&\quad + \mathbb{E}_{X_t}^{\mathbb{Y}^{(1)}_{\cdot+t}} \int_t^\infty f_u(Y^{(1)}_u)\,du - \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \mathbb{E}_{Z_t}^{\mathbb{Y}^{(2)}_{\cdot+t}} \int_t^\infty f_u(Y^{(2)}_u)\,du \\
&= \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \left(\int_s^t f_u(X_u) - f_u(Z_u)\,du \right) + \mathbb{E}_{X_s}^{\mathbb{Z}_{\cdot+s}} \Phi_t(X_t, Z_t).
\end{aligned}$$

□

The following lemma is crucial to obtain the concentration inequalities of Theorems 3.6 and 3.9 bellow. It expresses conditional moments of the increments in terms of the coupled function difference.

Lemma 3.5. *Fix $k \in \mathbb{N}$, $k \geq 2$. Assume that the f_t are k-regular and let $\Phi_t(\cdot, x)^k$ be in the domain of the generator A for all $x \in E$. Then*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbb{E}\left[\Delta^k_{t,t+\epsilon} \,\middle|\, \mathfrak{F}_t\right] = (A(\Phi_t(\cdot, X_t)^k))(X_t).$$

Proof We will use the following elementary fact repetitively. For $k \geq 2$, if $|\, b_\epsilon \,| \leq \epsilon \bar{b}_\epsilon$ and $\sup_{0 \leq \epsilon \leq \epsilon_0} \mathbb{E}\, \bar{b}^k_\epsilon < \infty$, then

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbb{E}\,(a_\epsilon + b_\epsilon)^k = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbb{E}\, a^k_\epsilon. \tag{2}$$

By Lemma 3.4,

$$\Delta_{t,t+\epsilon} = \mathbb{E}_{X_t}^{Z_{\cdot+t}} \left(\int_t^{t+\epsilon} f_s(X_s) - f_s(Z_s)\,ds \right) + \mathbb{E}_{X_t}^{Z_{\cdot+t}} \Phi_{t+\epsilon}(X_{t+\epsilon}, Z_{t+\epsilon}).$$

First, we show that we can neglect the first term. Indeed,

$$\left| \mathbb{E}_x^{\mathbb{Z}} \int_0^\epsilon f_{t+s}(X_s) - f_{t+s}(Z_s)\,ds \right| \leq \epsilon\, \mathbb{E}_x^{\mathbb{Z}} \sup_{0 \leq s \leq \epsilon} |\, f_{t+s}(X_s) - f_{t+s}(Z_s) \,|,$$

we can use part b) of the k-regularity to apply fact (2) and get

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\Delta^k_{t,t+\epsilon}\,|\,\mathfrak{F}_t\right] = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\left(\mathbb{E}^{Z_{\cdot+t}}_{X_t}\Phi_{t+\epsilon}(X_{t+\epsilon}, Z_{t+\epsilon})\right)^k\,\middle|\,\mathfrak{F}_t\right].$$

Next, by writing $\Phi_{t+\epsilon} = \Phi_t + (\Phi_{t+\epsilon} - \Phi_t)$, we will show that the difference can be neglected in the limit $\epsilon \to 0$. To this end, we observe that

$$\begin{aligned}|\,\Phi_{t+\epsilon}(x,y) - \Phi_t(x,y)\,| \le &\int_0^\infty \mathbb{E}^{\mathbb{X}}_x\,|\,f_{t+\epsilon+u}(X_u) - f_{t+u}(X_u)\,|\,du\\ &+\int_0^\infty \mathbb{E}^{\mathbb{Y}}_y\,|\,f_{t+\epsilon+u}(Y_u) - f_{t+u}(Y_u)\,|\,du.\end{aligned}$$

Part c) of the $k-$regularity condition allows us to invoke fact (2) again to obtain

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\Delta^k_{t,t+\epsilon}\,|\,\mathfrak{F}_t\right] = \frac{1}{\epsilon}\mathbb{E}\left[\left(\mathbb{E}^{Z_{\cdot+t}}_{X_t}\Phi_t(X_{t+\epsilon}, Z_{t+\epsilon})\right)^k\,\middle|\,\mathfrak{F}_t\right].$$

Finally, since $\Phi_t(y,\cdot)$ is in the domain of the generator, we can use the representation

$$\mathbb{E}^{\mathbb{Z}}_x\Phi_t(y, Z_\epsilon) = \Phi_t(y,x) + \epsilon A\Phi_t(y,\cdot)(x) + o(\epsilon)$$

and replace $Z_{t+\epsilon}$ by X_t by applying fact (2) for a third time. Now, the desired result is immediately achieved:

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\Delta^k_{t,t+\epsilon}\,|\,\mathfrak{F}_t\right] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}^{\widetilde{\mathbb{X}}}_{X_t}\left(\Phi_t(\widetilde{X}_\epsilon, X_t)\right)^k\\ &= A\Phi_t(\cdot, X_t)^k(X_t).\end{aligned}$$

□

We can now state our first main theorem, which is a bound of the exponential moment of $F(\mathbb{X})$ in terms of the coupled function difference Φ_t.

Theorem 3.6. *Assume that for all $k \in \mathbb{N}$, the f_t are k-regular and $\Phi_t(\cdot,x)^k \in dom(A)$ for all $x \in E$. Then, for any distributions μ and ν on E,*

$$\log \mathbb{E}_\mu\left[e^{F(\mathbb{X})-\mathbb{E}_\nu F(\mathbb{X})}\right] \le \log(c_0) + \int_0^\infty \sup_{x\in E}\sum_{k=2}^\infty \frac{1}{k!}(A(\Phi^k_t(\cdot,x)))(x)\,dt,$$

$$\log \mathbb{E}_\mu\left[e^{F(\mathbb{X})-\mathbb{E}_\nu F(\mathbb{X})}\right] \ge \log(c_0) + \int_0^\infty \inf_{x\in E}\sum_{k=2}^\infty \frac{1}{k!}(A(\Phi^k_t(\cdot,x)))(x)\,dt,$$

where the influence of the distributions μ and ν is only felt in the factor

$$c_0 = \mathbb{E}^X_\mu e^{\mathbb{E}^Y_\nu \Phi_0(X,Y)}.$$

Remark If $H_t : E \times E$ is an upper bound on $|\,\Phi_t\,|$ and $H_t(x,x) = 0$ for all $x \in E$, then the upper bound of the theorem remains valid if Φ_t is replaced by H_t. In particular, if $f_t \equiv f \mathbb{1}_{t \leq T}$, $H_t := |\,\Phi_0\,| \,\mathbb{1}_{t \leq T}$ serves as a good initial estimate to obtain the upper bound

$$\log \mathbb{E}_\mu \left[e^{F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})} \right] \leq \log(c_0) + T \sup_{x \in E} \sum_{k=2}^{\infty} \frac{1}{k!} A \,|\,\Phi_0\,|^k\,(\cdot, x)(x).$$

Further estimates on $|\,\Phi_0\,|$ specific to the particular process can can then be used without the need to keep a dependence on t.

Proof Define

$$\Psi(t) := \mathbb{E}_\mu \left[e^{\Delta_{\star,0} + \Delta_{0,t}} \right].$$

We see that for $\epsilon > 0$,

$$\begin{aligned}\Psi(t+\epsilon) - \Psi(t) &= \mathbb{E}_\mu \left(e^{\Delta_{\star,0} + \Delta_{0,t}} \mathbb{E} \left[e^{\Delta_{t,t+\epsilon}} - 1 \,\middle|\, \mathfrak{F}_t \right] \right) \\ &= \mathbb{E}_\mu \left(e^{\Delta_{\star,0} + \Delta_{0,t}} \mathbb{E} \left[e^{\Delta_{t,t+\epsilon}} - \Delta_{t,t+\epsilon} - 1 \,\middle|\, \mathfrak{F}_t \right] \right),\end{aligned}$$

where we used the fact that $\mathbb{E}\,[\Delta_{t,t+\epsilon} | \mathfrak{F}_t] = 0$. Hence, using Lemma 3.5, we can calculate the derivative of Ψ:

$$\Psi'(t) = \mathbb{E}_\mu \left(e^{\Delta_{\star,0} + \Delta_{0,t}} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t(\cdot, X_t)^k))(X_t) \right).$$

To get upper or lower bounds on Ψ', we move the sum out of the expectation as a supremum or infimum. Just continuing with the upper bound, as the lower bound is analogue,

$$\Psi'(t) \leq \Psi(t) \sup_{x \in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot, x)))(x).$$

After dividing by $\Psi(t)$ and integrating, we get

$$\ln \Psi(T) - \ln \Psi(0) \leq \int_0^T \sup_{x \in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot, x)))(x)\, dt,$$

which leads to

$$\lim_{T \to \infty} \Psi(T) = \mathbb{E}_\mu \left[e^{F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})} \right] \leq \Psi(0) e^{\int_0^\infty \sup_{x \in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot, x)))(x)\, dt}.$$

The value of $c_0 = \Psi(0) = \mathbb{E}_\mu e^{\Delta_{\star,0}}$ is obtained from the identity

$$\Delta_{\star,0} = \mathbb{E}_\nu^Y \Phi_0(X_0, Y). \qquad \square$$

Corollary 3.7. *Assume that* $F(\mathbb{X}) = \int_0^T f(X_t)\,dt$*, the conditions of Theorem 3.6 are satisfied, and* $\sup_{x\in E} A\,|\,\Phi_0\,|^k(\cdot,x)(x) \le c_1 c_2^k$ *for some* $c_1, c_2 > 0$*. Then, for any initial condition* $x \in E$*,*

$$\mathbb{P}_x(F(\mathbb{X}) - \mathbb{E}_x F(\mathbb{X}) > x) \le e^{\frac{-\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}}}.$$

Proof By Markov's inequality,

$$\begin{aligned}\mathbb{P}_x(F(\mathbb{X}) - \mathbb{E}_x F(\mathbb{X}) > x) &\le \mathbb{E}_x e^{\lambda F(\mathbb{X}) - \mathbb{E}_x \lambda F(\mathbb{X})} e^{-\lambda x} \\ &\le e^{Tc_1 \sum_{k=2}^{\infty} c_2^k - \lambda x},\end{aligned}$$

where the last line is the result from Theorem 3.6. Through optimizing λ, the exponent becomes

$$\frac{x}{c_2} - (Tc_1 + \frac{x}{c_2})\log(\frac{x}{Tc_1c_2} + 1).$$

To show that his term is less than $\frac{-\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}}$, we first rewrite it as the following inequality:

$$\log(\frac{x}{Tc_1c_2} + 1) \ge \frac{\frac{\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}} + \frac{x}{c_2}}{Tc_1 + \frac{x}{c_2}}.$$

Through comparing the derivatives, one concludes that the left hand side is indeed bigger than the right hand side. □

In applications one tries to find good estimates of Φ_t. When looking at the examples in Section 5, finding those estimates is where the actual work lies. In the case where the functions f_t are Lipschitz continuous with respect to a suitably chosen (semi)metric ρ, the problem can be reduced to questions about the generalized coupling time h, which is defined and discussed in detail in Section 4. In case that the exponential moment of $F(\mathbb{X}) - \mathbb{E}\,F(\mathbb{X})$ does not exist or the bound obtained from Theorem 3.6 is not useful, we turn to moment bounds. This is the content of the next theorem.

Lemma 3.8. *Assume that the* f_t *are* 2*-regular and* $\Phi_t^2(\cdot,x)$ *is in the domain of the generator* A*. Then the predictable quadratic variation of the martingale* $(\Delta_{0,t})_{t\ge 0}$ *is*

$$\langle \Delta_{0,\cdot} \rangle_t = \int_0^t A\Phi_s^2(\cdot, X_s)(X_s)\,ds.$$

Proof We have, using Lemma 3.5 for $k = 2$,

$$\frac{d}{dt}\langle \Delta_{0,\cdot} \rangle_t = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\mathbb{E}\left[\Delta_{t,t+\epsilon}^2 \,\middle|\, \mathfrak{F}_t\right] = A\Phi_t(\cdot, X_t)^2(X_t).$$ □

Theorem 3.9. *Let the functions f_t be 2-regular and $\Phi_t^2(\cdot, x)$ in the domain of the generator A. Then*

$$(\mathbb{E}_\mu \,|\, F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})\,|^p)^{\frac{1}{p}} \leq C_p \left[\left(\mathbb{E}_\mu \left(\int_0^\infty A\Phi_t^2(\cdot, X_t)(X_t)\, dt \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \right. \tag{3a}$$

$$\left. + \left(\mathbb{E}_\mu \left(\sup_{t \geq 0} |\, \Phi_t(X_t, X_{t-})\,| \right)^p \right)^{\frac{1}{p}} \right] \tag{3b}$$

$$+ \left(\mathbb{E}_\mu^X \left(\left| \mathbb{E}_\nu^Y \Phi_0(X, Y)\,\right| \right)^p \right)^{\frac{1}{p}} \tag{3c}$$

where the constant C_p only depends on p and behaves like $p/\log p$ as $p \to \infty$.

Proof By the triangle inequality,

$$(\mathbb{E}_\mu \,|\, F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})\,|^p)^{\frac{1}{p}} \leq (\mathbb{E}_\mu \,|\, \Delta_{0,\infty}\,|^p)^{\frac{1}{p}} + (\mathbb{E}_\mu \,|\, \Delta_{\star,0}\,|^p)^{\frac{1}{p}}\,.$$

Since $(\Delta_{0,t})_{t \geq 0}$ is a square integrable martingale starting at 0, a version of Rosenthal's inequality([10], Theorem 1) implies

$$(\mathbb{E}_\mu \,|\, \Delta_{0,T}\,|^p)^{\frac{1}{p}} \leq C_p \left[\left(\mathbb{E}_\mu \langle \Delta_{0,\cdot} \rangle_T^{\frac{p}{2}} \right)^{\frac{1}{p}} + \left(\mathbb{E}_\mu \sup_{0 \leq t \leq T} |\, \Delta_{0,t} - \Delta_{0,t-}\,|^p \right)^{\frac{1}{p}} \right].$$

Applying Lemma 3.8 to rewrite the predictable quadratic variation $\langle \Delta_{0,\cdot} \rangle_T$ and Lemma 3.4 to rewrite $\Delta_{t-,t}$, we end with the first two terms of our claim after letting $T \to \infty$. The last term is just a different way of writing $\Delta_{\star,0}$:

$$(\mathbb{E}_\mu \,|\, \Delta_{\star,0}\,|^p)^{\frac{1}{p}} = \left(\mathbb{E}_\mu^X \,|\, \mathbb{E}_\nu^Y \Phi_0(X, Y)\,|^p \right)^{\frac{1}{p}}.$$

□

Let us discuss the meaning of the three terms appearing on the right hand side in Theorem (3.9).

a) The first term gives the contribution, typically of order $T^{\frac{p}{2}}$, that one expects even in the simplest case of processes with independent increments.

   E.g. if μ is an invariant measure and $F(\mathbb{X}) = \int_0^T f(X_t)\, dt$, then

$$\mathbb{E}_\mu^{\mathbb{X}} \left(\int_0^\infty A\Phi_t^2(\cdot, X_t)(X_t)\, dt \right)^{\frac{p}{2}} \leq T^{\frac{p}{2}} \mathbb{E}_\mu^X \left(A\Phi_0^2(\cdot, X)(X) \right)^{\frac{p}{2}}.$$

   In many cases (see examples below), $\mathbb{E}_\mu^X \left(A\Phi_0^2(\cdot, X)(X) \right)^{\frac{p}{2}}$ can be treated as a constant, i.e., not depending on T. There are however relevant examples where this factor blows up as $T \to \infty$.

b) The second term measures rare events of possibly large jumps where it is very difficult to couple. If the process $\mathbb{X}$ has continuous paths, this term is not present. Usually this term is or bounded or is of lower order than the first term as $T \to \infty$.

c) The third term has only the hidden time dependence of Φ_0 on T. It measures the intrinsic variation given the starting measures μ and ν and it vanishes if and only if $\mu = \nu = \delta_x$.

It is also interesting to note that the estimate is sharp for small T: If one chooses $F(\mathbb{X}) = \frac{1}{T}\int_0^T f(X_t)\,dt$ and looks at the limit as $T \to 0$, the first two terms disappear and the third one becomes $(\mathbb{E}_\mu^X \left| f(X) - \mathbb{E}_\nu^Y f(Y) \right|^p)^{\frac{1}{p}}$, which is also the limit of the left hand side.

4 Generalized coupling time

As we had seen in the previous section, in the results we need estimates on Φ_t. We can obtain these if we know more about the coupling behaviour of the underlying process $\mathbb{X}$. To characterize this coupling behaviour, we will look at how close we can get two versions of the process started at different points measured with respect to a distance.

Let $\rho : E \times E \to [0, \infty]$ be a lower semi-continuous semi-metric. With respect to this semi-metric, we define

$$\| f \|_{Lip} := \inf \{ r \geq 0 \mid f(x) - f(y) \leq r\rho(x,y) \; \forall \, x, y \in E \} ,$$

the Lipschitz-seminorm of f corresponding to ρ. Now we introduce the main objects of study in this section.

Definition 4.1. *a) The optimal coupling distance at time t is defined as*

$$\rho_t(x,y) := \inf_{\pi \in \mathfrak{P}(\delta_x S_t, \delta_y S_t)} \int \rho(x',y')\,\pi(dx'\,dy'),$$

where the infimum ranges over the set of all possible couplings with marginals $\delta_x S_t$ and $\delta_y S_t$, i.e., the distribution of X_t started from x or y.

b) The generalized coupling time is defined as

$$h(x,y) := \int_0^\infty \rho_t(x,y)\,dt.$$

Now that we have introduced the generalized coupling time, a first application is a version of Theorem 3.6 for Lipschitz continuous functions:

Corollary 4.2. *Assume the functions f_t are Lipschitz continuous with respect to a semi-metric ρ, and that the conditions of Theorem 3.6 hold true. Then*

$$\mathbb{E}_\mu \left[e^{F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})} \right] \leq c_0 e^{\sum_{k=2}^{\infty} \frac{c_k}{k!} \sup_{x \in E} (A(h^k(\cdot,x)))(x)},$$

where

$$c_0 = \mathbb{E}_{\mu}^{x} e^{\sup\limits_{t\geq 0} \| f_t \|_{Lip} \mathbb{E}_{\nu}^{y} h(x,y)},$$

$$c_k = \int_0^\infty \sup_{t\geq 0} \| f_t \|_{Lip}^k \, dt.$$

In particular, if $f_t \equiv f$ *for* $t \leq T$ *and* $f_t \equiv 0$ *for* $t > T$*, then*

$$c_0 \leq \mathbb{E}_{\mu}^{x} e^{\| f \|_{Lip} \mathbb{E}_{\nu}^{y} h(x,y)},$$

$$c_k \leq T \| f \|_{Lip}^k .$$

Remark If $\overline{h}$ is an upper bound on the generalized coupling time h with $\overline{h}(x,x) = 0$, then the result holds true with h replaced by $\overline{h}$.

Proposition 4.3. *The optimal coupling distance* ρ_t *has the dual formulation*

$$\rho_t(x,y) = \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(y)).$$

Proof By the Kantorovich-Rubinstein theorem ([11], Theorem 1.14), we have

$$\begin{aligned} \inf_{\pi} \int \rho \, d\pi &= \sup_{\| f \|_{Lip}=1} \left[\int f \, d(S_t \delta_x) - \int f \, d(S_t \delta_y) \right] \\ &= \sup_{\| f \|_{Lip}=1} \left[(S_t f)(x) - (S_t f)(y) \right]. \end{aligned}$$

□

Also, it is easy to see that the semi-metric properties of ρ translate to ρ_t and thereby to the generalized coupling time h.

Proposition 4.4. *Both the optimal coupling distance* ρ_t *and the generalized coupling time* h *are semi-metrics.*

Proof We only have to prove the semi-metric properties of ρ_t, they translate naturally to h via integration.

Obviously, $\rho_t(x,x) = 0$ and $\rho_t(x,y) = \rho_t(y,x)$ is true for all $x, y \in E$ by definition of ρ_t. For the triangle inequality, we use the dual representation:

$$\begin{aligned} \rho_t(x,y) &= \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(y)) \\ &= \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(z) + S_t f(z) - S_t f(y)) \leq \rho_t(x,z) + \rho_t(y,z) \end{aligned}$$

□

A first result is a simple estimate on the decay of the semigroup S_t in terms of the optimal coupling distance.

Proposition 4.5. *Let μ be a stationary measure of the semigroup S_t. Then*

$$\| S_t f - \mu(f) \|_{L^p(\mu)} \leq \| f \|_{Lip} \left(\int \mu(dx) \left(\int \mu(dy) \rho_t(x,y) \right)^p \right)^{\frac{1}{p}}.$$

Remark When we choose the metric ρ to be the discrete metric $\mathbb{1}_{x \neq y}$ (a choice we can make even in a non-discrete setting), we can estimate $\rho_t(x,y)$ by $\widehat{\mathbb{P}}_{x,y}(\tau > t)$, the probability that the coupling time $\tau = \inf \left\{ t \geq 0 \mid X_s^1 = X_s^2 \ \forall s \geq t \right\}$ is larger than t in an arbitrary coupling $\widehat{\mathbb{P}}_{x,y}$ of the Markov process started in x and y. In this case, the result of Proposition 4.5 reads

$$\| S_t f - \mu(f) \|_{L^p(\mu)} \leq \| f \|_{osc} \left(\int \mu(dx) \left(\int \mu(dy) \widehat{\mathbb{P}}_{x,y}(\tau > t) \right)^p \right)^{\frac{1}{p}},$$

where $\| f \|_{osc} = \sup_{x,y}(f(x) - f(y))$ is the oscillation norm. Note that this can also be gained from the well-known coupling inequality

$$\| \delta_x S_t - \delta_y S_t \|_{TVar} \leq 2 \widehat{\mathbb{P}}_{x,y}(\tau > t).$$

Proof of Proposition 4.5 First,

$$\begin{aligned} | S_t f(x) - \mu(f) | &= | S_t f(x) - \mu(S_t f) | \\ &= \left| \mathbb{E}_x f(X_t) - \int \mu(dy) \mathbb{E}_y f(Y_t) \right| \\ &\leq \int \mu(dy) \, | \mathbb{E}_x f(X_t) - \mathbb{E}_y f(Y_t) | \\ &\leq \int \mu(dy) \, \| f \|_{Lip} \, \rho_t(x,y). \end{aligned}$$

This estimate can be applied directly to get the result:

$$\begin{aligned} \| S_t f - \mu(f) \|_{L^p(\mu)} &= \left(\int \mu(dx) \, | S_t f(x) - \mu(f) |^p \right)^{\frac{1}{p}} \\ &\leq \| f \|_{Lip} \left(\int \mu(dx) \left(\int \mu(dy) \rho_t(x,y) \right)^p \right)^{\frac{1}{p}}. \end{aligned}$$

□

The above result did not use the fact that S_t is a semigroup. When we use that fact, we can improve estimates considerably. The price is that from now on, ρ has to be a metric, and this metric must be compatible with the Markov process, which we will define a little bit later under the notion of contraction with respect to this metric. The aim is to show how the uniform boundedness of the generalized coupling time implies an exponential decay of the semigroup (S_t) in the Lipschitz seminorm. To this end, we need the following lemma:

Lemma 4.6. *Under the condition that ρ is a metric,*

$$\sup_{x\neq y} \frac{\rho_t(x,y)}{\rho(x,y)} = \| S_t \|_{Lip} .$$

Proof By the representation of the optimal coupling distance in Proposition 4.3,

$$\begin{aligned}\sup_{x\neq y} \frac{\rho_t(x,y)}{\rho(x,y)} &= \sup_{x\neq y} \sup_{\| f \|_{Lip}=1} \frac{S_t f(x) - S_t f(y)}{\rho(x,y)} \\ &= \sup_{\| f \|_{Lip}=1} \| S_t f \|_{Lip} = \| S_t \|_{Lip} .\end{aligned}$$

□

Definition 4.7. *We say that the process $\mathbb{X}$ acts as a* contraction for the distance ρ *if*

$$\rho_t(x,y) \leq \rho(x,y) \quad \forall\, t \geq 0, \tag{4}$$

or equivalently,

$$\| S_t \|_{Lip} \leq 1 \quad \forall\, t \geq 0.$$

This property is sufficient to show that the process is contracting the distance monotonely:

Lemma 4.8. *Assume that the process $\mathbb{X}$ acts as a contraction for the distance. Then*

$$\rho_{t+s}(x,y) \leq \rho_t(x,y) \quad \forall\, x, y \in E,\, s, t \geq 0.$$

Proof Using the dual representation,

$$\begin{aligned}\rho_{t+s}(x,y) &= \sup_{\| f \|_{Lip}=1} [S_{t+s} f(x) - S_{t+s} f(y)] \\ &= \sup_{\| f \|_{Lip}\leq 1} [S_t(S_s f)(x) - S_t(S_s f)(y)].\end{aligned}$$

By our assumption, the set of functions f with $\| f \|_{Lip} \leq 1$ are a subset of the set of functions f with $\| S_s f \|_{Lip} \leq 1$. Hence,

$$\begin{aligned}\rho_{t+s}(x,y) &\leq \sup_{f:\| S_s f \|_{Lip}\leq 1} [S_t(S_s f)(x) - S_t(S_s f)(y)] \\ &\leq \sup_{\| g \|\leq 1} [S_t g(x) - S_t g(y)] = \rho_t(x,y).\end{aligned}$$

□

With this property in mind, we can show the main theorem of this section.

Theorem 4.9. *Assume that ρ is a metric and that the process $\mathbb{X}$ acts as a contraction for the distance. Then the fact that the generalized coupling time h is bounded by the metric ρ is equivalent to the fact that the semigroup (S_t) is exponentially contracting. More precisely, for $\alpha > 1$ arbitrary,*

a) $\forall\, x, y \in E :\ h(x,y) \leq M\rho(x,y) \quad \Rightarrow \quad \forall\, t \geq M\alpha :\ \|\, S_t \,\|_{Lip} \leq \frac{1}{\alpha}$;

b) $\|\, S_T \,\|_{Lip} \leq \frac{1}{\alpha} \quad \Rightarrow \quad \forall\, x, y \in E :\ h(x,y) \leq \frac{\alpha T}{\alpha - 1}\rho(x,y)$.

Proof a) For $x, y \in E$, set

$$T_{x,y} := \inf\left\{ t \geq 0 \,\middle|\, \rho_t(x,y) \leq \frac{1}{\alpha}\rho(x,y) \right\}.$$

Then,

$$M\rho(x,y) \geq h(x,y) = \int_0^\infty \rho_t(x,y)\, dt \geq \int_0^{T_{x,y}} \rho_t(x,y)\, dt \geq T_{x,y}\frac{1}{\alpha}\rho(x,y).$$

Therefore $T_{x,y}$ is bounded by $M\alpha$. By Lemma 4.8, $\rho_t(x,y) \leq \rho_{T_{x,y}}(x,y)$ for all $t \geq T_{x,y}$. Hence $\rho_{M\alpha}(x,y) \leq \frac{1}{\alpha}\rho(x,y)$ uniformly, which implies $\|\, S_{M\alpha} \,\|_{Lip} \leq \frac{1}{\alpha}$.

b) Since $\rho_t(x,y) \leq \rho(x,y)\, \|\, S_t \,\|_{Lip}$,

$$\begin{aligned} h(x,y) &= \int_0^\infty \rho_t(x,y)\, dt \leq \rho(x,y) \int_0^\infty \|\, S_t \,\|_{Lip}\, dt \\ &\leq \rho(x,y) T \sum_{k=0}^\infty \|\, S_T \,\|_{Lip}^k \leq \frac{\alpha T}{\alpha - 1}\rho(x,y). \end{aligned}$$

□

When we apply this theorem to an arbitrary Markov process where we use the discrete distance, we get the following corollary:

Corollary 4.10. *The following two statements are equivalent:*

a) The generalized coupling time with respect to the discrete metric $\rho(x,y) = \mathbb{1}_{x \neq y}$ is uniformly bounded, i.e.

$$h(x,y) \leq M \quad \forall x, y \in E;$$

b) The semigroup is eventually contractive in the oscillation (semi)norm, i.e. $\|\, S_T \,\|_{osc} < 1$ for some $T > 0$.

Remark Theorem 4.9 actually gives us more information, namely how the constants M and T can be related to each other.

Proof Since obviously $\sup_{x \neq y} \rho_t(x,y) \le 1$, the process $\mathbb{X}$ acts as a contraction for the discrete distance and the result follows from Theorem 4.9, where we also use the fact that in the case of the discrete metric, $\| \cdot \|_{Lip} = \| \cdot \|_{osc}$. $\square$

Since Theorem 4.9 part a) implies that $\| S_t \|_{Lip}$ decays exponentially fast, it is of interest to get the best estimate on the speed of decay, which is the content of the following proposition:

Proposition 4.11. *Assume that ρ is a metric, the process $\mathbb{X}$ acts as a contraction for the distance and the generalized coupling time h satisfies $h(x,y) \le M\rho(x,y)$. Then*

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t \|_{Lip} \le -\frac{1}{M}.$$

Proof The proof uses the same structure as the proof of part a) in Theorem 4.9. First, fix ϵ between 0 and $\frac{1}{M}$. Then define

$$T_{x,y} = \inf \left\{ t > 0 \,\middle|\, \rho_t(x,y) \le \rho(x,y) e^{-(\frac{1}{M}-\epsilon)t} \right\}.$$

By our assumption,

$$M\rho(x,y) \ge h(x,y) \ge \rho(x,y) \int_0^{T_{x,y}} e^{-(\frac{1}{M}-\epsilon)t}\, dt = M\rho(x,y) \frac{1 - e^{-(\frac{1}{M}-\epsilon)T_{x,y}}}{1 - M\epsilon}.$$

Since the fraction on the right hand side becomes bigger than 1 if $T_{x,y}$ is too large, there exists an uniform upper bound $T(\epsilon)$ on $T_{x,y}$. Hence, for all $t \ge T(\epsilon)$, $\| S_t \|_{Lip} \le e^{-(\frac{1}{M}-\epsilon)t}$, which of course implies $\lim_{t\to\infty} \frac{1}{t} \| S_t \|_{Lip} \le -\frac{1}{M} + \epsilon$. By sending ϵ to 0, we finish our proof. $\square$

Again, we apply this result to the discrete metric to see what it contains.

Corollary 4.12. *Let $\widehat{\mathbb{P}}_{x,y}$ be a coupling of the process $\mathbb{X}$ started in x resp. y, and denote with $\tau := \inf \left\{ t \ge 0 \,\middle|\, X_s^1 = X_s^2 \; \forall s \ge t \right\}$ the coupling time. Set $M := \sup_{x,y \in E} \widehat{\mathbb{E}}_{x,y} \tau$. Then*

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t \|_{osc} \le -\frac{1}{M}.$$

Equivalently, for $f \in L^\infty$,

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t f - \mu(f) \|_\infty \le -\frac{1}{M} \| f \|_{osc},$$

where μ is the unique stationary distribution of $\mathbb{X}$.

Remarks a) If the the Markov process $\mathbb{X}$ is also reversible, then the above result extends to L^1 and hence to any L^p, where the spectral gap is then also at least of size $\frac{1}{M}$.

b) As an additional consequence, when a Markov process can be uniformly coupled, i.e. $\sup_{x,y\in E} \widehat{\mathbb{E}}_{x,y}\tau \leq M < \infty$ for a coupling $\widehat{\mathbb{E}}$, then there exists (a possibly different) coupling $\widetilde{\mathbb{E}}_{x,y}$, so that $\sup_{x,y\in E} \widetilde{\mathbb{E}}_{x,y} e^{\lambda\tau} < \infty$ for all $\lambda < \frac{1}{M}$.

5 Examples

5.1 Diffusions with a strictly convex potential

Let V be a twice continuously differentiable function on the real line with $V'' \geq c > 0$ and $\int e^{-V(x)}dx = Z_V < \infty$. To the potential V is associated the Gibbs measure

$$\mu_V(dx) = \frac{1}{Z_V} e^{-V(x)}dx$$

and a Markovian diffusion

$$dX_t = -V'(X_t) + \sqrt{2}dW_t$$

with μ_V as reversible measure.

To estimate the optimal coupling distance ρ_t at time t(see Definition 4.1), we couple two versions of the diffusion, X_t^x started in x and X_t^y started in y, by using the same Brownian motion $(W_t)_{t\geq 0}$. Then the difference process $X_t^x - X_t^y$ is deterministic, $x < y$ implies $X_t^x < X_t^y$ and by the convexity assumption

$$d(X_t^y - X_t^x) = -(V'(X_t^y) - V'(X_t^x)) \leq -c(X_t^y - X_t^x).$$

Using Gronwall's Lemma, we obtain the estimate

$$\rho_t(x, y) \leq |\, x - y \,| \, e^{-ct}$$

on the optimal coupling distance. By integration, the generalized coupling time h has the estimate $h(x, y) \leq \frac{1}{c} |\, x - y \,|$. As a consequence, Proposition 4.11 implies

$$\lim_{t\to\infty} \log \| \, S_t \, \|_{Lip} \leq -c.$$

Since the generator A of the diffusion is

$$A = \frac{d^2}{dx^2} - V' \cdot \frac{d}{dx},$$

we have

$$A\left(\frac{1}{c}|\cdot - x\,|\right)^k (x) = \begin{cases} \frac{2}{c^2}, & k = 2, \\ 0, & k > 2. \end{cases}$$

Therefore, for $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz-continuous, we can use Corollary 4.2 to get the estimate

$$\mathbb{E}_{\nu_1}\left[e^{\int_0^T f(X_t)\,dt-\mathbb{E}_{\nu_2}\int_0^T f(X_t)\,dt}\right] \le c_{\nu_1,\nu_2}e^{T\frac{\|f\|_{Lip}^2}{c^2}}, \tag{5}$$

with the dependence on the distributions ν_1 and ν_2 given by

$$c_{\nu_1,\nu_2} = \mathbb{E}^x_{\nu_1}e^{\mathbb{E}^y_{\nu_2}\frac{\|f\|_{Lip}}{c}|x-y|}.$$

Remark a) An alternative proof that strict convexity is sufficient for (5) to be true can be found in [12]. A proof via the log-Sobolev inequality can be found in [6]. Hence the result is of no surprise, but the method of obtaining it is new.

b) This example demonstrates nicely how in the case of diffusions the higher moments of $Ah^k(\cdot,x)(x)$ can disappear because the generalized coupling time is bounded by a multiple of the initial distance.

c) The generalization to higher dimensions under strict convexity is straightforward.

5.2 Independent Poisson random variables

This example nicely illustrates problems and effects occurring in even simple Markov processes when time is continuous and the state space unbounded(in this case $\mathbb{N}$).

Let $(X_n)_{n\in\mathbb{N}}$ be iid. with distribution $\mu = \text{poisson}(1)$, $(N_t)_{t\ge 0}$ be a standard Poisson process independent of (X_n) and $X_t := X_{N_t}$. We are considering the function $F(\mathbb{X}) = \int_0^T \lambda X_t\,dt$ with $\lambda > 0$.

First, let us look at the generalized coupling time. By the definition of the process,

$$S_tf(x) = \mathbb{P}(N_t = 0)f(x) + \mathbb{P}(N_t > 0)\mathbb{E}\,f(X_1)$$

and hence

$$|\,S_tf(x) - S_tf(y)\,| = \mathbb{P}(N_t = 0)\,|\,f(x) - f(y)\,|.$$

Given this, the generalized coupling time is

$$\begin{aligned} h(x,y) &= \int_0^\infty \rho_t(x,y)\,dt = \int_0^\infty \sup_{\|f\|_{Lip}\le 1}(S_tf(x) - S_tf(y)) \\ &= \int_0^\infty \sup_{\|f\|_{Lip}\le 1} \mathbb{P}(N_t = 0)\,|\,f(x) - f(y)\,|\,dt \\ &= |\,x-y\,|\,\mathbb{E}\,exp(1) = |\,x-y\,|. \end{aligned}$$

The generator A of the Markov process $\mathbb{X}$ is

$$Af(x) = \mathbb{E}_\mu f - f(x) = e^{-1}\sum_{y=0}^\infty \frac{1}{y!}[f(y) - f(x)]$$

and hence

$$Ah(\cdot,x)^k(x) = \mathbb{E}_\mu^Y |Y - x|^k,$$

which is unbounded as a function of x, making the upper bound of Theorem 3.6 useless in this context. This is unavoidable in this context, as we will show that there is no upper bound for $\mathbb{E}\, e^{\int_0^T \lambda X_t\,dt}$ of the form $e^{\lambda^2+O(\lambda^3)}$ due to the continuous time character(Poissonian jumps). We will now calculate upper and lower bounds directly to illustrate this.

We have the lower bound

$$\begin{aligned}
\mathbb{E}_\mu e^{F(\mathbb{X})} &\geq \mathbb{P}(N_T = 0)\mathbb{E}_\mu e^{\lambda T X_0} \\
&= e^{-T}\sum_{n=0}^{\infty} e^{\lambda T n}\frac{1}{n!}e^{-1} \\
&= e^{e^{\lambda T}-T-1}.
\end{aligned}$$

When we denote with $\mathbb{t} = (0 = t_0, \ldots, t_n = T)$ the partition induced by the jumps of the poisson process N_t, we have for the upper bound

$$\begin{aligned}
\mathbb{E}_\mu e^{F(\mathbb{X})} &= \mathbb{E}_\mu\left(\mathbb{E}\left[\prod_{k=1}^{n} e^{\lambda(t_k-t_{k-1})X_{k-1}}\,\middle|\,\mathbb{t}\right]\right) \\
&= \mathbb{E}_\mu \prod_{k=1}^{n} e^{e^{\lambda(t_k-t_{k-1})}-1} \\
&= \mathbb{E}_\mu \exp(\underbrace{\sum_{k=1}^{n}(e^{\lambda(t_k-t_{k-1})}-1)}_{\leq e^{\sum_{k=1}^{n}\lambda(t_k-t_{k-1})}-1}) \\
&\leq e^{e^{\lambda T}-1}.
\end{aligned}$$

As we can see, the expectation is of order $e^{e^{\lambda T}}$, which gives the length T of the integration interval a very large influence. We do not get this phenomena in the discrete version of this example:

$$\mathbb{E}\, e^{\sum_{k=1}^{n}\lambda X_k} = \prod_{k=1}^{n}\mathbb{E}\, e^{\lambda X_k} = e^{n(e^\lambda-1)}.$$

Therefore this example shows us why certain results in the discrete case cannot be generalized to continuous time. When correcting for the expectation, λn (or λT), and taking $\lambda = \alpha n^{-\frac{1}{2}}$ (or $\lambda = \alpha T^{-\frac{1}{2}}$),

$$\mathbb{E}\, e^{\sum_{k=1}^{n}\lambda X_k - \lambda n} \underset{n\to\infty}{\longrightarrow} e^{\frac{1}{2}\alpha^2},$$

which is the expected result, but in the continuous time case,

$$\mathbb{E}_{\mu} e^{\int_0^T \lambda X_t\, dt - \lambda T} \underset{T\to\infty}{\longrightarrow} \infty.$$

However, in the present context, we can still use Theorem 3.9 to get an estimate on the moments. By the remark following the theorem,

$$\begin{aligned}
\mathbb{E}_{\mu}\left(\int_0^T \overline{f}_t^2 Ah^2(\cdot, X_t)(X_t)\, dt\right)^{\frac{p}{2}} &\le \lambda^p T^{\frac{p}{2}} \mathbb{E}_{\mu}^{X}\left(\mathbb{E}_{\mu}^{Y} \mid X - Y \mid^2\right)^{\frac{p}{2}} \\
&\le \lambda^p T^{\frac{p}{2}} \mathbb{E}_{\mu}^{X}\left(X^2 + \mathbb{E}_{\mu}^{Y} Y^2\right)^{\frac{p}{2}} \\
&\le 2^{\frac{p}{2}+1} \lambda^p T^{\frac{p}{2}} \mathbb{E}_{\mu}^{X} X^p.
\end{aligned}$$

Next,

$$\begin{aligned}
\mathbb{E}_{\mu} \sup_{0\le t\le T} h(X_{t-}, X_t)^p &= \sum_{n=0}^{\infty} \frac{T^n}{n!} e^{-T} \mathbb{E} \sup_{0\le k\le n} \mid X_k - X_{k-1} \mid^p \\
&\le \sum_{n=0}^{\infty} \frac{T^n}{n!} e^{-T} (n+1) \mathbb{E}_{\mu}^{X} X^p \\
&= (T+1) \mathbb{E}_{\mu}^{X} X^p
\end{aligned}$$

As a last estimate, we have

$$\mathbb{E}_{\mu}^{X}\left(\mathbb{E}_{\mu}^{Y} \mid X - Y \mid\right)^p \le 2^{p+1} \mathbb{E}_{\mu}^{X} X^p.$$

Putting all things together, we have

$$\begin{aligned}
\left(\mathbb{E}_{\mu}\left(\int_0^T \lambda X_t\, dt - \lambda T\right)^p\right)^{\frac{1}{p}} &\le \left(C_p 2^{\frac{1}{2}+\frac{1}{p}} T^{\frac{1}{2}} + C_p (T+1)^{\frac{1}{p}} + 2^{1+\frac{1}{p}}\right) \lambda (\mathbb{E}_{\mu}^{X} X^p)^{\frac{1}{p}} \\
&\le \widetilde{C}_p T^{\frac{1}{2}} \lambda (\mathbb{E}_{\mu} X^p)^{\frac{1}{p}},
\end{aligned}$$

where the last line holds only for $T \ge 1$ and a new constant $\widetilde{C}_p$ depending only on p.

5.3 Interacting particle systems

Let $E = \{0,1\}^{\mathbb{Z}^d}$ be the state space of the interacting particle system with a generator L given by

$$Lf(\eta) = \sum_x \sum_{\Delta \subset \mathbb{Z}^d} c(\eta, x + \Delta)[f(\eta^{x+\Delta}) - f(\eta)],$$

where η^{Δ} denotes the configuration η with all spins in Δ flipped. This kind of particle system is extensively treated in [7]. For $f : E \to \mathbb{R}$, we denote with $\delta_f(x) :=$

$\sup_{\eta\in E} f(\eta^x) - f(\eta)$ the maximal influence of a single flip at site x, and with $\delta_f = (\delta_f(x))_{x\in E}$ the vector of all those influences.

If there is a way to limit how flips in the configuration affect the system as time progresses, then we can obtain a concentration estimate. Again, denote with $F(\eta_\cdot) = \int_0^T f(\eta_t)\,dt$ the additive functional of the function f and the particle system $\eta_\cdot$.

Theorem 5.1. *Assume there exists a family of operators A_t so that $\delta_{S_t f} \leq A_t \delta_f$, and write*

$$G := \int_0^\infty A_t\,dt,$$

which is assumed to exist. Denote with

$$c_k := \sup_{\eta\in E, x\in\mathbb{Z}^d} \sum_{\Delta\subset\mathbb{Z}^d} c(\eta, x+\Delta)\,|\,\Delta\,|^k$$

the weighted maximal rate of spin flips. If $\|\,G\,\|_{p\to 2} < \infty$ for some $p \geq 1$, then for any f with $\delta_f \in \ell^p$ and any initial condition $\eta \in E$,

$$\mathbb{E}_\eta e^{F(\eta_\cdot)-\mathbb{E}_\eta F(\eta_\cdot)} \leq \exp\left[T\sum_{k=2}^\infty \frac{c_k\,\|\,G\,\|_{p\to 2}^k\,\|\,\delta_f\,\|_p^k}{k!}\right].$$

If additionally $\|\,G\,\|_1 < \infty$ and $|||\,f\,||| := \|\,\delta_f\,\|_1 < \infty$, then for any two distributions ν_1, ν_2,

$$\mathbb{E}_{\nu_1} e^{F(\eta_\cdot)-\mathbb{E}_{\nu_2} F(\eta_\cdot)} \leq \exp\left[\|\,G\,\|_1\ |||\ f\ |||\ + T\sum_{k=2}^\infty \frac{c_k\,\|\,G\,\|_{p\to 2}^k\,\|\,\delta_f\,\|_p^k}{k!}\right].$$

Applications of this Theorem are for example spin flip dynamics in the so-called $M < \epsilon$ regime, where there exists an operator Γ with $\|\,\Gamma\,\|_1 = M$, so that

$$\delta_{S_t f} \leq e^{-t(\epsilon-\Gamma)}\delta_f$$

holds. Since $G = \int_0^\infty e^{-t(\epsilon-\Gamma)}\,dt = (\epsilon-\Gamma)^{-1}$, $\|\,G\,\|_1 \leq (\epsilon - M)^{-1}$. Hence $\|\,G\,\|_{1\to 2} \leq (\epsilon - M)^{-1}$ for a first application of the Theorem. If the process is reversible as well, $\|\,G\,\|_\infty = \|\,G\,\|_1$, and by Riesz-Thorin's Theorem, we have $\|\,G\,\|_2 \leq (\epsilon - M)^{-1}$, hence we get the result for functions f with $\|\,\delta_f\,\|_2 < \infty$.

Another example is the exclusion process. As a singe discrepancy is preserved and moves like a random walk, $A_t(x,y) = p_t(x,y)$, the transition probability of the random walk. In high dimensions, $G(x,y) = \int_0^\infty p_t(x,y)\,dt$ has bounded $\ell^1 \to \ell^2$-norm:

$$\begin{aligned}
\|\,G\,\|_{1\to 2} &= \sup_{\|\,g\,\|_1=1} \sum_x \Big(\sum_y G(x,y)g(y)\Big)^2 \\
&\leq \sup_{\|\,g\,\|_1=1} \sum_x \sum_y |\,g(y)\,|\,G(x,y)^2 \leq \sum_x G(x,0)^2 \infty \\
&= \int_0^\infty \int_0^\infty \sum_x p_t(0,x)p_s(0,x)\,ds\,dt = \int_0^\infty \int_0^\infty p_{s+t}(0,0)\,ds\,dt < \infty
\end{aligned}$$

in dimension 5 and higher. As the exclusion process switches two sites, $c_k \leq 2^k$, and hence

$$\mathbb{E}_\eta e^{F(\eta.)-\mathbb{E}_\eta F(\eta.)} \leq \exp\left[T\sum_{k=2}^{\infty}\frac{2^k \,\|\, G \,\|_{1\to 2}^k \,|||\, f \,|||^k}{k!}\right].$$

However, this is only a quick result exploiting the strong diffusive behaviour in high dimensions. In the last section we will deal with the exclusion process in much more detail to obtain results for lower dimensions as well.

Proof of Theorem 5.1 First, we notice that the coupled function difference Φ_t for a single flip can be bounded by

$$\begin{aligned}
\Phi_t(\eta^x, \eta) &\leq \int_0^\infty |\, S_t f(\eta^x) - S_t f(\eta) \,|\, dt \\
&\leq \int_0^\infty \delta_{S_t f}(x)\, dt \leq \int_0^\infty (A_t \delta_f)(x)\, dt \\
&\leq (G\delta_f)(x)
\end{aligned}$$

uniformly in η. To estimate the coupled function difference Φ_t we telescope over single site flips,

$$\Phi_t^k(\eta^{x+\Delta}, \eta) \leq |\, \Delta \,|^k \,((G\delta_f)(x))^k,$$

and therefore

$$\begin{aligned}
L\Phi_t^k(\cdot, \eta)(\eta) &= \sum_x \sum_{\Delta\subset\mathbb{Z}^d} c(\eta, x+\Delta)\Phi_t^k(\eta^{x+\Delta}, \eta) \\
&\leq \sum_x \sum_{\Delta\subset\mathbb{Z}^d} c(\eta, x+\Delta) \,|\, \Delta \,|^k \,(G\delta_f)^k(x) \\
&\leq c_k \,\|\, G\delta_f \,\|_k^k \leq c_k \,\|\, G\delta_f \,\|_2^k \leq c_k \,\|\, G \,\|_{p\to 2}^k \,\|\, \delta_f \,\|_p^k
\end{aligned}$$

Hence the first part is proven by applying these estimates to Theorem 3.6 for fixed and identical initial conditions. To prove the estimate for arbitrary initial distributions, we simply observe that, again by telescoping over single site flips,

$$\Phi_0(\eta, \xi) \leq \sum_x \sup_\zeta \Phi_0(\zeta^x, \zeta) \leq \sum_x (G\delta_f)(x) \leq \|\, G \,\|_1 \,\|\, f \,\|_1 .$$

□

5.4 Simple symmetric random walk

The aim of this example is to show that we can get concentration estimates even if the process $\mathbb{X}$, in this example a simple symmetric nearest neighbour random walk in $\mathbb{Z}^d$, has no stationary distribution. We will consider three cases, $f \in \ell^1(\mathbb{Z}^d)$, $\ell^2(\mathbb{Z}^d)$

and $\ell^\infty(\mathbb{Z}^d)$, and $F(\mathbb{X}) = \int_0^T f(X_t)\, dt$. To apply Theorem 3.6, our task is to estimate $|\,\Phi_t(x,y)\,|$ where y is a neighbour of x. We will denote with $p_t(x,z)$ the transition probability from x to z in time t. We start with the estimate on the coupled function difference

$$
\begin{aligned}
|\,\Phi_t(x,y)\,| &= \left| \int_0^{T-t} \mathbb{E}_x f(X_s) - \mathbb{E}_y f(X_s)\, ds \right| \\
&= \left| \int_0^{T-t} \sum_{z\in\mathbb{Z}^d} f(z)(p_s(x,z) - p_s(y,z))\, ds \right| \\
&\le \sum_z |\,f(z)\,| \left| \int_0^{T-t} p_s(x,z) - p_s(y,z)\, ds \right| \\
&\le \sum_z |\,f(z)\,| \left| \int_0^{T} p_s(x,z) - p_s(y,z)\, ds \right| .
\end{aligned}
$$

Now, depending on the three cases of f, we proceed differently. First, let $f \in \ell^1$. Then,

$$
\begin{aligned}
|\,\Phi_t(x,y)\,| &\le \sum_z |\,f(z)\,| \left| \int_0^{T} p_s(x,z) - p_s(y,z)\, ds \right| \\
&\le \|\,f\,\|_1 \sup_z \left| \int_0^{T} p_s(x,z) - p_s(y,z)\, ds \right| \\
&= \|\,f\,\|_1 \int_0^{T} p_s(0,0) - p_s(y-x,0)\, ds \le C_1 \,\|\,f\,\|_1 \,.
\end{aligned}
$$

Since $|\,x-y\,| = 1$, the constant $C_1 = \int_0^\infty p_s(0,0) - p_s(y-x,0)\, ds$ depends on the dimension but nothing else.

Second, let $f \in \ell^\infty$. Then,

$$
\begin{aligned}
|\,\Phi_t(x,y)\,| &\le \sum_z |\,f(z)\,| \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&\le \|\,f\,\|_\infty \sum_z \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&= \|\,f\,\|_\infty \int_0^T \sum_z |\,p_s(x,z) - p_s(y,z)\,|\;ds \\
&= \|\,f\,\|_\infty \int_0^T \frac{1}{2}\,\|\,p_s(x,\cdot) - p_s(y,\cdot)\,\|_{TVar}\;ds \\
&\le \|\,f\,\|_\infty \int_0^T \widehat{\mathbb{P}}_{x,y}(\tau > s)\,ds
\end{aligned}
$$

In the last line, we used the coupling inequality. The coupling $\widehat{\mathbb{P}}_{x,y}$ is the Ornstein coupling, i.e., the different coordinates move independently until they meet. Since x and y are equal in all but one coordinate, the probability of not having succeeded at time t is of order $t^{-\frac{1}{2}}$. Hence we end up with

$$|\,\Phi_t(x,y)\,| \le C_\infty\,\|\,f\,\|_\infty\,\sqrt{T}.$$

Third, let $f \in \ell^2$. This is the most interesting case.

Lemma 5.2. *Let $x, y \in \mathbb{Z}^d$ be neighbours. Then*

$$\sum_{z\in\mathbb{Z}^d} \left(\int_0^T p_t(x,z) - p_t(y,z)\,dt \right)^2 \le \alpha(T)$$

with

$$\alpha(T) \in \begin{cases} O(\sqrt{T}), & d = 1; \\ O(\log T), & d = 2; \\ O(1), & d \ge 3. \end{cases}$$

Proof By expanding the product and using the fact that $\sum\limits_z p_t(a,z)p_s(b,z) = p_{t+s}(a,b) =$

$p_{t+s}(a-b,0)$, we get

$$\sum_{z\in\mathbb{Z}^d}\left(\int_0^T p_t(x,z)-p_t(y,z)\,dt\right)^2 = 2\int_0^T\int_0^T p_{t+s}(0,0)-p_{t+s}(x-y,0)\,dt\,ds$$
$$= 2\int_0^T\int_0^T (-\Delta)p_{t+s}(\cdot,0)(0)\,dt\,ds = 2\int_0^T p_s(0,0)-p_{T+s}(0,0)\,ds$$
$$\le 2\int_0^T p_s(0,0)\,ds =: \alpha(T).$$

□

Using first the Cauchy-Schwarz inequality and then Lemma 5.2,

$$|\,\Phi_t(x,y)\,|^k \le \|\,f\,\|_2^k\left(\sum_z\left(\int_0^T p_t(x,z)-p_t(y,z)\,dt\right)^2\right)^{\frac{k}{2}} \le \|\,f\,\|_2^k\,\alpha(T)^{\frac{k}{2}}.$$

To conclude this example, we finally use the uniform estimates on Φ_t to apply Theorem 3.6 and obtain

$$\mathbb{E}_x\exp\left[\int_0^T f(X_t)\,dt-\mathbb{E}_x\int_0^T f(X_t)\,dt\right]\le\exp\left[T\sum_{k=2}^{\infty}\frac{C_1^k\,\|\,f\,\|_1^k}{k!}\right],\qquad f\in\ell^1;$$
$$\mathbb{E}_x\exp\left[\int_0^T f(X_t)\,dt-\mathbb{E}_x\int_0^T f(X_t)\,dt\right]\le\exp\left[T\sum_{k=2}^{\infty}\frac{\|\,f\,\|_2^k}{k!}\alpha(T)^{\frac{k}{2}}\right],\qquad f\in\ell^2;$$

and

$$\mathbb{E}_x\exp\left[\int_0^T f(X_t)\,dt-\mathbb{E}_x\int_0^T f(X_t)\,dt\right]\le\exp\left[T\sum_{k=2}^{\infty}\frac{C_\infty^k\,\|\,f\,\|_\infty^k}{k!}T^{\frac{k}{2}}\right],\qquad f\in\ell^\infty.$$

Since the generator is $Af(x)=\frac{1}{2d}\sum_{y\sim x}(f(y)-f(x))$, we use the estimates $2d$ times and divide by $2d$, so no additional constants appear in the results.

6 Application: Simple symmetric exclusion process

This example is somewhat more involved(because of the conservation law), and shows the full power of our approach in the context where classical functional inequalities such as log-Sobolev inequality do not hold.

The simple symmetric exclusion process is defined via its generator

$$Af(\eta) = \sum_{x \sim y} \frac{1}{2d}(f(\eta^{xy}) - f(\eta)).$$

It is known that the large deviation behaviour of the occupation time of the origin $\int_0^T \eta_t(0)\,dt$ is dependent on the dimension[5]. Its variance is of order $T^{\frac{3}{2}}$ in dimension $d = 1$, $T\log(T)$ in dimension $d = 2$ and T in dimensions $d \geq 3$([1]). Here we will show the same kind of time dependence for the exponential moments, in dimension $d = 1$ for functionals of any quasi-local function f, and in dimension $d \geq 2$ for the occupation time of a finite set A.

Theorem 6.1. *Let $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ be such that $|||\, f \,||| < \infty$, and fix an initial configuration $\eta_0 \in \{0,1\}^{\mathbb{Z}}$. Then*

$$\mathbb{E}_{\eta_0} \exp\left(\int_0^T f(\eta_t)\,dt - \mathbb{E}_{\eta_0} \int_0^T f(\eta_t)\,dt \right) \leq \exp\left[T^{\frac{3}{2}} c_1 \sum_{k=2}^{\infty} \frac{(c_2 \,|||\, f \,|||)^k}{k!} \right],$$

and the constants $c_1, c_2 > 0$ are independent of f, η_0 and T.

While it is natural to assume the same kind of result in all dimensions(with a properly adjusted dependence on T), we can only prove it in high dimensions($d \geq 5$, see application of Theorem 5.1) or for a subset of the local functions, the occupation indicator $H_A(\eta) := \prod_{a \in A} \eta(a)$ of a finite set $A \subset \mathbb{Z}^d$, with a slightly worse dependence on the function(i.e. $|\, A \,|$).

Theorem 6.2. *Let $A \subset \mathbb{Z}^d$ be a finite, and fix an initial configuration $\eta_0 \in \{0,1\}^{\mathbb{Z}^d}$. Then, for all $\lambda > 0$,*

$$\mathbb{E}_{\eta_0} \exp\left(\int_0^T \lambda H_A(\eta_t)\,dt - \mathbb{E}_{\eta_0} \int_0^T \lambda H_A(\eta_t)\,dt \right) \leq e^{T\alpha(T) \sum_{k=2}^{\infty} \frac{(c\lambda |\, A \,|^3)^k}{k!}},$$

where $\alpha(T) \in O(T^{\frac{1}{2}}), O(\log T)$ or $O(1)$ in dimensions $d = 1$, $d = 2$ or $d \geq 3$. The constant $c > 0$ is independent of A, η_0 and T, but may depend on the dimension d.

The proofs of Theorems 6.1 and 6.2 are subject of the two subsections below. For Theorem 6.2, we will only look at $d \geq 2$, the case $d = 1$ is contained in Theorem 6.1.

6.1 Concentration of quasi-local functions in $d = 1$

Let f be a quasi-local function. To derive an exponential estimate, we will create a coupling between the exclusion process started in η and started in η^{xy}:

Proposition 6.3. *There exists a coupling $\widehat{\mathbb{P}}_{\eta,\eta^{xy}}$ of $\mathbb{P}_\eta$ and $\mathbb{P}_{\eta^{xy}}$ for which*

$$\widehat{\mathbb{E}}_{\eta,\eta^{xy}} \mathbb{1}_{\eta_t^1(z) \neq \eta_t^2(z)} \leq C\,|\, p_t(x,z) - p_t(y,z) \,|$$

holds for some constant $C > 0$.

Proof To couple two exclusion processes with almost identical initial conditions, we use a variation of the graphical representation to describe their development, which is the following: at each edge between two consecutive integer numbers, we put an independent Poissonian clock of rate 1, and whenever this clock rings we exchange the occupation status of the sites which are connected by the edge associated to the clock, which is represented by a double sided arrow. Now, to couple $\mathbb{P}_\eta$ with $\mathbb{P}_{\eta^{xy}}$, we instead take Poissonian clocks of rate 2, and additionally a sequence of independent fair coin flips associated to the arrows. For both η^1 and η^2, which use the same arrow configuration, if the coin flip corresponding to an arrow is tails, that arrow is ignored, with one exception explained a bit lated. First, we notice that this leads to effective rates of 1. Second, since we start with just two discrepancies(one at x and one at y), those remain the only discrepancies, and they perform independent random walk movements until they encounter the same arrow, which leads us to the only exception of the mechanics described above: When there is an arrow connecting the two discrepancies, the exchange of process η^1 is suppressed if the coin flip is tails, but then η^2 performs the exchange, and if the coin flip is heads, η^1 performs the exchange and η^2 does not. After this event, η^1 and η^2 are identical.

If we denote the position of the discrepancies by X_t and Y_t, those perform independent random walks of rate 1 until they meet, then they stay together. Hence

$$\widehat{\mathbb{E}}_{\eta,\eta^{xy}} \mathbb{1}_{\eta_t^1(z)\neq\eta_t^2(z)} = \widehat{\mathbb{E}}_{x,y} \mathbb{1}_{X_t\neq Y_t, z\in\{X_t,Y_t\}} \leq C\,|\,p_t(x,z) - p_t(y,z)\,|\,,$$

where we used the fact that in dimension 1, the independent coupling of two random walks is optimal and hence

$$\widehat{\mathbb{P}}_{x,y}(X_t = Y_t = z) = p_t(x,z) \wedge p_t(y,z).$$

$\square$

To apply Theorem 3.6, we have to estimate

$$\begin{aligned}
L\Phi^k(\cdot,\eta)(\eta) &\leq \sum_{x\in\mathbb{Z}}\sum_{j=\pm1}\left|\int_0^T S_t f(\eta^{x,x+j}) - S_t f(\eta)\,dt\right|^k \\
&\leq \sum_{x\in\mathbb{Z}}\sum_{j=\pm1}\left(\int_0^T \sum_z \delta_f(z)\mathbb{E}_{\eta,\eta^{x,x+j}} \mathbb{1}_{\eta_t^1(z)\neq\eta_t^2(z)}\,dt\right)^k \\
&\leq C^k \sum_{x\in\mathbb{Z}}\sum_{j=\pm1}\left(\int_0^T \sum_z \delta_f(z)\,|\,p_t(x,z) - p_t(y,z)\,|\,dt\right)^k,
\end{aligned}$$

where we used Proposition 6.3 to obtain the last line. To continue, we calculate

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}}\sum_{j=\pm 1}\int_0^T \sum_z \delta_f(z)\,|\,p_t(x,z)-p_t(y,z)\,|\;dt \\
&= \sum_{j=\pm 1}\sum_z \delta_f(z)\int_0^T \|\,p_t(0,\cdot)-p_t(j,\cdot)\,\|_{TVar}\;dt \\
&\le \tilde{C}\,|||\;f\;|||\,\sqrt{T}.
\end{aligned}$$

Next,

$$\begin{aligned}
&\sup_{x\in\mathbb{Z}}\sup_{j=\pm 1}\int_0^T \sum_z \delta_f(z)\,|\,p_t(x,z)-p_t(y,z)\,|\;dt \\
&\le |||\,f\;|||\sup_{j=\pm 1}\sup_z \int_0^T |\,p_t(0,z)-p_t(j,z)\,|\;dt \\
&= |||\,f\;|||\sup_{z\ge 0}\int_0^T p_t(0,z)-p_t(-1,z)\,dt.
\end{aligned}$$

In order to control the supremum over z on the right hand side of the last line, let τ_0 denote the first time a simple symmetric random walk $(X_t)_{t\ge 0}$ hits 0. Then

$$\begin{aligned}
\int_0^T p_t(0,z)-p_t(-1,z)\,dt &= \int_0^T p_t(0,z)\,dt - \mathbb{E}\left[\int_{\tau_0\wedge T}^T p_{t-\tau_0}(0,z)\,dt\,\middle|\,X_0=-1\right] \\
&= \mathbb{E}\left[\int_{T-\tau\wedge T}^T p_t(0,z)\,dt\,\middle|\,X_0=-1\right] \\
&\le \mathbb{E}\left[\int_{T-\tau\wedge T}^T p_t(0,0)\,dt\,\middle|\,X_0=-1\right] \\
&= \int_0^T p_t(0,z)-p_t(-1,z)\,dt =: \overline{C} < \infty.
\end{aligned}$$

Hence

$$L\Phi^k(\cdot,\eta)(\eta)\le |||\,f\;|||^k\,\sqrt{T}C_1C_2^k$$

for suitable constants C_1 and C_2, and Theorem 3.6 implies

$$\mathbb{E}_\eta \exp\left(\int_0^T f(\eta_t)\,dt - \mathbb{E}_\eta\int_0^T f(\eta_t)\,dt\right)\le e^{T^{\frac{3}{2}}C_1\sum_{k=2}^{\infty}\frac{(C_2|||\,f\,|||)^k}{k!}}$$

for any initial configuration η.

6.2 Concentration of the occupation time of a finite set in $d \geq 2$

Now, we want to show that the occupation time $\int_0^T H_A(\eta_t)\,dt$, $H_A(\eta) := \prod_{a\in A} \eta(a)$, of a finite set $A \subset \mathbb{Z}^d$ has the same time asymptotic behaviour as the occupation time of a single site. As a stating point to estimate $L\,|\,\Phi_t\,|^k(\cdot,\eta)$, we use the following result by Ferrari, Galves and Landim:

Theorem 6.4. *[2], Theorem 2.2*

$$\begin{aligned}
&\mathbb{E}_\eta \prod_{a\in A} \eta_t(a) - \prod_{a\in A} \rho_t^\eta(a) \\
&= -\frac{1}{2}\int_0^t ds \sum_{\substack{Z\subset\mathbb{Z}^d \\ |\,Z\,|=|\,A\,|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1,z_2\in Z \\ z_1\neq z_2}} p(z_1,z_2)(\rho_{t-s}^\eta(z_1) - \rho_{t-s}^\eta(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^\eta(z_3)
\end{aligned}$$

Here $\mathbb{P}_A(X_s = Z)$ is the probability of exclusion walkers started in A occupying the set Z at time s, and $\rho_t^\eta(z) = \mathbb{E}_\eta\eta_t(z)$ is the occupation probability of z at time t given the initial configuration η.

By using this comparison of exclusion dynamics with independent random walkers, we get

$$\begin{aligned}
&\mathbb{E}_{\eta^{xy}} \prod_{a\in A} \eta_t(a) - \mathbb{E}_\eta \prod_{a\in A} \eta_t(a) \\
&= \mathbb{E}_{\eta^{xy}} \prod_{a\in A} \eta_t(a) - \prod_{a\in A} \rho_t^{\eta^{xy}}(a) + \prod_{a\in A} \rho_t^{\eta^{xy}}(a) - \prod_{a\in A} \rho_t^\eta(a) + \prod_{a\in A} \rho_t^\eta(a) - \mathbb{E}_\eta \prod_{a\in A} \eta_t(a) \\
&= \left(\prod_{a\in A} \rho_t^{\eta^{xy}}(a) - \prod_{a\in A} \rho_t^\eta(a)\right) - \frac{1}{2}\int_0^t ds \sum_{\substack{Z\subset\mathbb{Z}^d \\ |\,Z\,|=|\,A\,|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1,z_2\in Z \\ z_1\neq z_2}} p(z_1,z_2)\cdot \\
&\quad \cdot \left[(\rho_{t-s}^{\eta^{xy}}(z_1) - \rho_{t-s}^{\eta^{xy}}(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^{\eta^{xy}}(z_3) - (\rho_{t-s}^\eta(z_1) - \rho_{t-s}^\eta(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^\eta(z_3)\right]
\end{aligned}$$

Taking absolute values, we start to estimate the first difference:

$$\left|\prod_{a\in A} \rho_t^{\eta^{xy}}(a) - \prod_{a\in A} \rho_t^\eta(a)\right| \leq \sum_{a\in A} \left|\rho_t^{\eta^{xy}}(a) - \rho_t^\eta(a)\right| = \sum_{a\in A} |\,p_t(x,a) - p_t(y,a)\,|\,.$$

The next part is the big difference inside the integral. It is estimated by

$$\begin{aligned}
&\left|(\rho_{t-s}^{\eta^{xy}}(z_1) - \rho_{t-s}^{\eta^{xy}}(z_2))^2 - (\rho_{t-s}^\eta(z_1) - \rho_{t-s}^\eta(z_2))^2\right| \\
&+ \sum_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \left|\rho_{t-s}^{\eta^{xy}}(z_3) - \rho_{t-s}^\eta(z_3)\right| (\rho_{t-s}^\eta(z_1) - \rho_{t-s}^\eta(z_2))^2
\end{aligned}$$

Now we come back to the original task of estimating $L\,|\,\Phi_t\,|^k\,(\cdot,\eta)$. From now on, multiplicative constants are ignored on a regular basis, which results in an omitted factor of the form $c_1 c_2^k$. However warning is given by using $\lesssim$ instead of $\leq$. By using the above estimates, we obtain the upper bound

$$\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T \sum_{a\in A} |\,p_t(x,a)-p_t(y,a)\,|\,dt\right)^k \tag{6}$$

$$+\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T dt\int_0^t ds \sum_{\substack{Z\subset\mathbb{Z}^d\\ |\,Z\,|=|\,A\,|}} \mathbb{P}_A(X_s=Z)\sum_{\substack{z_1,z_2\in Z\\ z_1\neq z_2}} p(z_1,z_2)\cdot\right.$$
$$\left.\cdot\left|\,(\rho_{t-s}^{\eta^{xy}}(z_1)-\rho_{t-s}^{\eta^{xy}}(z_2))^2-(\rho_{t-s}^{\eta}(z_1)-\rho_{t-s}^{\eta}(z_2))^2\,\right|\right)^k \tag{7}$$

$$+\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T dt\int_0^t ds \sum_{\substack{Z\subset\mathbb{Z}^d\\ |\,Z\,|=|\,A\,|}} \mathbb{P}_A(X_s=Z)\sum_{\substack{z_1,z_2\in Z\\ z_1\neq z_2}} p(z_1,z_2)\cdot\right.$$
$$\left.\cdot\sum_{\substack{z_3\in Z\\ z_3\neq z_1,z_2}}\left|\,\rho_{t-s}^{\eta^{xy}}(z_3)-\rho_{t-s}^{\eta}(z_3)\,\right|(\rho_{t-s}^{\eta}(z_1)-\rho_{t-s}^{\eta}(z_2))^2\right)^k, \tag{8}$$

which we will treat individually.

For term (6), we estimate sum over A by the maximum times $|\,A\,|$. Hence

$$(6)\leq |\,A\,|^k\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T |\,p_t(x,a_0)-p_t(y,a_0)\,|\,dt\right)^k.$$

We note that

$$\sup_{x\in\mathbb{Z},y\sim x}\int_0^T |\,p_t(x,a_0)-p_t(y,a_0)\,|\,dt$$
$$\leq \sup_{|\,j\,|=1}\int_0^\infty p_t(0,0)-p_t(j,0)\,dt<\infty$$

and

$$\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T |\, p_t(x,a_0)-p_t(y,a_0)\,|\; dt\right)^2$$
$$=\frac{1}{2d}\sum_{|\,j\,|=1}\sum_{x\in\mathbb{Z}^d}\left(\int_0^T p_t(x,a_0)-p_t(x,a_0+j)\,dt\right)^2$$
$$\le \alpha(T)$$

by Lemma 5.2. Hence

$$(6)\lesssim |\,A\,|^k\,\alpha(T).$$

Next, we must treat (7). In the case $k=1$,

$$(7)\lesssim \int_0^T dt\int_0^t ds \sum_{x\in\mathbb{Z}^d}\sum_{y\sim x}\sum_{z_1\in\mathbb{Z}^d}\sum_{z_2\sim z_1}\left(\sum_{Z:z_1,z_2\in Z}\mathbb{P}_A(X_s=Z)\right) \tag{9a}$$
$$\cdot\left|\,\rho^{\eta^{xy}}_{t-s}(z_1)-\rho^{\eta^{xy}}_{t-s}(z_2)-\rho^{\eta}_{t-s}(z_1)+\rho^{\eta}_{t-s}(z_2)\right| \tag{9b}$$
$$\cdot\left|\,\rho^{\eta^{xy}}_{t-s}(z_1)-\rho^{\eta^{xy}}_{t-s}(z_2)+\rho^{\eta}_{t-s}(z_1)-\rho^{\eta}_{t-s}(z_2)\right|. \tag{9c}$$

Regarding the exclusion walkers X_s in (9a), we can simplify by using Liggett's correlation inequality ([7], chapter 8):

$$\sum_{Z:z_1,z_2\in Z}\mathbb{P}_A(X_s=Z)=\mathbb{P}_A(z_1,z_2\in X_s)\le\mathbb{P}_A(z_1\in X_s)\mathbb{P}_A(z_2\in X_s)$$
$$=\left(\sum_{a\in A}p_s(z_1,a)\right)\left(\sum_{a\in A}p_s(z_2,a)\right).$$

Lemma 6.5. *For* $|\,i\,|,|\,j\,|=1$*,*

a) For any η*,*

$$\left|\,\rho_t^{\eta^{x,x+j}}(z)-\rho_t^{\eta^{x,x+j}}(z+i)-\rho_t^{\eta}(z)+\rho_t^{\eta}(z+i)\right|$$
$$\le |\,p_t(x,z)-p_t(x+j,z)-p_t(x,z+i)+p_t(x+j,z+i)\,|,$$

b) $\sum_{x\in\mathbb{Z}^d}|\,p_t(x,z)-p_t(x+j,z)-p_t(x,z+i)+p_t(x+j,z+i)\,|\lesssim(1+t)^{-1}.$

Part b) holds as well when we sum over z instead of x.

Proof First we notice that

$$\rho_t^{\eta^{xy}}(z) - \rho_t^{\eta}(z) = \begin{cases} p_t(y,z) - p_t(x,z), & \eta(x)=1, \eta(y)=0; \\ p_t(x,z) - p_t(y,z), & \eta(x)=0, \eta(y)=1; \\ 0, & otherwise, \end{cases}$$

which immediately proves a). To show b),

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^d} |\, p_t(x,z) - p_t(x+j,z) - p_t(x,z+i) + p_t(x+j,z+i)\,| \\
&\quad = \sum_x \Big| \sum_u p_{t/2}(x,u)p_{t/2}(u,z) - p_{t/2}(x+j,u)p_{t/2}(u,z) \\
&\qquad\qquad - p_{t/2}(x,u)p_{t/2}(u,z+i) + p_{t/2}(x+j,u)p_{t/2}(u,z+i) \Big| \\
&\quad \le \sum_x \sum_u |\,(p_{t/2}(x,u) - p_{t/2}(x+j,u))(p_{t/2}(u,z) - p_{t/2}(u,z+i))\,| \\
&\quad = \sum_u |\, p_{t/2}(u,z) - p_{t/2}(u,z+i)\,| \sum_x |\, p_{t/2}(x,u) - p_{t/2}(x+j,u)\,| \\
&\quad = 4\, \|\, p_{t/2}(0,\cdot) - p_{t/2}(i,\cdot)\,\|_{TVar}\, \|\, p_{t/2}(0,\cdot) - p_{t/2}(j,\cdot)\,\|_{TVar} \\
&\quad \lesssim (1+t/2)^{-\frac{1}{2}}(1+t/2)^{-\frac{1}{2}} \le 2(1+t)^{-1},
\end{aligned}$$

where the last line relies on optimal coupling of two random walks, see for example [8]. □

As a third observation,

$$\begin{aligned}
|\, \rho_t^{\eta}(z_1) - \rho_t^{\eta}(z_2)\,| &= \left| \sum_x (p_t(z_1,x) - p_t(z_2,x))\eta(x) \right| \\
&\le \|\, p_t(z_1,\cdot) - p_t(z_2,\cdot)\,\|_{TVar}\,, \qquad (10)
\end{aligned}$$

which leads to the estimate

$$(9c) \le 2\, \|\, p_{t-s}(z_1,\cdot) - p_{t-s}(z_2,\cdot)\,\|_{TVar} \lesssim (1+t-s)^{-\frac{1}{2}}.$$

Applying the estimates for (9a) to (9c), we have (for $k=1$)

$$\begin{aligned}
(7) &\lesssim \int_0^T dt \int_0^t ds \sum_{z_1\in\mathbb{Z}^d} \sum_{z_2\sim z_1} \left(\sum_{a\in A} p_s(z_1,a)\right) \left(\sum_{a\in A} p_s(z_2,a)\right) (1+t-s)^{-\frac{3}{2}} \\
&\le 2d\,|\,A\,|^2 \int_0^T dt \int_0^t ds\ p_s(0,0)(1+t-s)^{-\frac{3}{2}} \\
&\lesssim |\,A\,|^2 \int_0^T dt \int_0^t ds\ (1+s)^{-\frac{d}{2}}(1+t-s)^{-\frac{3}{2}} \\
&\lesssim |\,A\,|^2\, \alpha(T),
\end{aligned}$$

where the last line is due to the following lemma:

Lemma 6.6.

$$\int_0^T dt \int_0^t ds\ (1+s)^{-\frac{n}{2}}(1+t-s)^{-\frac{3}{2}} \lesssim \begin{cases} \sqrt{T}, & n=1; \\ \log(1+T), & n=2; \\ 1, & n \geq 3. \end{cases}$$

Proof Write

$$f(m,n) := \int_0^t (1+s)^{-\frac{m}{2}}(1+t-s)^{-\frac{n}{2}}\, dt.$$

Then f satisfies $f(m,n) \leq (1+t)^{-\frac{1}{2}}(f(m-1,n)+f(m,n-1))$ for $m,n \geq 1$:

$$\begin{aligned} f(m,n) &= (1+t)^{-\frac{1}{2}} \int_0^t \frac{(1+t)^{\frac{1}{2}}}{(1+s)^{\frac{m}{2}}(1+t-s)^{\frac{n}{2}}}\, dt \\ &\leq (1+t)^{-\frac{1}{2}} \int_0^t \frac{(1+s)^{\frac{1}{2}}+(1+t-s)^{\frac{1}{2}}}{(1+s)^{\frac{m}{2}}(1+t-s)^{\frac{n}{2}}}\, dt \\ &= (1+t)^{-\frac{1}{2}}(f(m-1,n)+f(m,n-1)). \end{aligned}$$

Also, $f(n,0) = f(0,n) \lesssim (1+t)^{\frac{1}{2}}, \log(1+t)$ or 1 for $n=1, n=2$ or $n \geq 3$. Using these two rules we obtain the given estimates. $\square$

As we have already dealt with (7) when $k=1$, we use the simple fact

$$\sum_x h(x)^k \leq (\sum_x h(x))(\sup_x h(x))^{k-1}, \quad h \geq 0,$$

to generalize to any k. However, we must show that (7) is bounded by a constant when we replace the sum by the supremum. When we use the same initial estimates as above, we get

$$\begin{aligned} &\sup_{x\in\mathbb{Z}^d} \sup_{y\sim x} \int_0^T dt \int_0^t ds \sum_{z_1\in\mathbb{Z}^d} \sum_{z_2\sim z_1} \left(\sum_{a\in A} p_s(a,z_1)\right)\left(\sum_{a\in A} p_s(a,z_2)\right) \\ &\cdot |\, p_{t-s}(x,z_1) - p_{t-s}(x,z_2) - p_{t-s}(y,z_1) + p_{t-s}(y,z_2)\,|\, \|\, p_{t-s}(z_1,\cdot) - p_{t-s}(z_2,\cdot)\, \|_{TVar}, \end{aligned}$$

and by taking the sum over z_1 over the p_{t-s} differences,

$$\begin{aligned} &\lesssim \int_0^T dt \int_0^t ds\ (|\,A\,|\,p_s(0,0))^2(1+t-s)^{-\frac{3}{2}} \\ &\lesssim |\,A\,|^2 \int_0^T dt \int_0^t ds\ (1+s)^{-d}(1+t-s)^{-\frac{3}{2}} \lesssim |\,A\,|^2 \qquad \text{if } d \geq 2. \end{aligned}$$

Hence, finally, we have obtained the estimate

$$(7) \lesssim |A|^{2k} \alpha(T).$$

Part (8) is treated in a similar way:

$$\sum_{\substack{Z \subset \mathbb{Z}^d \\ |Z|=|A|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1,z_2 \in Z \\ z_1 \neq z_2}} p(z_1,z_2) \sum_{\substack{z_3 \in Z \\ z_3 \neq z_1,z_2}} \left| \rho_{t-s}^{\eta^{xy}}(z_3) - \rho_{t-s}^{\eta}(z_3) \right| (\rho_{t-s}^{\eta}(z_1) - \rho_{t-s}^{\eta}(z_2))^2$$

$$\lesssim \sum_{z_1} \sum_{z_2 \sim z_1} \sum_{z_3} \prod_{i=1}^{3} \left(\sum_{a \in A} p_s(a,z_i) \right) |p_{t-s}(y,z_3) - p_{t-s}(x,z_3)| \, \| p_{t-s}(z_1,\cdot) - p_{t-s}(z_2,\cdot) \|_{TVar}^2$$

By using the fact that

$$\sum_x |p_{t-s}(x+j,z) - p_{t-s}(x,z)| = 2 \, \| p_{t-s}(j,\cdot) - p_{t-s}(0,\cdot) \|_{TVar}$$

we can sum over x to obtain another power of the total variation distance. Also,

$$\sum_{z_1} \sum_{z_2 \sim z_1} \sum_{z_3} \prod_{i=1}^{3} \left(\sum_{a \in A} p_s(a,z_i) \right) \leq 2d \, |A|^3 p_s(0,0),$$

hence we obtain the compound estimate

$$|A|^3 p_s(0,0)(1+t-s)^{-3/2},$$

which after integrating over s and t is again of order $\alpha(T)$. When we take the supremum over x, we can instead take the sum over z_3 on the middle term. Hence we keep another $p_s(0,0)$ and we get and get

$$|A|^3 p_s(0,0)^2(1+t-s)^{-3/2},$$

which after integration is of order 1 if $d \geq 2$. Hence,

$$(8) \lesssim |A|^{3k} \alpha(T).$$

Returning to the original question,

$$L \, |\Phi_t|^k (\cdot,\eta)(\eta) \lesssim |A|^k \alpha(T) + |A|^{2k} \alpha(T) + |A|^{3k} \alpha(T) \lesssim |A|^{3k} \alpha(T),$$

and after replacing $\lesssim$ with $\leq$,

$$L \, |\Phi_t|^k (\cdot,\eta)(\eta) \leq c_1 c_2^k \, |A|^{3k} \alpha(T).$$

Now that we have this estimate, Theorem 3.6 gives us the estimate

$$\mathbb{E}_\eta \exp \left(\int_0^T \lambda H_A(\eta_t) \, dt - \mathbb{E}_\eta \int_0^T \lambda H_A(\eta_t) \, dt \right) \leq \exp \left(T\alpha(T) c_1 \sum_{k=2}^{\infty} \frac{(c_2 \lambda \, |A|^3)^k}{k!} \right),$$

where the constants c_1 and c_2 do not depend on T or A.

References

[1] R. Arratia. Symmetric exclusion processes: a comparison inequality and a large deviation result. *Annals of Probability*, 13(1):53–61, 1985.

[2] P.A. Ferrari, A. Galves, and C. Landim. Rate of converge to equilibrium of symmetric simple exclusion processes. *Markov processes and related fields*, 6:73–88, 2000.

[3] F. Redig J.R. Chazottes. Concentration inequalities for markov processes via coupling. *Electronic Journal of Probability*, 14:1162–1180, 2009.

[4] N. Yao L. Wu. Large deviation principles for markov processes via ϕ-sobolev inequalities. *Electronic Communications in Probability*, (13):10–23, 2008.

[5] C. Landim. Occupation time large deviations for the symmetric simple exclusion process. *The Annals of Probability*, 20(1):206–231, 1992.

[6] M. Ledoux. *The Concentration of Measure Phenomenon*. American Mathematical Society, 2001.

[7] T.M. Liggett. *Interacting Particle Systems*. Springer, 2005.

[8] T. Lindvall. *Lectures on the coupling method*. Wiley, 1992.

[9] A. Guillin P. Cattiaux. Deviation bounds for additive functionals of markov processes. *ESAIM. Probability and Statistics*, 12:12–29, 2008.

[10] Yao-Feng Ran and Fan-Ji Tian. On the Rosenthal's inequality for locally square integrable martingales. *Stochastic processes and their applications*, 104:107–116, 2003.

[11] C. Villani. *Topics in optimal transportation*. American Mathematical Society, 2003.

[12] C. Villani. *Optimal transport. Old and new*. Springer-Verlag, Berlin, 2009.