On the construction of Lyapunov functions for nonlinear Markov processes via relative entropy

Paul Dupuis* and Markus Fischer†

9th January 2011, revised 13th July 2011

Abstract

We develop an approach to the construction of Lyapunov functions for the forward equation of a finite state nonlinear Markov process. Nonlinear Markov processes can be obtained as a law of large number limit for a system of weakly interacting processes. The approach exploits this connection and the fact that relative entropy defines a Lyapunov function for the solution of the forward equation for the many particle system. Candidate Lyapunov functions for the nonlinear Markov process are constructed via limits, and verified for certain classes of models.

1 Introduction

Stochastic processes that go under the unusual title "nonlinear Markov processes" arise as the limits of large collections of exchangeable and weakly interacting "particles." Although these processes originally appeared as models in mathematical physics (Kac, 1956), they now appear in many other contexts, including communication networks (Dawson et al., 2005; McDonald and Reynier, 2006; Antunes et al., 2008; Benaïm and Le Boudec, 2008; Graham and Robert, 2009), economics (Dai Pra et al., 2009), and neural networks (Laughton and Coolen, 1995).

---

*Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, RI, 02912, USA. Research supported in part by the National Science Foundation (DMS-0706003 and DMS-1008331), the Air Force Office of Scientific Research (FA9550-09-1-0378), and the Army Research Office (W911NF-09-1-0155).

†Dipartimento di Matematica Pura ed Applicata, Università di Padova, via Trieste 63, 35121 Padova, Italy. Research supported in part by the Air Force Office of Scientific Research (FA9550-09-1-0378).

A nonlinear Markov process involves a pair of quantities that evolve in time. One quantity is a particle that corresponds to a "typical" or "canonical" particle at the prelimit level and the other is a probability measure. If the measure component were fixed then the particle component would evolve as an ordinary Markov process. The role of the measure-valued component is to account for the aggregate effect of all other particles on the evolution of the canonical particle. At the prelimit, each of these particles interacts with the canonical particle through their state values. Under standard scalings this effect can be summarized in terms of the empirical measure on the particle states, and since many particles contribute equally to this empirical measure, the effect of any one is in some sense weak. In the limit as the number of particles tends to infinity, this empirical measure converges to the measure-valued component of the corresponding nonlinear Markov process. Because of exchangeability one can interpret the empirical measure as an approximation to the distribution of the canonical particle. In the limit the measure component is identified with the exact distribution of the canonical particle, and the two components evolve together. Thus the measure component feeds into the dynamics of the particle component, which in turn determines the evolution of the measure since it coincides with the distribution of the particle. Although the two component picture is useful for interpretative purposes, one can eliminate the particle from the dynamics and obtain an evolution equation for just the measure component.

An important but difficult issue in the study of nonlinear Markov processes is stability. Here, what is meant by stability is deterministic stability of the measure component. For example, one can ask if there is a unique, globally attracting fixed point for the dynamics of the measure-valued component. When this is not the case, all the usual questions regarding stability of deterministic systems, such as existence of multiple fixed points, their local stability properties, etc., arise here as well. One is also interested in the connection between these sorts of stability properties of the limit model and related stability and meta-stability (in the sense of small noise stochastic systems) questions for the prelimit model.

There are several features which make stability analysis particularly difficult for these models. One is that the state space of the system, being the set of probability measures on some Polish space, is not a linear space (although it is a closed, convex subset of a Banach space). Obvious first choices for Lyapunov functions or even local Lyapunov functions, such as quadratic functions, are not naturally suited to such state spaces. Related to the structure of the state space is the fact that the dynamics, linearized at any point in the state space, always have a zero eigenvalue, which also complicates the stability analysis. Based on one's physical intuition, it is sometimes possible to guess a Lyapunov function. See for example Frank (2001), and also Subsection 5.4 below.

The purpose of the present paper is to introduce and develop a somewhat more systematic approach to the construction of Lyapunov functions for nonlinear Markov processes. The starting point is the observation that given any ergodic Markov process, the relative entropy function in some sense always defines a globally attracting Lyapunov function. Now of course a nonlinear Markov process is not a Markov process in the usual sense. Indeed, as will be shown with examples there can be multiple fixed points, locally stable and otherwise, that are associated with the measure component. The fact that relative entropy provides a Lyapunov function for ordinary Markov processes is not directly applicable. Instead, we will temporarily lift the problem to the level of the prelimit processes. Under mild conditions the N-particle process will be ergodic, and thus relative entropy can be used to define a Lyapunov function for the joint distribution of these N particles. The scaling properties of relative entropy and convergence properties of the weakly interacting system then suggest that the limit of suitably normalized relative entropy for the N-particle system, assuming it exists, is a natural candidate Lyapunov function for the corresponding nonlinear Markov process.

To simplify the exposition, in this paper we focus on the relatively simple setting in which each of the interacting particles is modeled by a finite state Markov process. This setting is only relatively simple and, as we will see, even with additional strong assumptions there are many possible behaviors for the corresponding nonlinear process.

An outline of the rest of this paper is as follows. Section 2 provides the notation and basic results on families of weakly interacting Markov processes and the associated limit systems. In Section 3 we recall basic properties of relative entropy and describe different approaches to the construction of Lyapunov functions for the limit systems based on relative entropy. Section 4 studies a class of limit systems where the effect of the nonlinear part due to interaction is in some sense slow. In Section 5 a class of weakly interacting Markov processes of Gibbs type is studied. Section 6 contains concluding remarks.

2 Weakly interacting chains and nonlinear Markov processes

Here we describe the structure and basic asymptotic properties of families of weakly interacting Markov processes in continuous time. To avoid all technical complications let us assume that the state of any individual particle takes values in a finite set $\mathcal{X}$. It is without loss and sometimes convenient to use the particular choice $\mathcal{X} = \{1, \ldots, d\}$. For $N \in \mathbb{N}$, the N-particle system is described as an $\mathcal{X}^N$-valued Markov process. We will use boldface to denote objects that take values in spaces that are N-fold products such as $\mathcal{X}^N$, and ordinary typeface for objects taking values in one copy, such as $\mathcal{X}$.

A heuristic description of the N-particle process is a follows. Let $\mathcal{P}(\mathcal{X})$ denote the space of probability vectors on $\mathcal{X}$. We are given a family of rate matrices $\Gamma(p)$ indexed by $p \in \mathcal{P}(\mathcal{X})$, and assume for simplicity that $p \mapsto \Gamma(p)$ is Lipschitz continuous. Given the states $X^{1,N}(t), \ldots, X^{N,N}(t)$ of all N particles at any time t, form the empirical measure

$$\mu^N(t,\omega) \doteq \frac{1}{N}\sum_{i=1}^{N} \delta_{X^{i,N}(t,\omega)}, \quad t \geq 0,\ \omega \in \Omega, \tag{2.1}$$

where δ_x is the Dirac measure at x. Then the evolution of any particular particle, say $X^{i,N}$, over the interval $[t, t+\delta]$ with $\delta > 0$ small should be (approximately) conditionally independent of the evolution of all other particles, with the law of each particle governed by the rate matrix $\Gamma(\mu^N(t))$.

This description is only heuristic because the particles are continuously recoupled through the evolution of the empirical measure, and so to give a precise description we consider the entire $\mathcal{X}^N$-valued process. Let $\mathbf{x} = (x_1, \ldots, x_N)$, $\mathbf{y} = (y_1, \ldots, y_N)$ be two different states of the N-particle system. We assume that instantaneous transitions can occur only if $\mathbf{x}$ and $\mathbf{y}$ differ in exactly one component. [This is consistent with the special case of a Markovian system made up of N independent particles.] The rate of an admissible transition from $\mathbf{x}$ to $\mathbf{y}$ is assumed to depend on $\mathbf{x}$ and $\mathbf{y}$ only through the values of the component that is changing as well as the empirical measure of $\mathbf{x}$, i.e.,

$$\mu^N_{\mathbf{x}} \doteq \frac{1}{N}\sum_{j=1}^{N} \delta_{x_j}. \tag{2.2}$$

Set $\mathcal{P}_N(\mathcal{X}) \doteq \{\mu^N_{\mathbf{x}} : \mathbf{x} \in \mathcal{X}^N\} \subset \mathcal{P}(\mathcal{X})$.

Let $\mathbf{\Gamma}_N(\mathbf{x},\mathbf{y})$ denote the rate of transition from $\mathbf{x}$ to $\mathbf{y}$, $\mathbf{x}\neq\mathbf{y}$. If $\mathbf{x}$, $\mathbf{y}$ differ in more than one component then $\mathbf{\Gamma}_N(\mathbf{x},\mathbf{y})=0$. If $\mathbf{x}$ and $\mathbf{y}$ differ in exactly one component, that is, if there is $l\in\{1,\ldots,N\}$ such that $x_l\neq y_l$ and $x_j=y_j$ for all $j\neq l$, then according to the assumptions made above

$$\mathbf{\Gamma}_N(\mathbf{x},\mathbf{y})=\Phi_N\left(x_l,y_l,\mu^N_{\mathbf{x}}\right) \tag{2.3}$$

for some function $\Phi_N:\mathcal{X}\times\mathcal{X}\times\mathcal{P}_N(\mathcal{X})\mapsto[0,\infty)$. The N-particle system described previously in a heuristic fashion corresponds to $\Phi_N\left(x_l,y_l,\mu^N_{\mathbf{x}}\right)=\Gamma_{x_l,y_l}(\mu^N_{\mathbf{x}})$. (A situation that requires Φ to also depend on N will appear in Section 5.) Set

$$\mathbf{\Gamma}_N(\mathbf{x},\mathbf{x})\doteq-\sum_{\mathbf{y}\neq\mathbf{x}}\mathbf{\Gamma}_N(\mathbf{x},\mathbf{y}),\quad \mathbf{x}\in\mathcal{X}^N.$$

The matrix $\mathbf{\Gamma}_N=(\mathbf{\Gamma}_N(\mathbf{x},\mathbf{y}))_{\mathbf{x},\mathbf{y}\in\mathcal{X}^N}$ then is a rate matrix, and it corresponds to the infinitesimal generator of the Markov family describing an N-particle system.

Let $\mathbf{X}^N=(X^{1,N},\ldots,X^{N,N})$ be an $\mathcal{X}^N$-valued Markov process with rate matrix $\mathbf{\Gamma}_N$ defined on some probability space $(\Omega,\mathcal{F},\mathbf{P})$. Then $\mathbf{X}^N$ provides a precise construction of the desired interacting particle system, with component $X^{i,N}$ representing the state of the i-th particle. We may assume that $\mathbf{X}^N$ has piecewise constant càdlàg trajectories. The assumption on the non-zero off-diagonal entries of $\mathbf{\Gamma}_N$ implies that, with probability one, $X^{i,N}$, $X^{j,N}$ do not jump simultaneously if $i\neq j$. The empirical measure process corresponding to the N-particle state process $\mathbf{X}^N$ can be written as in (2.1). Notice that μ^N has $\mathcal{P}(\mathcal{X})$-valued càdlàg trajectories. Actually, μ^N is a Markov process with values in $\mathcal{P}_N(\mathcal{X})$. Let $\mathcal{L}_N$ denote the infinitesimal generator of μ^N. Then $\mathcal{L}_N$ acts on bounded measurable functions $f:\mathcal{P}_N(\mathcal{X})\mapsto\mathbb{R}$ according to

$$\mathcal{L}_N(f)(p)=N\sum_{x,y\in\mathcal{X}}\Phi_N(x,y,p)\left(f\left(p+\tfrac{1}{N}\delta_y-\tfrac{1}{N}\delta_x\right)-f(p)\right)p_x.$$

For later use we note that since $\mathbf{X}^N$ is a Markov process with infinitesimal generator $\mathbf{\Gamma}_N$, the evolution of its law obeys the linear Kolmogorov forward equation

$$\frac{d}{dt}\boldsymbol{p}^N(t)=\boldsymbol{p}^N(t)\mathbf{\Gamma}_N,\quad t\geq 0, \tag{2.4}$$

with initial condition $\boldsymbol{p}^N(0)=\mathrm{Law}(\mathbf{X}^N(0))$. Note that $\boldsymbol{p}^N(t)\in\mathcal{P}(\mathcal{X}^N)$ for all $t\geq 0$. By convention, probability vectors are interpreted as row vectors.

Since $\mathcal{X}$ and hence $\mathcal{X}^N$ are finite, Eq. (2.4) is a finite-dimensional ordinary differential equation (ODE).

The structure assumption (2.3) on $\mathbf{\Gamma}_N$ has two important consequences. The first is that the random variables $X^{1,N}(t), \ldots, X^{N,N}(t)$ are exchangeable for all $t > 0$ whenever $X^{1,N}(0), \ldots, X^{N,N}(0)$ are exchangeable. In this sense particles are indistinguishable. The second consequence is the desired property that the evolution of the i-th particle depends on the state of the other particles only through the empirical measure of all particles. More precisely, with probability one, for all $i \in \{1, \ldots, N\}$, $0 \leq s \leq t$, and all functions $f : \mathcal{X} \mapsto \mathbb{R}$,

$$\mathbf{E}\left[f(X^{i,N}(t))|\mathbf{X}^N(s)\right] = \mathbf{E}\left[f(X^{i,N}(t))|X^{i,N}(s), \mu^N(s)\right].$$

As noted previously, the interaction between particles is said to be weak because the contribution of any individual particle to the N-particle empirical measure μ^N is of order $1/N$. Also note that the pair $(X^{i,N}, \mu^N)$ is an $\mathcal{X} \times \mathcal{P}_N(\mathcal{X})$-valued Markov process.

The nonlinear Markov processes mentioned in the introduction arise as the limit as $N \to \infty$ of $(X^{i,N}, \mu^N)$. Since the particles are exchangeable the choice of i is not important, and can be taken to be $i = 1$. We have also noted that the inclusion of the particle component is superfluous, in that one can derive an evolution equation for the limit of the μ^N component alone. Although μ^N is random, its limit (under suitable conditions) will be deterministic, and thus the limit is a form of the law of large numbers. The stability properties of the limit model are the main object of study, though ultimately one is also interested in various properties of the prelimit processes as well.

To simplify the discussion we make the particular choice $\Phi_N(x, y, p) = \Gamma_{x,y}(p)$ for $y \neq x$ for the rest of this section. Since $\Gamma(p)$ is a rate matrix, automatically $\Gamma_{x,x}(p) = -\sum_{y \neq x} \Gamma_{x,y}(p)$ for all $x \in \mathcal{X}$. The result stated below continues to hold if $\Phi_N(x, y, p) \to \Gamma_{x,y}(p)$ uniformly in $p \in \mathcal{P}(\mathcal{X})$ for all $x \neq y \in \mathcal{P}(\mathcal{X})$, a situation that will be encountered in Section 5. The limit of μ^N will be characterized via the nonlinear Kolmogorov forward equation

$$\frac{d}{dt}p(t) = p(t)\Gamma(p(t)). \tag{2.5}$$

Since $\mathcal{X}$ is finite, Eq. (2.5) is simply a finite-dimensional ODE. It is convenient here to recall the identification of $\mathcal{X}$ with $\{1, \ldots, d\}$ and $\mathcal{P}(\mathcal{X})$ with $\{p \in \mathbb{R}^d : p_x \geq 0 \text{ and } \sum_{x=1}^d p_x = 1\}$. Thus $\mathcal{P}(\mathcal{X})$ is a compact and convex subset of $\mathbb{R}^d$ equipped with its standard norm, and the solutions to (2.5) take values in $\mathcal{P}(\mathcal{X})$.

Laws of large numbers for the empirical measures of interacting processes can be efficiently established by using a martingale problem formulation, see Oelschläger (1984), for instance. In the present situation, due to the fact that $\mathcal{X}$ is finite, we can rely on a classical convergence theorem for pure jump Markov processes with state space contained in Euclidean space.

Theorem 2.1. *Suppose that $\Gamma_{x,y}(\cdot)$ is Lipschitz continuous for all $x, y \in \mathcal{X}$, and assume that $\mu^N(0)$ converges in probability to $q \in \mathcal{P}(\mathcal{X})$ as N tends to infinity. Then $(\mu^N(\cdot))_{N\in\mathbb{N}}$ converges uniformly on compact time intervals in probability to $p(\cdot)$, where $p(\cdot)$ is the unique solution to Eq.* (2.5) *with $p(0) = q$.*

Proof. The assertion follows from Theorem 2.11 in Kurtz (1970). In the notation of that work, $E = \mathcal{P}(\mathcal{X})$, $E_N = \mathcal{P}_N(\mathcal{X})$, $N \in \mathbb{N}$,

$$F_N(p) = \sum_{x,y\in\mathcal{X}} N \cdot p_x \left(\tfrac{1}{N}e_y - \tfrac{1}{N}e_x\right)\Gamma_{x,y}(p), \qquad p \in E_N,$$

$$F(p) = \sum_{x,y\in\mathcal{X}} p_x(e_y - e_x)\Gamma_{x,y}(p), \qquad p \in E,$$

where e_x is the unit vector with component x equal to 1. Note that $F_N(p) = \mathcal{L}_N(f)(p)$, $p \in \mathcal{P}_N(\mathcal{X})$, when f is the identity function $f(\tilde{p}) \doteq \tilde{p} \in \mathbb{R}^d$. Moreover, the z-th component of the d-dimensional vector $F(p)$ is equal to $\sum_{x\neq z} p_x\Gamma_{x,z}(p) - \sum_{y\neq z} p_z\Gamma_{z,y}(p)$, which in turn is equal to $\sum_x p_x\Gamma_{x,z}(p)$, the z-th component of the row vector $p\Gamma(p)$, since $\Gamma_{x,z}(p) = \Phi(x,z,p)$ for $x \neq z$ and $\Gamma_{z,z}(p) = -\sum_{y\neq z}\Phi(z,y,p)$. The ODE $\frac{d}{dt}p(t) = F(p(t))$ is therefore the same as Eq. (2.5). $\square$

Under the assumptions of Theorem 2.1, Eq. (2.5) has a unique solution for each initial condition in $\mathcal{P}(\mathcal{X})$, and the solution stays in $\mathcal{P}(\mathcal{X})$. Although we have considered just the convergence in distribution $\mu^N \to p$, one could also consider the limit of the pair $(X^{i,N}, \mu^N)$ and thereby show that Eq. (2.5) is the Kolmogorov forward equation of a nonlinear Markov process. However, one can also directly construct such a process. To see this, let $q \in \mathcal{P}(\mathcal{X})$ and let $p(\cdot)$ be the unique solution to Eq. (2.5) with $p(0) = q$. Set $\Gamma_q(t) \doteq \Gamma(p(t))$, $t \geq 0$. Then $(\Gamma_q(t))_{t\geq 0}$ is a time-dependent rate matrix, and we can find a time-inhomogeneous $\mathcal{X}$-valued Markov process X with $\mathrm{Law}(X(0)) = q$ (Stroock, 2005, Section 5.5.2). The evolution of the law of X at time t obeys the linear time-inhomogeneous forward equation

$$\frac{d}{dt}\tilde{p}(t) = \tilde{p}(t)\Gamma_q(t), \quad t \geq 0, \tag{2.6}$$

with initial condition $\tilde{p}(t) = q$. Under the assumptions of Theorem 2.1, uniqueness holds also for Eq. (2.6). Since $\Gamma_q(t) = \Gamma(p(t))$ and $p(0) = q$, $p(\cdot)$ solves Eq. (2.6) with the same initial condition as $\tilde{p}(\cdot) = \mathrm{Law}(X(\cdot))$. It follows that $\mathrm{Law}(X(\cdot)) = p(\cdot)$, and hence Eq. (2.5) is indeed the forward equation for the nonlinear Markov process X.

A fundamental property of interacting systems that will play a role in the discussion below is propagation of chaos; see Gottlieb (1998) for an exposition and characterization. Propagation of chaos means that the first k components of the N-particle system over any finite time will be asymptotically independent and identically distributed (i.i.d.) as N tends to infinity, whenever the initial distributions of all components are asymptotically i.i.d. In the present context, propagation of chaos for the family $(\mathbf{X}^N)_{N\in\mathbb{N}}$ (or $(\mathbf{\Gamma}_N)_{N\in\mathbb{N}}$) means the following. If $q \in \mathcal{P}(\mathcal{X})$ and if for all $k \in \mathbb{N}$ $\mathrm{Law}(X^{1,N}(0), \ldots, X^{k,N}(0))$ converges to the product measure $\otimes^k q$ weakly as N goes to infinity, then for all $k \in \mathbb{N}$ and all $t \geq 0$,

$$\mathrm{Law}\left(X^{1,N}(t), \ldots, X^{k,N}(t)\right) \overset{N\to\infty}{\longrightarrow} \otimes^k p(t) \text{ weakly in } \mathcal{P}(\mathcal{X}), \tag{2.7}$$

where $p(\cdot)$ is the solution to Eq. (2.5) with $p(0) = q$. Note that the number of components whose joint distribution is asymptotically i.i.d. is fixed. Instead of a particular time t a finite time interval may be considered. Clearly, (2.7) can be rewritten in terms of $\boldsymbol{p}^N$. Under the assumptions of Theorem 2.1, propagation of chaos holds for the family of N-particle systems determined by $(\mathbf{\Gamma}_N)$. See, for instance, Theorem 4.1 in Graham (1992).

3 Relative entropy as a Lyapunov function

In this section we outline an approach to the construction of Lyapunov functions for nonlinear Markov processes. As noted in the introduction, various features of the deterministic system (2.5) make standard forms of Lyapunov functions that might be considered unsuitable. Indeed, one of the most challenging problems in the construction of Lyapunov functions for any system is the identification of natural forms that reflect the particular features and structure of the system. Here the ODE (2.5) is naturally related to a flow of probability measures, and for this reason one might consider constructions based on relative entropy. However, the use of relative entropy as a Lyapunov function is limited to processes that are stationary and ergodic. The ODE (2.5) corresponds to a nonstationary process, though as was discussed in the previous section it arises naturally as a certain limit of stationary processes. The general approach is to identify ways to exploit this connection to suggest good forms for the limit, which in general are obtained as limits of relative entropies for the prelimit model.

3.1 Descent property for linear Markov processes

Here we recall the descent property of Markov processes with respect to relative entropy. This fact, which is well known, has a straightforward proof in the setting of finite-state continuous-time Markov processes. The earliest proof the authors have been able to locate is in Spitzer (1971, pp. I-16-17). Let $G = (G_{x,y})_{x,y\in\mathcal{X}}$ be an irreducible rate matrix over the finite state space $\mathcal{X}$, and denote by π its unique stationary distribution. The forward equation for the family of Markov processes with rate matrix G is the linear ODE

$$\frac{d}{dt}\tilde{p}(t) = \tilde{p}(t)G. \tag{3.1}$$

Define $\psi : [0,\infty) \mapsto [0,\infty)$ by $\psi(z) \doteq z\log(z) - z + 1$. Recall that the relative entropy of $p \in \mathcal{P}(\mathcal{X})$ with respect to $q \in \mathcal{P}(\mathcal{X})$ is given by

$$R\left(p\|q\right) \doteq \sum_{x\in\mathcal{X}} p_x \log\left(\frac{p_x}{q_x}\right).$$

Lemma 3.1. *Let $p(\cdot)$, $q(\cdot)$ be solutions to Eq.* (3.1) *with initial conditions $p(0), q(0) \in \mathcal{P}(\mathcal{X})$. Then for all $t \geq 0$,*

$$\frac{d}{dt}R\left(p(t)\|q(t)\right) = -\sum_{x,y\in\mathcal{X}:x\neq y} \psi\left(\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)}\right) p_x(t)\frac{q_y(t)}{q_x(t)}G_{y,x} \leq 0.$$

Moreover, $\frac{d}{dt}R\left(p(t)\|q(t)\right) = 0$ if and only if $p(t) = q(t)$.

Proof. It is well known (and easy to check) that ψ is strictly convex on $[0,\infty)$, with $\psi(0) = 1$ and $\psi(z) = 0$ if and only if $z = 1$. Owing to the irreducibility of G, $p(t)$ and $q(t)$ have no zero components and hence are equivalent probability vectors for all $t > 0$. Assume for the time being that $p(0)$, $q(0)$ are equivalent probability vectors.

By assumption $\tilde{p}_x'(t) = \sum_{y\in\mathcal{X}} \tilde{p}_y(t)G_{y,x}$ for all $x \in \mathcal{X}$ and all $t \geq 0$. Since

G is a rate matrix, $\sum_{x\in\mathcal{X}} G_{y,x} = 0$ for all $y \in \mathcal{X}$. Thus

$$\begin{aligned}
&\frac{d}{dt} R\left(p(t)\|q(t)\right)\\
&= \frac{d}{dt}\sum_{x\in\mathcal{X}} p_x(t)\log\left(\frac{p_x(t)}{q_x(t)}\right)\\
&= \sum_{x\in\mathcal{X}} p'_x(t)\log\left(\frac{p_x(t)}{q_x(t)}\right) + \sum_{x\in\mathcal{X}} p'_x(t) - \sum_{x\in\mathcal{X}} p_x(t)\frac{q'_x(t)}{q_x(t)}\\
&= \sum_{x,y\in\mathcal{X}} \left(p_y(t) + p_y(t)\log\left(\frac{p_x(t)}{q_x(t)}\right) - p_x(t)\frac{q_y(t)}{q_x(t)}\right) G_{y,x}\\
&\quad - \sum_{x,y\in\mathcal{X}} p_y(t)\log\left(\frac{p_y(t)}{q_y(t)}\right) G_{y,x}\\
&= \sum_{x,y\in\mathcal{X}} \left(p_y(t) - p_y(t)\log\left(\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)}\right) - p_x(t)\frac{q_y(t)}{q_x(t)}\right) G_{y,x}\\
&= \sum_{x,y\in\mathcal{X}} \left(\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)} - \frac{p_y(t)q_x(t)}{p_x(t)q_y(t)}\log\left(\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)}\right) - 1\right) p_x(t)\frac{q_y(t)}{q_x(t)} G_{y,x}\\
&= -\sum_{x,y\in\mathcal{X}:x\neq y} \psi\left(\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)}\right) p_x(t)\frac{q_y(t)}{q_x(t)} G_{y,x}.
\end{aligned}$$

Recall that $\psi \geq 0$. Clearly, $q_x, p_x \geq 0$ for all $x \in \mathcal{X}$, and $G_{y,x} \geq 0$ for all $x \neq y$. It follows that $\frac{d}{dt}R(p(t)\|q(t)) \leq 0$.

Next observe that $\psi(1/z)z$ tends to infinity as z goes to zero. If the probability vectors $p(0)$, $q(0)$ are not equivalent, that is, if there is $x \in \mathcal{X}$ such that $p_x(0) = 0$ and $q_x(0) > 0$, or $p_x(0) > 0$ and $q_x(0) = 0$, then $\frac{d}{dt}\left. R(p(t)\|q(t))\right|_{t=0} = -\infty$ and the above equalities continue to hold in the sense that $-\infty = -\infty$.

It remains to show that $\frac{d}{dt}R(p(t)\|q(t)) = 0$ if and only if $p(t) = q(t)$. We claim that this follows from the fact that $\psi \geq 0$ with $\psi(z) = 0$ if and only if $z = 1$, and from the irreducibility of G. Indeed, $p(t) = q(t)$ if and only if $\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)} = 1$ for all $x, y \in \mathcal{X}$ with $x \neq y$. Thus $p(t) = q(t)$ implies $\frac{d}{dt}R(p(t)\|q(t)) = 0$. If $\frac{d}{dt}R(p(t)\|q(t)) = 0$ then immediately $\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)} = 1$ for all $x, y \in \mathcal{X}$ such that $G_{y,x} > 0$. If y does not directly communicate with x then, by irreducibility, there is a chain of directly communicating states leading from y to x, and using those states it follows that $\frac{p_y(t)q_x(t)}{p_x(t)q_y(t)} = 1$. $\square$

If $q(0) = \pi$ then, by stationarity, $q(t) = \pi$ for all $t \geq 0$. Lemma 3.1 then implies that the mapping

$$p \mapsto R\,(p\|\pi) \tag{3.2}$$

is a strict Lyapunov function for the linear forward equation (3.1). This is, however, just one of many possibilities. Indeed, a second example is

$$p \mapsto R\,(\pi\|p)\,.$$

Yet a third can be constructed as follows. Let $T > 0$ and consider the mapping

$$p \mapsto R\,(p\|q_p(T))\,, \tag{3.3}$$

where $q_p(\cdot)$ is the solution to Eq. (3.1) with $q_p(0) = p$. Lemma 3.1 also implies that the mapping given by (3.3) is a strict Lyapunov function for Eq. (3.1). This is so because $R\left(p(t)\|q_{p(t)}(T)\right) = R\,(p(t)\|q(t))$, where $q(\cdot)$ is the solution to Eq. (3.1) with $q(0) = p(T)$. Note that (3.2) arises as the limit of (3.3) as T goes to infinity.

3.2 Approximations for nonlinear Markov processes

Since a nonlinear Markov process is not in general equivalent to a time inhomogeneous Markov process, the descent property for relative entropy does not directly apply. Indeed, a review of the proof shows that the descent property relies crucially on the fact that $p(\cdot)$ and $q(\cdot)$ satisfy a forward equation defined in terms of the same rate matrix. However, nonlinear Markov processes can be related to linear Markov processes through the law of large numbers and propagation of chaos, and this suggests an indirect way to obtain natural forms of Lyapunov functions. There are in fact many different ways this connection could be used to suggest forms that are suitable, and this paper will consider two that are arguably the most straightforward.

Consider a fixed point π of (2.5), i.e., $\pi\Gamma(\pi) = 0$, and suppose for the sake of argument that π is either a global or local attractor for the limit dynamics $p'(t) = p(t)\Gamma(p(t))$. Consider a weakly interacting system as in Section 2. In particular, for $N \in \mathbb{N}$, let Γ_N be the infinitesimal generator of a linear Markov process with values in $\mathcal{X}^N$, describing the N-particle system, and suppose that $\Gamma(\cdot)$ is the nonlinear generator of the corresponding limit model. For simplicity, let us assume that the initial distributions for the N-particle systems are of the product form $\otimes^N q$ for some $q \in \mathcal{P}(\mathcal{X})$, and let $\mathbf{p}^N(t)$ be the law of the N-particle system at time t for this initial condition. As before, let $X^{i,N}$ and μ^N denote the ith particle and empirical measure.

We next use the fact that π is a local or global attractor for the limit dynamics. The discussion here is formal, being meant only to motivate the possible forms for Lyapunov functions, which will later be justified by direct verification. We claim that when π is a local or global attractor, a good starting point for the construction of a Lyapunov function is the mapping

$$q \mapsto \frac{1}{N} R\left(\mathbf{p}^N(0) \| \mathbf{p}^N(T)\right), \tag{3.4}$$

where $T \in (0, \infty]$ is chosen to make certain approximations to $\mathbf{p}^N(T)$ plausible and approximately independent of the starting distribution q. The scaling in N is natural, given that the relative entropy would be expected to scale proportionally with N.

As discussed in Lemma 3.1, the mapping $t \mapsto R\left(\mathbf{p}^N(t) \| \mathbf{p}^N(T+t)\right)$ is non-increasing at $t = 0$ (and if $\mathbf{p}^N(0)$ is not the stationary distribution strictly decreasing) for solutions of (2.4). Suppose that (3.4) has a well-defined limit which, ignoring all other dependencies, we write as $F(q)$. Then one might conjecture that $F(\cdot)$ could serve in some sense as a Lyapunov function for $p'(t) = p(t)\Gamma(p(t))$. There are (at least) two potential problems. One is that although $R\left(\mathbf{p}^N(0) \| \mathbf{p}^N(T)\right)$ is strictly decreasing along solutions to (2.4), the strict decrease may be lost in the limit as $N \to \infty$. This can in fact happen, as will be seen in examples considered in Section 5. However, in those examples there are multiple fixed points to (2.5), and F still serves as a Lyapunov function (though not a global one), with a strict decrease at all points that are not fixed points. A second (and more significant) potential problem is that in the interchange of differentiation of time and limits in N, even the non-increasing property may be lost. Suppose for simplicity that $T = \infty$, i.e., that $\mathbf{p}^N(T)$ is the stationary distribution $\boldsymbol{\pi}_N$ of the N-particle system. Let $\mathbf{p}^N(0) = \otimes^N p(0)$. For $\delta > 0$

$$R\left(\mathbf{p}^N(\delta) \| \boldsymbol{\pi}_N\right) \leq R\left(\otimes^N p(0) \| \boldsymbol{\pi}_N\right).$$

Thus the non-increasing property will be inherited by $F(p(t))$ if it turns out that

$$R\left(\otimes^N p(\delta) \| \boldsymbol{\pi}_N\right) \leq R\left(\mathbf{p}^N(\delta) \| \boldsymbol{\pi}_N\right) + o(\delta)N, \tag{3.5}$$

since this will imply $F(p(\delta)) - F(p(0)) \leq o(\delta)$. When (3.5) does not hold, then additional information on the "error" $R\left(\otimes^N p(\delta) \| \boldsymbol{\pi}_N\right) - R\left(\mathbf{p}^N(\delta) \| \boldsymbol{\pi}_N\right)$ would likely be needed to construct a candidate Lyapunov function. Letting again $\mathbf{p}^N(0) = \otimes^N q$, we note that this discussion is also relevant when considering use of the mapping

$$q \mapsto \frac{1}{N} R\left(\mathbf{p}^N(T) \| \mathbf{p}^N(0)\right) \tag{3.6}$$

as a potential Lyapunov function.

Thus two issues that should be addressed if this approach is to be used successfully are (i) how to construct approximations to $\mathbf{p}^N(T)$ that capture as much of the dynamics of the N-particle system as possible, and yet for which the calculation of the relative entropy as $N \to \infty$ is possible, and (ii) ensuring that a bound such as (3.5) is valid. It turns out that the second issue is not symmetric for the two forms (3.4) and (3.6). In fact, based on large deviation calculations one can show that the bound (3.5) typically holds for the former, and that it usually fails for the latter. However, we will not give this calculation since for the problems we consider F is obtained in an explicit form for both cases, and we simply test whether or not F has the desired Lyapunov function property. The asymmetry is discussed further in Remark 5.6.

Thus we return to the first issue, the construction of approximations to $\mathbf{p}^N(T)$. The simplest possible situation is when $\boldsymbol{\pi}_N$, the unique stationary distribution of the N-particle system, is available in some tractable form. One class of models for which this is the case are those connected with Gibbs measures. This class of models will be discussed in Section 5. The next simplest may be to assume that the propagation of chaos approximation is valid over a large enough time interval that the marginals of $\mathbf{p}^N(T)$ are nearly equal to $p(T)$, which is itself approximately equal to π. Now in fact the propagation of chaos considers only the joint distribution of a *fixed* number of particles as $N \to \infty$ (see (2.7)), but we ignore that issue and assume in fact that a good approximation for $\mathbf{p}^N(T)$ is $\otimes^N p(T) \approx \otimes^N \pi$. This crude approximation suggests the candidate Lyapunov function

$$\frac{1}{N} R\left(\mathbf{p}_N(0) \| \mathbf{p}_N(T)\right) \approx \frac{1}{N} R\left(\otimes^N q \| \otimes^N \pi\right) = R\left(q \| \pi\right) \doteq F(q). \tag{3.7}$$

One would expect this to yield a useful Lyapunov function only when the interaction between particles is limited (which is related to but distinct from the strength of the interaction as measured by $1/N$). One can formulate such a class of problems in terms of what we refer to as "slow adaptation," for which (3.7) generically gives a global Lyapunov function. The analysis for this case complements results proved in Veretennikov (2006). While this class of models is clearly not as broad as one would like, it is useful in that when compared with the Lyapunov functions obtained for the models of Section 5 (where both methods apply), one can identify important "correction terms" that are lost by neglecting the correlations between particles.

4 Systems with slow adaptation

Here we consider a class of finite-state nonlinear Markov processes of the type introduced in Section 2 as limit systems, but with a structure we call *slow adaptation*, for which the strength of the nonlinear component is adjusted through a small parameter. The long-time behavior of systems of this type, in the context of nonlinear diffusions arising as limits of weakly interacting Itô diffusions, is studied in Veretennikov (2006) based on coupling arguments and hitting times (and not in terms of a Lyapunov function).

Recall that $\mathcal{X}$ is a finite set with $d \geq 2$ elements. Let $\Gamma(p)$ be the rate matrix for a nonlinear Markov process, and assume that $p \mapsto \Gamma(p)$ is Lipschitz continuous for $p \in \mathcal{P}(\mathcal{X})$. Assume also that $\pi \in \mathcal{P}(\mathcal{X})$ is a fixed point of the corresponding ODE. Then for any $\lambda \in [0,1]$, π is also a fixed point for the ODE defined by $p' = \Gamma(\lambda(p-\pi)+\pi)$. The rate matrix $\Gamma^{\lambda}(p) = \Gamma(\lambda(p-\pi)+\pi)$ corresponds to a version of the original system but with slow adaptation when $\lambda > 0$ is small. With $\lambda \in (0,1]$ fixed, the rate matrices $\Gamma^{\lambda}(p)$, $p \in \mathcal{P}(\mathcal{X})$, determine a family of nonlinear Markov processes. The corresponding forward equation

$$\frac{d}{dt}p^{\lambda}(t) = p^{\lambda}(t)\Gamma^{\lambda}(p^{\lambda}(t)) \tag{4.1}$$

has a unique solution given any initial distribution $p(0) \in \mathcal{P}(\mathcal{X})$. We are interested in the question of when π is a stable fixed point for sufficiently slow adaptation.

Following the discussion of the last section, the mapping

$$F(p) = R\left(p\|\pi\right) \tag{4.2}$$

is a natural candidate for $\lambda > 0$ but small.

Proposition 4.1. *Let $p^{\lambda}(\cdot)$ be defined by (4.1). Then there is $\lambda_0 > 0$ such that if $\lambda \in [0, \lambda_0]$, then for all $t \geq 0$*

$$\frac{d}{dt}F(p^{\lambda}(t)) \leq 0,$$

with a strict inequality if and only if $p^{\lambda}(t) \neq \pi$.

Proof. By construction and hypothesis, there is $C < \infty$ such that for all $x, y \in \mathcal{X}$, all $\lambda > 0$, all $p \in \mathcal{P}(\mathcal{X})$,

$$\left|\Gamma^{\lambda}_{yx}(p) - \Gamma_{yx}(\pi)\right| \leq \lambda C \|p - \pi\|,$$

where $\|p - \pi\| \doteq \sum_x |p_x - \pi_x|$. Using calculations similar to those in the proof of Lemma 3.1, with $p^{\lambda}(t) = p$ we have

$$\begin{aligned}
\frac{d}{dt} R\big(p^{\lambda}(t)\|\pi\big) &= \sum_{x,y\in\mathcal{X}} p_y \left(\log\left(\frac{p_x}{\pi_x}\right) + 1\right) \Gamma^{\lambda}_{yx}(p) \\
&= \sum_{x,y\in\mathcal{X}:x\neq y} p_y \left(\log\left(\frac{p_x\pi_y}{p_y\pi_x}\right) - \frac{p_x\pi_y}{p_y\pi_x} + 1\right) \Gamma_{yx}(\pi) \\
&\qquad + \sum_{x,y\in\mathcal{X}} p_y \log\left(\frac{p_x}{\pi_x}\right) \left(\Gamma^{\lambda}_{yx}(p) - \Gamma_{yx}(\pi)\right) \\
&= \sum_{x,y\in\mathcal{X}:x\neq y} p_y \left(\log\left(\frac{p_x\pi_y}{p_y\pi_x}\right) - \frac{p_x\pi_y}{p_y\pi_x} + 1\right) \Gamma_{yx}(\pi) \\
&\qquad + \sum_{x,y\in\mathcal{X}:x\neq y} p_y \log\left(\frac{p_x\pi_y}{p_y\pi_x}\right) \left(\Gamma^{\lambda}_{yx}(p) - \Gamma_{yx}(\pi)\right).
\end{aligned}$$

For $x, y \in \mathcal{X}$ with $x \neq y$ set

$$\begin{aligned}
\gamma_{yx}(p) &\doteq p_y \left(\log\left(\frac{p_x\pi_y}{p_y\pi_x}\right) - \frac{p_x\pi_y}{p_y\pi_x} + 1\right) \Gamma_{yx}(\pi), \\
\rho^{\lambda}_{yx}(p) &\doteq p_y \log\left(\frac{p_x\pi_y}{p_y\pi_x}\right) \left(\Gamma^{\lambda}_{yx}(p) - \Gamma_{yx}(\pi)\right).
\end{aligned}$$

Thus, we have to show that there is $\lambda_0 > 0$ such that

$$\sum_{x,y:x\neq y} \left(\gamma_{yx}(p) + \rho^{\lambda}_{yx}(p)\right) \leq 0 \quad \text{for all } \lambda \in [0, \lambda_0], \tag{4.3}$$

with equality if and only if $p = \pi$.

We first use the fact that $s \mapsto \log(s) - s + 1$ is smooth and concave for $s \in [0, \infty)$, with the maximum value of zero at $s = 1$. Since $\mathcal{P}(\mathcal{X})$ is compact, we find that there is $c > 0$, not depending on p, such that

$$\sum_{x,y\in\mathcal{X}:x\neq y} \gamma_{yx}(p) \leq -c\|p - \pi\|^2.$$

This clearly implies that

$$\sum_{x,y:x\neq y} \gamma_{yx}(p) \leq -\frac{c}{2}\|p-\pi\|^2 + \frac{1}{2}\sum_{x,y:x\neq y} \gamma_{yx}(p). \tag{4.4}$$

Set $\gamma_{\min} \doteq \min\{\Gamma_{yx}(\pi) : \Gamma_{yx}(\pi) > 0\}$. Then $\gamma_{\min} > 0$ since $\Gamma(\pi)$, being irreducible and finite, has a minimal strictly positive entry. Notice that $\max_{x\in\mathcal{X}} \frac{1}{\pi_x} < \infty$ since $\pi_{\min} \doteq \min_{x\in\mathcal{X}} \pi_x > 0$.

Let $x, y \in \mathcal{X}$, $x \neq y$, and set $z \doteq \frac{p_x \pi_y}{p_y \pi_x}$. We distinguish two cases. If $z \geq \frac{1}{2}$ or $\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi) \geq 0$, then, since $|\log(s)| \leq 2|s-1|$ for all $s \geq \frac{1}{2}$,

$$\begin{aligned}
\rho^\lambda_{yx}(p) &= p_y \log(z)\big(\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi)\big) \\
&\leq 2p_y \left|\frac{p_x\pi_y}{p_y\pi_x} - 1\right| \big|\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi)\big| \\
&= \frac{2}{\pi_x}|p_x\pi_y - p_y\pi_x|\big|\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi)\big| \\
&\leq \frac{2}{\pi_{\min}}\left(\pi_y|p_x - \pi_x| + \pi_x|\pi_y - p_y|\right) C\lambda\|p-\pi\|.
\end{aligned}$$

If $\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi) < 0$ and $z \in [0, \frac{1}{2})$, then $\Gamma_{yx}(\pi) \geq \gamma_{\min}$ since $\Gamma^\lambda_{yx}(p) \geq 0$ for $x \neq y$, and, since $\log(s) - s + 1 \leq 0$ for all $s \geq 0$,

$$\begin{aligned}
\tfrac{1}{2}\gamma_{yx}(p) + \rho^\lambda_{yx}(p) &= p_y\left(\tfrac{1}{2}\left(\log(z) - z + 1\right)\Gamma_{yx}(\pi) + \log(z)\big(\Gamma^\lambda_{yx}(p) - \Gamma_{yx}(\pi)\big)\right) \\
&\leq p_y\left(\tfrac{1}{2}\left(\log(z) - z + 1\right)\gamma_{\min} + |\log(z)|2C\lambda\right) \\
&\leq \tfrac{1}{2}p_y\left(-|\log(z)|\left(\gamma_{\min} - 4C\lambda\right) + (1-z)\gamma_{\min}\right).
\end{aligned}$$

This quantity is non-positive for $z \in [0, \frac{1}{2})$ whenever $\lambda \leq \frac{\gamma_{\min}}{16C}$. Consequently, recalling inequality (4.4), we have for $\lambda \in [0, \frac{\gamma_{\min}}{16C}]$,

$$\begin{aligned}
&\sum_{x,y:x\neq y}\left(\gamma_{yx}(p) + \rho^\lambda_{yx}(p)\right) \\
\leq\ & -\frac{c}{2}\|p-\pi\|^2 + \frac{2C\lambda}{\pi_{\min}}\|p-\pi\| \sum_{x,y:x\neq y}\left(\pi_y|p_x - \pi_x| + \pi_x|\pi_y - p_y|\right) \\
\leq\ & -\frac{c}{2}\|p-\pi\|^2 + \frac{2C\lambda}{\pi_{\min}}\|p-\pi\| \left(\sum_{x\in\mathcal{X}}|p_x - \pi_x| + \sum_{y\in\mathcal{X}}|\pi_y - p_y|\right) \\
\leq\ & -\frac{c}{2}\|p-\pi\|^2 + \frac{4C\lambda}{\pi_{\min}}\|p-\pi\|^2.
\end{aligned}$$

This last quantity is strictly negative if $\lambda < \min\{\frac{\gamma_{\min}}{16C}, c\frac{\pi_{\min}}{8C}\}$ and $p \neq \pi$, and zero for $p = \pi$. Choosing $\lambda_0 \in (0, \min\{\frac{\gamma_{\min}}{16C}, c\frac{\pi_{\min}}{8C}\})$, we find that (4.3) holds, with equality if and only if $p = \pi$. $\square$

The bound on λ is obviously conservative, and better bounds that depend on $\Gamma(\pi)$ and π can be found.

5 Systems of Gibbs type

In this section we carry out the program outlined in Section 3 for a class of problems where the stationary distribution for the N-particle system is known. Subsection 5.1 introduces the class of weakly interacting Markov processes and the corresponding nonlinear Markov processes. The construction starts from the definition of the stationary distribution as a Gibbs measure for the N-particle system. In Subsection 5.2 we derive candidate Lyapunov functions for the limit systems as limits of relative entropy. As we will see, (3.4) is suitable while (3.6) is not, and there are interesting terms that are not needed when one restricts to systems with slow adaptation as in the last section. In Subsection 5.3, the functions obtained from (3.4) are shown to be indeed strict Lyapunov functions for the limit systems. Subsection 5.4 contains a discussion of Tamura (1987), where the somewhat analogous situation of nonlinear Itô diffusions of Gibbs type is studied.

5.1 The prelimit and limit systems

Recall that $\mathcal{X}$ is a finite set with $d \geq 2$ elements. Let $V : \mathcal{X} \mapsto \mathbb{R}$, $W : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ be functions, representing the *environment potential* and the *interaction potential*, respectively. Assume throughout that W is zero on the diagonal, that is, $W(x,x) = 0$ for all $x \in \mathcal{X}$. It turns out that the same models at both the prelimit and limit are obtained for the symmetrized version of W as for W itself, and hence we assume that $W(x,y) = W(y,x)$ for all $(x,y) \in \mathcal{X} \times \mathcal{X}$.

Let $\beta > 0$ and $N \in \mathbb{N}$. The functions W, V induce a Gibbs distribution with interaction parameter β on the N-particle state space $\mathcal{X}^N$ given by

$$\text{(5.1)} \qquad \boldsymbol{\pi}_N(\mathbf{x}) \doteq \frac{1}{Z_N} \exp\left(-\sum_{i=1}^{N} V(x_i) - \frac{\beta}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} W(x_i, x_j) \right),$$

where

$$Z_N \doteq \sum_{\mathbf{x}\in\mathcal{X}^N} \exp\left(-\sum_{i=1}^N V(x_i) - \frac{\beta}{N}\sum_{i=1}^N\sum_{j=1}^N W(x_i,x_j)\right)$$

is the normalizing constant (also known as the *partition function*).

There are standard methods for identifying $\mathcal{X}^N$-valued Markov processes for which $\boldsymbol{\pi}_N$ is the stationary distribution. The resulting rate matrices are often called Glauber dynamics; see, for instance, Stroock (2005, § 5.4) or Martinelli (1999, § 3.1). To be precise, we seek an $\mathcal{X}^N$-valued Markov process which has the structure of a weakly interacting N-particle system and is reversible with respect to $\boldsymbol{\pi}_N$. To start, define a function $U_N\colon \mathcal{X}^N \mapsto \mathbb{R}$, the N-particle *energy function*, by

$$U_N(\mathbf{x}) \doteq \sum_{i=1}^N V(x_i) + \frac{\beta}{N}\sum_{i,j=1}^N W(x_i,x_j).$$

In view of (5.1), for all $\mathbf{x},\mathbf{y}\in\mathcal{X}^N$,

$$\frac{\boldsymbol{\pi}_N(\mathbf{x})}{\boldsymbol{\pi}_N(\mathbf{y})} = e^{U_N(\mathbf{y})-U_N(\mathbf{x})}. \tag{5.2}$$

Let $A = (\alpha(x,y))_{x,y\in\mathcal{X}}$ be an irreducible and symmetric matrix with diagonal entries equal to zero and off-diagonal entries either one or zero. A will identify those states of a single particle that can be reached in one jump from any given state. For $N\in\mathbb{N}$, define a matrix $\mathbf{A}_N = (\mathbf{A}_N(\mathbf{x},\mathbf{y}))_{\mathbf{x},\mathbf{y}\in\mathcal{X}^N}$ indexed by elements of $\mathcal{X}^N$ according to $\mathbf{A}_N(\mathbf{x},\mathbf{y}) = \alpha(x_l,y_l)$ if $\mathbf{x}$ and $\mathbf{y}$ differ in exactly one index $l\in\{1,\ldots,N\}$, and $\mathbf{A}_N(\mathbf{x},\mathbf{y}) = 0$ otherwise. Then $\mathbf{A}_N$ determines which states of the $\boldsymbol{N}$-particle system can be reached in one jump, as discussed in Section 2. Observe that $\mathbf{A}_N$ is symmetric and irreducible with values in $\{0,1\}$.

Recall that $W(x,x) = 0$ for all $x\in\mathcal{X}$. Let $\mathbf{x},\mathbf{y}\in\mathcal{X}^N$ be such that $\mathbf{A}_N(\mathbf{x},\mathbf{y}) = 1$. Then $x_l \neq y_l$ for exactly one $l\in\{1,\ldots,N\}$. Using this property and the symmetry of W

$$\begin{aligned}
&U_N(\mathbf{y}) - U_N(\mathbf{x}) \\
&= \sum_{i=1}^{N} (V(y_i) - V(x_i)) + \frac{\beta}{N} \sum_{i,j=1}^{N} (W(y_i, y_j) - W(x_i, x_j)) \\
&= V(y_l) - V(x_l) + \frac{\beta}{N} \sum_{j=1}^{N} (W(y_l, y_j) - W(x_l, x_j) + W(y_j, y_l) - W(x_j, x_l)) \\
&= V(y_l) - V(x_l) - \frac{\beta}{N} (W(y_l, x_l) + W(x_l, y_l)) \\
&\quad + \frac{\beta}{N} \sum_{j=1}^{N} (W(y_l, x_j) - W(x_l, x_j) + W(x_j, y_l) - W(x_j, x_l)) \\
&= V(y_l) - V(x_l) \\
&\quad + \frac{2\beta}{N} \sum_{z \in \mathcal{X}} (W(y_l, z) - W(x_l, z)) \left(\mu_{\mathbf{x}}^N - \frac{1}{N} \delta_{x_l} \right) (\{z\}),
\end{aligned}$$

where δ_{x_l} denotes the Dirac measure at x_l and $\mu_{\mathbf{x}}^N$ is the empirical measure of $\mathbf{x}$ as in (2.2), i.e., $\mu_{\mathbf{x}}^N(\{z\}) = \#\{j \in \{1, \dots, N\} : x_j = z\}/N$. Let $\mathcal{M}_1(\mathcal{X})$ denote the set of subprobability measures on $\mathcal{X}$. Define $\Psi\colon \mathcal{X} \times \mathcal{X} \times \mathcal{M}_1(\mathcal{X}) \mapsto \mathbb{R}$ by

$$\Psi(x, y, \nu) \doteq V(y) - V(x) + 2\beta \sum_{z \in \mathcal{X}} \big(W(y, z) - W(x, z)\big) \nu_z.$$

Then for $\mathbf{x}, \mathbf{y} \in \mathcal{X}^N$ as above,

$$U_N(\mathbf{y}) - U_N(\mathbf{x}) = \Psi\left(x_l, y_l, \mu_{\mathbf{x}}^N - \tfrac{1}{N} \delta_{x_l}\right). \tag{5.3}$$

There are many ways one can define a rate matrix $\mathbf{\Gamma}_N$ that corresponds to $\boldsymbol{\pi}_N$. Three standard ones are as follows. For $\mathbf{x}, \mathbf{y} \in \mathcal{X}^N$, $\mathbf{x} \neq \mathbf{y}$, set either

$$\mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y}) \doteq e^{-(U_N(\mathbf{y}) - U_N(\mathbf{x}))^+} \mathbf{A}_N(\mathbf{x}, \mathbf{y}) \tag{5.4a}$$

$$\text{or} \qquad \mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y}) \doteq \left(1 + e^{U_N(\mathbf{y}) - U_N(\mathbf{x})}\right)^{-1} \mathbf{A}_N(\mathbf{x}, \mathbf{y}) \tag{5.4b}$$

$$\text{or} \qquad \mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y}) \doteq \frac{1}{2} \left(1 + e^{-(U_N(\mathbf{y}) - U_N(\mathbf{x}))}\right) \mathbf{A}_N(\mathbf{x}, \mathbf{y}). \tag{5.4c}$$

In all three cases set $\mathbf{\Gamma}_N(\mathbf{x}, \mathbf{x}) \doteq -\sum_{\mathbf{y} \neq \mathbf{x}} \mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in \mathcal{X}^N$. The model defined by (5.4a) is sometimes referred to as *Metropolis dynamics*, and (5.4b) as *heat bath dynamics* (Martinelli, 1999, p. 110). In what follows we will consider only (5.4a), the analysis for the other dynamics being completely analogous.

The matrix $\mathbf{\Gamma}_N$ is the generator of an irreducible continuous-time finite-state Markov process with state space $\mathcal{X}^N$. In view of Eq. (5.3), for $\mathbf{x}, \mathbf{y} \in \mathcal{X}^N$ such that $x_l \neq y_l$ for some $l \in \{1, \ldots, N\}$, $x_j = y_j$ for $j \neq l$,

$$\mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y}) = e^{-\left(\Psi\left(x_l, y_l, \mu^N_{\mathbf{x}} - \frac{1}{N}\delta_{x_l}\right)\right)^+} \mathbf{A}_N(\mathbf{x}, \mathbf{y}).$$

Thus the jump rate depends on the components x_j, $j \neq l$, only through the empirical measure $\mu^N_{\mathbf{x}}$, and so for each of the three choices in (5.4), $\mathbf{\Gamma}_N$ is the generator of a family of weakly interacting Markov processes in the sense of Section 2. Indeed, in the notation of that section, $\mathbf{\Gamma}_N$ is defined in terms of a function $\Phi_N \colon \mathcal{X} \times \mathcal{X} \times \mathcal{P}_N(\mathcal{X}) \mapsto [0, \infty)$ given by

$$\Phi_N(x, y, p) = e^{-\left(\Psi\left(x, y, p - \frac{1}{N}\delta_x\right)\right)^+}.$$

The rate matrix $\mathbf{\Gamma}_N$ has $\boldsymbol{\pi}_N$ as its stationary distribution. Let $\mathbf{x}, \mathbf{y} \in \mathcal{X}^N$. By symmetry, $\mathbf{A}_N(\mathbf{x}, \mathbf{y}) = \mathbf{A}_N(\mathbf{y}, \mathbf{x})$. Taking into account Eq. (5.2), it is easy to see that for any of the three choices of $\mathbf{\Gamma}_N$ according to (5.4) we have $\boldsymbol{\pi}_N(\mathbf{x})\mathbf{\Gamma}_N(\mathbf{x}, \mathbf{y}) = \boldsymbol{\pi}_N(\mathbf{y})\mathbf{\Gamma}_N(\mathbf{y}, \mathbf{x})$. Thus $\mathbf{\Gamma}_N$ satisfies the detailed balance condition with respect to $\boldsymbol{\pi}_N$, and $\boldsymbol{\pi}_N$ is its unique stationary distribution.

In order to illustrate the construction we consider an example with three states and near-neighbor transitions.

Example 5.1. Suppose that $\mathcal{X} = \{1, 2, 3\}$. Set $\alpha(x, y) = 1$ if and only if $|x - y| = 1$ and zero otherwise. Let $r_1, r_2 \in (0, 1)$, $w_3 \in \mathbb{R}$. Define functions V, W according to

$$V(1) \doteq 0, \quad V(2) \doteq -\log(r_1), \quad V(3) \doteq -\log(r_2) - \log(r_1),$$

$$W(3, 2) = W(2, 3) \doteq w_3, \quad W(x, y) \doteq 0 \textit{ otherwise.}$$

Then

$$\begin{aligned} &\Psi(1, 2, p) = -\log(r_1) + 2\beta w_3 p_3, && \Psi(2, 3, p) = -\log(r_2) + 2\beta w_3 (p_2 - p_3), \\ &\Psi(2, 1, p) = -\Psi(1, 2, p), && \Psi(3, 2, p) = -\Psi(2, 3, p). \end{aligned}$$

Consider the models (5.4a) and (5.4b), respectively, in the limit $N \to \infty$. For the model (5.4a), if $w_3 \geq 0$ and $r_2 < e^{-2\beta w_3}$ then the transition rates will asymptotically be

$$\begin{array}{ll} r_1 e^{-2\beta w_3 p_3} & \textit{for transition from } x_l = 1 \textit{ to } y_l = 2, \\ r_2 e^{-2\beta w_3 (p_2 - p_3)} & \textit{for transition from } x_l = 2 \textit{ to } y_l = 3, \\ 1 & \textit{for transition from } x_l = 2 \textit{ to } y_l = 1, \\ 1 & \textit{for transition from } x_l = 3 \textit{ to } y_l = 2, \end{array}$$

where p corresponds to the empirical measure. For the model (5.4b), if $r_1, r_2 < e^{-2\beta w_3}$ then the transition rates will asymptotically be

$$\begin{array}{ll} \frac{1}{2}\left(1 + r_1 e^{-2\beta w_3 p_3}\right) & \textit{for transition from } x_l = 1 \textit{ to } y_l = 2, \\ \frac{1}{2}\left(1 + r_2 e^{-2\beta w_3 (p_2 - p_3)}\right) & \textit{for transition from } x_l = 2 \textit{ to } y_l = 3, \\ \frac{1}{2}\left(1 - r_1 e^{2\beta w_3 p_3}\right) & \textit{for transition from } x_l = 2 \textit{ to } y_l = 1, \\ \frac{1}{2}\left(1 - r_2 e^{2\beta w_3 (p_2 - p_3)}\right) & \textit{for transition from } x_l = 3 \textit{ to } y_l = 2. \end{array}$$

For $p \in \mathcal{P}(\mathcal{X})$ let $\Gamma(p)$ denote the rate matrix $(\Gamma_{x,y}(p))_{x,y\in\mathcal{X}}$ given by

$$\Gamma_{x,y}(p) \doteq e^{-(\Psi(x,y,p))^+} \alpha(x,y), \quad x \neq y, \tag{5.5}$$

and with diagonal entries $\Gamma_{x,x}(p) \doteq -\sum_{y\in\mathcal{X}} \Gamma_{x,y}(p)$, $x \in \mathcal{X}$. If $p \in \mathcal{P}(\mathcal{X})$ is fixed then $\Gamma(p)$ is the generator of an ergodic finite-state Markov process, and the unique invariant distribution on $\mathcal{X}$ is given by $\pi(p)$ with

$$\pi(p)_x \doteq \frac{1}{Z(p)} \exp\left(-V(x) - 2\beta \sum_{y\in\mathcal{X}} W(x,y) p_y\right), \tag{5.6}$$

where

$$Z(p) \doteq \sum_{x\in\mathcal{X}} \exp\left(-V(x) - 2\beta \sum_{y\in\mathcal{X}} W(x,y) p_y\right).$$

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space and for $N \in \mathbb{N}$ let $\mathbf{X}^N = (X^{1,N}, \ldots, X^{N,N})$ be an $\mathcal{X}^N$-valued càdlàg Markov process with generator $\mathbf{\Gamma}_N$. As in Section 2, μ^N is the empirical measure process corresponding to $\mathbf{X}^N$.

Theorem 2.1 implies that the sequence $(\mu^N)_{N\in\mathbb{N}}$ of $D([0,\infty), \mathcal{P}(\mathcal{X}))$-valued random variables satisfies a law of large numbers with limit determined by Eq. (2.5), the nonlinear forward equation. More precisely, if $\mu^N(0)$

converges in distribution to $q \in \mathcal{P}(\mathcal{X})$ as N goes to infinity then $\mu^N(\cdot)$ converges in distribution to the (deterministic) solution $p(\cdot)$ of

$$\frac{d}{dt}p(t) = p(t)\Gamma(p(t)), \quad p(0) = q.$$

The nonlinear generator $\Gamma(\cdot)$ above, or in Eq. (2.5), is now defined according to (5.5). Thus $\Gamma(\cdot)$ describes the limit model for the families of weakly interacting Markov processes of Gibbs type introduced above.

5.2 Limit of relative entropies

We will construct a candidate Lyapunov function for the nonlinear limit model as the limit of normalized relative entropies according to (3.4). To this end, for $N \in \mathbb{N}$, define $\tilde{F}_N : \mathcal{P}(\mathcal{X}) \mapsto [0, \infty]$ by

$$\tilde{F}_N(p) \doteq \frac{1}{N} R\left(\otimes^N p \middle\| \boldsymbol{\pi}_N\right). \tag{5.7}$$

Theorem 5.2. *For $N \in \mathbb{N}$, let $\tilde{F}_N$ be defined according to* (5.7)*. Then there is a constant $C \in \mathbb{R}$ such that for all $p \in \mathcal{P}(\mathcal{X})$,*

$$\lim_{N\to\infty} \tilde{F}_N(p) = R\left(p \| \pi(p)\right) - \log(Z(p)) - \beta \sum_{x,y\in\mathcal{X}} W(x,y)p_x p_y - C.$$

Proof. Let $p \in \mathcal{P}(\mathcal{X})$. By definition of relative entropy and (5.6),

$$\begin{aligned}
\tilde{F}_N(p) &= \frac{1}{N} \sum_{\mathbf{x}\in\mathcal{X}^N} \prod_{i=1}^N p_{x_i} \log\left(\frac{\prod_{i=1}^N p_{x_i}}{\boldsymbol{\pi}_N(x)}\right) \\
&= \frac{1}{N} \sum_{\mathbf{x}\in\mathcal{X}^N} \prod_{i=1}^N p_{x_i} \left(\sum_{i=1}^N \log(p_{x_i})\right) + \frac{1}{N}\log(Z_N) \\
&\quad + \frac{1}{N} \sum_{\mathbf{x}\in\mathcal{X}^N} \prod_{i=1}^N p_{x_i} \left(\sum_{i=1}^N V(x_i) + \frac{\beta}{N}\sum_{i,j=1}^N W(x_i,x_j)\right).
\end{aligned}$$

Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. $\mathcal{X}$-valued random variables with common distribution p defined on some probability space. Then

$$\frac{1}{N} \sum_{\mathbf{x}\in\mathcal{X}^N} \prod_{i=1}^N p_{x_i} \left(\sum_{i=1}^N \log(p_{x_i})\right) = \mathbf{E}\left[\frac{1}{N}\sum_{i=1}^N \log\left(p_{X_i}\right)\right] = \mathbf{E}\left[\log\left(p_{X_1}\right)\right],$$

$$\frac{1}{N}\sum_{\mathbf{x}\in\mathcal{X}^N}\prod_{i=1}^N p_{x_i}\left(\sum_{i=1}^N V(x_i)\right) = \mathbf{E}\left[\frac{1}{N}\sum_{i=1}^N V(X_i)\right] = \mathbf{E}\left[V(X_1)\right],$$

and as $N \to \infty$

$$\begin{aligned}\frac{\beta}{N}\sum_{\mathbf{x}\in\mathcal{X}^N}\prod_{i=1}^N p_{x_i}\left(\frac{1}{N}\sum_{i,j=1}^N W(x_i,x_j)\right) &= \mathbf{E}\left[\frac{\beta}{N^2}\sum_{i,j=1}^N W(X_i,X_j)\right]\\ &= \beta\frac{N^2-N}{N^2}\,\mathbf{E}\left[W(X_1,X_2)\right]\\ &\to \beta\,\mathbf{E}\left[W(X_1,X_2)\right].\end{aligned}$$

In order to compute the limit of $\frac{1}{N}\log(Z_N)$, define a bounded and continuous mapping $\Phi\colon \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ by

$$\Phi(q) \doteq \beta \sum_{x,y\in\mathcal{X}} W(x,y)q_x q_y.$$

Let $\tilde{\mu}^N$ be the empirical measure of N independent and identically distributed $\mathcal{X}$-valued random variables with common distribution ν given by $\nu_x \doteq \exp(-V(x))/\tilde{Z}$, $x \in \mathcal{X}$, where $\tilde{Z} \doteq \sum_{x\in\mathcal{X}}\exp(-V(x))$. Then

$$\log(Z_N) = \log \mathbf{E}\left[\exp\left(-N\Phi(\tilde{\mu}^N)\right)\right] + N\log(\tilde{Z}).$$

By Sanov's theorem and Varadhan's lemma on the asymptotic evaluation of exponential integrals (Dupuis and Ellis, 1997), it follows that

$$\lim_{N\to\infty}\frac{1}{N}\log(Z_N) = -\inf_{q\in\mathcal{P}(\mathcal{X})}\left\{R(q\|\nu)+\Phi(q)\right\} + \log(\tilde{Z}) \doteq -C.$$

Note that C is finite and does not depend on p.

Recalling that X_1, X_2 are independent random variables with common distribution p, we find

$$\begin{aligned}\lim_{N\to\infty}\tilde{F}_N(p) &= \mathbf{E}\left[\log\left(p_{X_1}\right)\right] + \mathbf{E}\left[V(X_1)\right] + \beta\,\mathbf{E}\left[W(X_1,X_2)\right] - C\\ &= \sum_{x\in\mathcal{X}} p_x\log(p_x) + \sum_{x\in\mathcal{X}} V(x)p_x + \beta\sum_{x,y\in\mathcal{X}} W(x,y)p_x p_y - C.\end{aligned}$$

On the other hand, by the definition of relative entropy and (5.1),

$$\begin{aligned}R\left(p\|\pi(p)\right) = \log(Z(p)) &+ \sum_{x\in\mathcal{X}} p_x\log(p_x) + \sum_{x\in\mathcal{X}} V(x)p_x\\ &+ 2\beta\sum_{x,y\in\mathcal{X}} W(x,y)p_x p_y.\end{aligned}$$

Therefore

$$\lim_{N\to\infty} \tilde{F}_N(p) = R\left(p \| \pi(p)\right) - \log(Z(p)) - \beta \sum_{x,y\in\mathcal{X}} W(x,y) p_x p_y - C.$$

$\square$

The constant C appearing in the expression for $\lim_{N\to\infty} \tilde{F}_N(p)$ in Theorem 5.2 does not depend on p. We therefore define a candidate Lyapunov function F for the limit model by

$$F(p) \doteq R\left(p \| \pi(p)\right) - \log(Z(p)) - \beta \sum_{x,y\in\mathcal{X}} W(x,y) p_x p_y. \tag{5.8}$$

In the proof of Theorem 5.2 it was incidentally noted that F also takes the form

$$F(p) = \sum_{x\in\mathcal{X}} p_x \log(p_x) + \sum_{x\in\mathcal{X}} V(x) p_x + \beta \sum_{x,y\in\mathcal{X}} W(x,y) p_x p_y. \tag{5.9}$$

Since the first sum in (5.9) equals minus entropy of p, F is the sum of a convex function, an affine function and a quadratic function on $\mathcal{P}(\mathcal{X})$. This fact is useful in determining if the fixed points of F are stable or not. Here we will use it to characterize the fixed points. Recall that W has zero diagonal elements, and as in Section 2 identify $\mathcal{X}$ with $\{1, \ldots, d\}$ and $\mathcal{P}(\mathcal{X})$ with $\{p \in \mathbb{R}^d : p_x \geq 0 \text{ and } \sum_{x=1}^d p_x = 1\}$. Let ∂_x denote ∂/∂_{p_x}. From (5.9) it follows that for $p \in \mathcal{P}(\mathcal{X})$ such that $p_x > 0$ for all $x \in \mathcal{X}$,

$$\partial_x F(p) = \log(p_x) + 1 + V(x) + 2\beta \sum_{y\neq x} W(x,y) p_y. \tag{5.10}$$

Set $\mathcal{P}_{tan}(\mathcal{X}) \doteq \{\nu \in \mathbb{R}^d : \sum_{i=1}^d \nu_i = 0\}$. The directional derivative in any direction $\nu \in \mathcal{P}_{\tan}(\mathcal{X})$ is given by

$$\frac{\partial}{\partial \nu} F(p) = \sum_{x\in\mathcal{X}} \nu_x \left(\log(p_x) + V(x) + 2\beta \sum_{y\neq x} W(x,y) p_y \right).$$

If the derivatives of F in all directions in $\mathcal{P}_{tan}(\mathcal{X})$ vanish at some point $p \in \mathcal{P}(\mathcal{X})$ then p is an equilibrium point for Eq. (2.5), the nonlinear forward equation. The converse is true as well.

Theorem 5.3. *Let $p \in \mathcal{P}(\mathcal{X})$. Then p is an equilibrium point for* (2.5) *if and only if p is non-degenerate and $\frac{\partial}{\partial \nu} F(p) = 0$ for all $\nu \in \mathcal{P}_{tan}(\mathcal{X})$.*

Proof. Recall that $\pi(p)$ is the unique invariant probability associated with $\Gamma(p)$, and hence $\pi(p)\Gamma(p) = 0$. First observe that p is an equilibrium point for Eq. (2.5) if and only if $p\Gamma(p) = 0$, which can be true if and only if $p = \pi(p)$. If $p = \pi(p)$ then $p_x > 0$ for all $x \in \mathcal{X} \equiv \{1, \ldots, d\}$ and so (5.10) holds. Let $x, y \in \mathcal{X}$, $x \neq y$. Then by (5.10),

$$(5.11) \qquad \begin{aligned}(\partial_x - \partial_y)F(p) = \log(p_x) + V(x) + 2\beta \sum_{z \neq x} W(x,z)p_z \\ - \log(p_y) - V(y) + 2\beta \sum_{z \neq y} W(y,z)p_z.\end{aligned}$$

Note that $(\partial_x - \partial_y)$ is the derivative in direction $v \in \mathcal{P}_{tan}(\mathcal{X})$ with $v_x \doteq 1$, $v_y \doteq -1$, $v_z \doteq 0$ for $z \in \mathcal{X} \setminus \{x, y\}$.

Now assume that $\frac{\partial}{\partial v}F(p) = 0$ for all $v \in \mathcal{P}_{tan}(\mathcal{X})$. In view of (5.1), the definition of $\pi(p)$ in (5.6), and (5.11) it follows that for all $i, j \in \{1, \ldots, d\}$, $i \neq j$,

$$\log\left(\frac{p_i}{p_j}\right) = \log\left(\frac{\pi(p)_i}{\pi(p)_j}\right),$$

which implies $p = \pi(p)$ since p, $\pi(p)$ are both probability distributions.

On the other hand, if $p = \pi(p)$ then, by (5.11), $(\partial_i - \partial_j)F(p) = 0$ for all $i, j \in \{1, \ldots, d\}$, $i \neq j$, which implies $\frac{\partial}{\partial v}F(p) = 0$ for all $v \in \mathcal{P}_{tan}(\mathcal{X})$. $\square$

According to Theorem 5.3, the equilibrium points of the forward equation (2.5) are precisely the critical points of F on $\mathcal{P}(\mathcal{X})$. If there is no interaction (i.e., $\beta = 0$ or $W \equiv 0$) then the generator $\Gamma(p)$ does not depend on p and Eq. (2.5) has a unique equilibrium point, namely the unique stationary distribution associated with Γ. From Eq. (5.9) and the strict concavity of entropy it follows that in this case F is a strictly convex function.

In general, Eq. (2.5) can have multiple equilibria, stable and unstable. Let us illustrate this by a simple example.

Example 5.4. Assume that $\mathcal{X} = \{1, 2\}$, $V \equiv 0$ and $W(1,2) = 1 = W(2,1)$. Then $F(p) = f(p_1)$ with

$$f(x) \doteq x\log(x) + (1-x)\log(1-x) + 2\beta(1-x)x, \quad x \in [0,1].$$

The critical points of F on $\mathcal{P}(\{1,2\})$ are in a one-to-one correspondence with the critical points of f on $[0,1]$. Clearly, $f(0) = 0 = f(1)$, and for $x \in (0,1)$,

$$f'(x) = \log(x) - \log(1-x) + 2\beta - 4\beta x, f''(x) = \frac{1}{x} + \frac{1}{1-x} - 4\beta.$$

Moreover, $f'(x) \to -\infty$ as x tends to zero, $f'(x) \to \infty$ as x tends to one.

If $\beta \leq 1$ then f has exactly one critical point, namely a global minimum at $x = \frac{1}{2}$. If $\beta > 1$ then there are three critical points, one local maximum at $x = \frac{1}{2}$ and two minima at x_β and $1 - x_\beta$, respectively, for some $x_\beta \in (0, \frac{1}{2})$, where $x_\beta \to \frac{1}{2}$ as $\beta \searrow 1$, $x_\beta \to 0$ as β goes to infinity.

As a consequence of Theorem 5.3 and Theorem 5.5 below, if $\beta > 1$ then the two minima of f correspond to stable equilibria of the forward equation, while the local maximum corresponds to an unstable equilibrium.

5.3 Lyapunov function property

The function F defined in (5.8) is a strict Lyapunov function for the dynamics of the limit model as given by Eq. (2.5), the nonlinear forward equation.

Theorem 5.5. *Let $p(\cdot)$ be a solution to the forward equation* (2.5) *with some initial distribution $p(0) \in \mathcal{P}(\mathcal{X})$. Then for all $t \geq 0$,*

$$\frac{d}{dt} F(p(t)) = \left. \frac{d}{dt} R\left(p(t) \| \pi(q)\right) \right|_{q=p(t)} \leq 0.$$

Moreover, $\frac{d}{dt} F(p(t)) = 0$ if and only if $p(t) = \pi(p(t))$.

Proof. We will show that if $p(\cdot)$ is the solution to Eq. (2.5) with $p(0) = q$ then

$$\left. \frac{d}{dt} F(p(t)) \right|_{t=0} = \left. \frac{d}{dt} R\left(p(t) \| \pi(q)\right) \right|_{t=0}. \tag{5.12}$$

In view of the semigroup property of solutions to the ODE (2.5), the forward equation, and since q is arbitrary, validity of (5.12) implies $\frac{d}{dt} F(p(t)) = \frac{d}{dt} R\left(p(t) \| \pi(q)\right)$, $q = p(t)$, for all $t \geq 0$.

Let $p(0) = q$. By Eq. (5.9) and since $\sum_{x \in \mathcal{X}} p'_x(0) = 0$,

$$\begin{aligned} \left. \frac{d}{dt} F(p(t)) \right|_{t=0} &= \sum_{x \in \mathcal{X}} \log(q_x) p'_x(0) + \sum_{x \in \mathcal{X}} V(x) p'_x(0) \\ &\quad + \beta \sum_{x,y \in \mathcal{X}} W(x,y) p'_x(0) q_y + \beta \sum_{x,y \in \mathcal{X}} W(x,y) p'_y(0) q_x. \end{aligned}$$

On the other hand, by the definition of relative entropy and (5.6),

$$\begin{aligned} \left. \frac{d}{dt} R\left(p(t) \| \pi(q)\right) \right|_{t=0} &= \sum_{x \in \mathcal{X}} \log(q_x) p'_x(0) + \sum_{x \in \mathcal{X}} V(x) p'_x(0) \\ &\quad + \beta \sum_{x,y \in \mathcal{X}} W(x,y) p'_x(0) q_y + \beta \sum_{x,y \in \mathcal{X}} W(y,x) p'_x(0) q_y. \end{aligned}$$

Since $\sum_{x,y} W(y,x)p'_x(0)q_y = \sum_{x,y} W(x,y)p'_y(0)q_x$, it follows that (5.12) holds.

The rest of the assertion follows from the observation that $\pi(q)$ is the stationary distribution for the (linear) Markov family associated with $\Gamma(q)$ and from the strict Lyapunov property of relative entropy in case of ergodic Markov processes, see Lemma 3.1 in Section 3. $\square$

Remark 5.6. It was noted in Subsection 3.2 that the reversed form of relative entropy is not as suitable a starting point for the construction of Lyapunov functions. In the present setting use of the reverse form would mean defining a candidate Lyapunov function by

$$\hat{F}(p) \doteq \lim_{N\to\infty} \frac{1}{N} R\left(\boldsymbol{\pi}_N \| \otimes^N p\right).$$

Suppose that $\sum_{x,y\in\mathcal{X}} W(x,y)p_x p_y$ is convex and that $\bar{p}$ is the unique solution to $p\Gamma(p) = 0$ in $\mathcal{P}(\mathcal{X})$. Using the Lyapunov function F it is not hard to check that under π_N the empirical measure converges in probability to $\bar{p}$, and that

$$\hat{F}(p) = R\left(\bar{p} \| p\right) + \log Z(\bar{p}).$$

In contrast to $F(p)$, this function does not always serve as a global Lyapunov function, confirming the asymmetry in this setting as well.

Remark 5.7. Consider the slow adaptation setting of Section 4 for the limit dynamics associated with a Gibbs measure. Choosing Metropolis dynamics as above, we thus start from a family of rate matrices $\Gamma(p)$, $p \in \mathcal{X}$, defined according to (5.5). Suppose that p^* is a fixed point of the mapping $p \mapsto \pi(p)$. For $\lambda \in [0,1]$, $p \in \mathcal{X}$, set $\Gamma^\lambda(p) \doteq \Gamma(p^* + \lambda(p - p^*))$. The rate matrices $\Gamma^\lambda(p)$ are again of Gibbs type, that is, $\Gamma^\lambda(p)$ is defined according to (5.5) but with different potentials in place of V, W, namely $V^{\lambda,\beta}$, W^λ given by

$$V^{\lambda,\beta}(x) \doteq V(x) + 2\beta(1-\lambda)\sum_{z\in\mathcal{X}} W(x,z)p_z^*, \quad W^\lambda(x,y) \doteq \lambda W(x,y).$$

Fix $\lambda \geq 0$. Then Eq. (5.9) and Theorem 5.5 imply that

$$\begin{aligned}
F^\lambda(p) &\doteq \sum_{x\in\mathcal{X}} p_x \log(p_x) + \sum_{x\in\mathcal{X}} V^{\lambda,\beta}(x)p_x + \beta \sum_{x,y\in\mathcal{X}} W^\lambda(x,y)p_x p_y \\
&= \sum_{x\in\mathcal{X}} p_x \log(p_x) + \sum_{x\in\mathcal{X}} \left(V(x) + 2\beta(1-\lambda)\sum_{z\in\mathcal{X}} W(x,z)p_z^* \right) p_x \\
&\quad + \lambda\beta \sum_{x,y\in\mathcal{X}} W(x,y)p_x p_y
\end{aligned}$$

is a (strict) Lyapunov function for the nonlinear forward equation

$$\frac{d}{dt}p(t) = p(t)\Gamma^{\lambda}(p(t))$$

associated with $\Gamma^{\lambda}(p)$, $p \in \mathcal{X}$. Proposition 4.1, on the other hand, implies that $\bar{F}(p) = R\left(p\|p^*\right)$ is also a strict Lyapunov function when λ is positive but sufficiently small. Since $p^* = \pi(p^*)$, by (5.6)

$$\begin{aligned} R\left(p\|p^*\right) &= \sum_{x\in\mathcal{X}} p_x \log(p_x) - \sum_{x\in\mathcal{X}} p_x \log(p_x^*) \\ &= \sum_{x\in\mathcal{X}} p_x \log(p_x) + \log(Z(p^*)) + \sum_{x\in\mathcal{X}} \left(V(x) + 2\beta \sum_{z\in\mathcal{X}} W(x,z) p_z^* \right) p_x, \end{aligned}$$

which is equal to $F^{\lambda}(p) + \log(Z(p^*))$ for $\lambda = 0$. Observe that the term $\log(Z(p^*))$ has no impact on the Lyapunov function property as it does not depend on p.

5.4 Comparison with existing results for Itô diffusions

A situation analogous to that of this section is considered in Tamura (1987), where the author studies the long-time behavior of "nonlinear" Itô-McKean diffusions of the form

(5.13)
$$dX(t) = -\left(\nabla V\big(X(t)\big) + 2\beta \int_{\mathbb{R}^d} \nabla_1 W\big(X(t), y\big)\big)\mu_t(dy) \right) dt + \sqrt{2} dB(t),$$

where $\mu_t = \mathrm{Law}(X(t))$, B is a standard d-dimensional Wiener process, V a function $\mathbb{R}^d \mapsto \mathbb{R}$, the environment potential, and W a *symmetric* function $\mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ with zero diagonal, the interaction potential. Signs and constants have been chosen in analogy with the finite-state models considered here. Solutions of Eq. (5.13) arise as weak limits of the empirical measure processes associated with weakly interacting Itô diffusions. The N-particle model is described by the system

$$dX^{i,N}(t) = -\nabla V\big(X^{i,N}(t)\big)dt - \frac{2\beta}{N} \sum_{j=1}^{N} \nabla_1 W\big(X^{i,N}(t), X^{j,N}(t)\big)dt + \sqrt{2} dB^i(t),$$

where $i \in \{1, \ldots, N\}$, $B^1, \ldots, B^N$ are independent standard Brownian motions, and ∇_1 denotes gradient with respect to the first $\mathbb{R}^d$-valued variable.

In Tamura (1987) a candidate Lyapunov function $F\colon \mathcal{P}(\mathbb{R}^d) \mapsto [0,\infty]$ is introduced without explicit motivation, and then shown to be in fact a valid Lyapunov function. This function, which appears to be a close analogue of the functions derived in this section as the limits of relative entropies, takes the following form. If μ is a probability measure that is absolutely continuous with respect to Lebesgue measure and of the form $f_\mu(x)dx$, then

(5.14)
$$F(\mu) \doteq \int \log\left(f_\mu(x)\right) f_\mu(x)dx + \int V(x)\mu(dx) + \int\int W(x,y)\mu(dx)\mu(dy).$$

In all other cases $F(\mu) \doteq \infty$.

There are, however, some interesting differences. The most significant of these is the form taken by the derivative of the composition of the Lyapunov function with the solution to the forward equation. In Tamura (1987) the descent property is established by expressing the orbital derivatives of F in terms of the Donsker-Varadhan rate function associated with the empirical measures of solutions to Eq. (5.13), when the measure μ_t is frozen at $\mu \in \mathcal{P}(\mathbb{R}^d)$. In contrast, in our case the orbital derivative of the Lyapunov function is equal to the orbital derivative of relative entropy with respect to the invariant distribution $\pi(p)$ that is obtained when the dynamics of the nonlinear Markov process are frozen at p. The relationship we have derived in fact carries over to the diffusion case in the sense that for all $t \geq 0$,

$$\frac{d}{dt}F(\mu_t) = \frac{d}{dt}R\left(\mu_t \| \pi_\nu\right)_{\nu=\mu_t}, \tag{5.15}$$

where $\mu_t = \mathrm{Law}(X(t))$, X being the solution to Eq. (5.13) for some (absolutely continuous) initial condition, and $\pi_\nu \in \mathcal{P}(\mathbb{R}^d)$ is given by

$$\pi_\nu(dx) \doteq \frac{1}{Z_\nu}\exp\left(-V(x) - 2\beta\int_{\mathbb{R}^d} W(x,y)\nu(dy)\right)dx \tag{5.16}$$

with Z_ν the normalizing constant. Clearly, the probability measures given by (5.16) correspond to the distributions $\pi(p) \in \mathcal{P}(\mathcal{X})$ defined in (5.6). Relationship (5.15) can be established in a way analogous to the proof of Theorem 5.3. On the other hand, the relationship between orbital derivative of the Lyapunov function and Donsker-Varadhan rate function as established in Tamura (1987) for the diffusion case does *not* carry over to the finite-state Gibbs models studied above.

6 Concluding remarks

In this paper we have exploited a connection between nonlinear Markov process and approximating N-particle systems to generate Lyapunov functions as a limit of normalized relative entropies. Explicit identification was carried out for two classes of models, based on very different approximations. However, there are other quite natural approximations that should be considered so that the technique can be developed more broadly. The goal of these approximations should be to bring in more information on the correlations between particles when the N-particle system is "quasi-stationary," but hopefully without requiring that full information on the stationary distribution be available. For example, as an improvement on the very crude product form approximation used for the case of slow adaptation, one might consider the mapping

$$p \mapsto \lim_{N\to\infty} \frac{1}{N} R\left(\otimes^N p \| q_N(T)\right),$$

where $q_N(\cdot)$ is the solution to Eq. (2.4) with $q_N(0) = \otimes^N \pi$, with π a fixed point of $\Gamma(\cdot)$.

References

N. Antunes, C. Fricker, P. Robert, and D. Tibi. Stochastic networks with multiple stable points. *Ann. Probab.*, 36(1):255–278, 2008.

M. Benaïm and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65 (11-12):823–838, 2008.

P. Dai Pra, W. J. Runggaldier, E. Sartori, and M. Tolotti. Large portfolio losses: A dynamic contagion model. *Ann. Appl. Probab.*, 19(1):347–394, 2009.

D. A. Dawson, J. Tang, and Y. Q. Zhao. Balancing queues by mean field interaction. *Queueing Syst.*, 49:335–361, 2005.

P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, 1997.

T. D. Frank. Lyapunov and free energy functionals of generalized Fokker-Planck equations. *Phys. Lett. A*, 290:93–100, 2001.

A. D. Gottlieb. *Markov transitions and the propagation of chaos.* PhD thesis, Lawrence Berkeley National Laboratory, 1998.

C. Graham. McKean-Vlasov Itô-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes Appl.*, 40(1):69–82, 1992.

C. Graham and P. Robert. Interacting multi-class transmissions in large stochastic networks. *Ann. Appl. Probab.*, 19(6):2334–2361, 2009.

M. Kac. Foundations of kinetic theory. In J. Neyman, editor, *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, volume 3, pages 171–197, Berkeley, California, 1956.

T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.*, 7(1):49–58, 1970.

S. N. Laughton and A. C. C. Coolen. Macroscopic Lyapunov functions for separable stochastic neural networks with detailed balance. *J. Statist. Phys.*, 80(1-2):375–387, 1995.

F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In P. Bernard, editor, *Lectures on probability theory and statistics (Saint-Flour XXVII, 1997)*, volume 1717 of *Lecture Notes in Mathematics*, pages 93–191, Berlin, 1999. Springer-Verlag.

D. R. McDonald and J. Reynier. Mean field convergence of a model of multiple TCP connections through a buffer implementing RED. *Ann. Appl. Probab.*, 16(1):244–294, 2006.

K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984.

F. Spitzer. *Random Fields and Interacting Particle Systems.* Mathematical Association of America, Washington, D.C., 1971. Notes on lectures given at the 1971 MAA Summer Seminar, Williamstown, Mass.

D. W. Stroock. *An Introduction to Markov Processes*, volume 230 of *Graduate Texts in Mathematics.* Springer, Berlin, 2005.

Y. Tamura. Free energy and the convergence of distributions of diffusion processes of McKean type. *J. Fac. Sci. Univ. Tokyo, Sect. IA, Math.*, 34 (2):443–484, 1987.

A. Y. Veretennikov. On ergodic measures for McKean-Vlasov stochastic equations. In H. Niederreiter and D. Talay, editors, *Monte Carlo and quasi-Monte Carlo methods 2004*, pages 471–486, Berlin, 2006. Springer.