Progress in Probability, Vol. 59, 19–22

A Bakry-Emery Criterion for Self-Interacting Diffusions

Michel Benaïm and Olivier Raimond

Abstract. We give a Bakry-Emery type criterion for self-interacting diffusions on a compact manifold.

Mathematics Subject Classification (2000). missing.

Keywords. missing.

Let M be a smooth compact connected Riemannian manifold without boundary and $V : M \times M \to \mathbb{R}$ a smooth function. For every Borel probability measure μ on M let $V\mu : M \to \mathbb{R}$ denote the function defined by $V\mu(x) = \int_M V(x,u)\mu(du)$, and let $\nabla(V\mu)$ denote its gradient.

A *self-interacting diffusion process* associated to V is a continuous-time stochastic process $\{X_t\}$ which is a solution on M to the stochastic differential equation

$$dX_t = dW_t(X_t) - \frac{1}{2}\nabla(V\mu_t)(X_t)dt, \quad X_0 = x \in M,$$

where (W_t) is a Brownian vector field on M, and $\mu_t = \frac{1}{t}\int_0^t \delta_{X_s} ds$ is the *empirical occupation measure* of $\{X_t\}$.

This type of process with reinforcement was introduced in [2] and further studied in [3], [4], with the ultimate goal to:

- **(a)** provide tools allowing us to analyze the long term behavior of $\{\mu_t\}$,
- **(b)** understand the relations connecting this behavior to the nature of V, and,
- **(c)** the geometry of M.

Let $\mathcal{P}(M)$ denote the space of Borel probability measures over M, λ the Riemannian probability on M and $\mathcal{P}_{cd}(M) \subset \mathcal{P}(M)$ the set of measures having a continuous density with respect to λ. Let X_V be the vector field defined on $\mathcal{P}_{cd}(M)$ by

$$X_V(\mu) = -\mu + \Pi_V(\mu)$$

We acknowledge financial support from the Swiss National Science Foundation grant 200021-1036251/1.

where

$$\frac{d\Pi_V(\mu)}{d\lambda} = \frac{e^{-V\mu}}{\int_M e^{-V\mu(y)}\lambda(dy)}.$$

Point **(a)** was mainly addressed in [2] where it was shown that the asymptotic behavior of $\{\mu_t\}$ can be precisely[1] described in terms of the deterministic dynamical system induced by X_V.

Depending on the nature of V, the dynamics of X_V can either be convergent, globally convergent or non-convergent, leading to a similar behavior for $\{\mu_t\}$. A key step toward **(b)** is the next result recently proved in [4].

Theorem 1. *Suppose* V *is a symmetric function. Then the limit set of* $\{\mu_t\}$ *(for the topology of weak* convergence) is almost surely a connected subset of* $X_V^{-1}(0) = \mathsf{Fix}(\Pi_V)$.

In (the generic) case where the equilibrium set $X_V^{-1}(0)$ is finite, Theorem 1 implies that $\{\mu_t\}$ converges almost surely. If furthermore, $X_V^{-1}(0)$ reduces to a singleton $\{\mu^*\}$, then $\{\mu_t\}$ converges almost surely to μ^* and we say that $\{\mu_t\}$ is *globally convergent*.

A function $K : M \times M \to \mathbb{R}$ is a *Mercer kernel* provided K is continuous symmetric and defines a positive operator in the sense that

$$\int_{M\times M} K(x,y)f(x)f(y)\lambda(dx)\lambda(dy) \geq 0$$

for all $f \in L^2(\lambda)$. The following result is proved in [4].

Theorem 2. *Assume that (up to an additive constant)* V *is a Mercer Kernel. Then* $\{\mu_t\}$ *is* globally convergent.

Example. Suppose $M \subset \mathbb{R}^n$ and $V(x,y) = f(-||x-y||^2)$ where $||\cdot||$ is the Euclidean norm of $\mathbb{R}^n$ and $f : \mathbb{R} \mapsto \mathbb{R}^+$ is a smooth function whose derivatives of all order $f', f'', \ldots$ are nonnegative. Then it was proved by Schoenber [6] that V is a Mercer Kernel.

As observed in [4] the assumption that V is a Mercer Kernel seems well suited to describe *self-repelling diffusions*. On the other hand, it is not clearly related to the geometry of M (see, e.g., the preceding example).

The next theorem has a more geometrical flavor and is robust to smooth perturbations (of M and V). It can be seen as a Bakry-Emery type condition [1] for self-interacting diffusions and is a first step toward **(c)**.

Theorem 3. *Assume that* V *is symmetric and that for all* $x \in M, y \in M, u \in T_xM, v \in T_yM$

$$\mathsf{Ric}_x(u,u) + \mathsf{Ric}_y(v,v) + \mathsf{Hess}_{x,y}V((u,v),(u,v)) \geq K(||u||^2 + ||v||^2)$$

where K *is some positive constant. Then* $\{\mu_t\}$ *is* globally convergent.

[1] We refer the reader to this paper for more details and mathematical statements.

Proof. Let $\mathcal{P}_{ac}(M)$ denote the set of probabilities which are absolutely continuous with respect to λ and let J be the nonlinear free energy function defined on $\mathcal{P}_{ac}(M)$ by

$$J(\mu) = \mathsf{Ent}(\mu) + \frac{1}{2}\int_{M\times M} V(x,y)\mu(dx)\mu(dy)$$

where

$$\mathsf{Ent}(\mu) = \int_M \log\left(\frac{d\mu}{d\lambda}\right) d\mu.$$

The key point is that $X_V^{-1}(0)$ is the critical set of J (restricted to $\mathcal{P}_{cd}(M)$) as shown in [4] (Proposition 2.9). On the other hand, the condition given in the theorem makes J a *displacement K-convex* function in the sense of McCann [5]. Let us briefly explain this latter statement.

Let d_2^W denote the L^2 Wasserstein distance on $\mathcal{P}(M)$ (see, e.g., [7] or [8]). Given $\nu^0, \nu^1 \in \mathcal{P}_{ac}(M)$ McCann [5] proved that there exists a unique geodesic path $t \to \nu^t$ in $(\mathcal{P}_{ac}(M), d_2^W)$ and that ν^t is the image of ν^0 by a map of the form $F_t(x) = \exp_x(t\Phi)$ where Φ is some vector field. Moreover,

$$d_2^W(\nu^0,\nu^t)^2 = \int_M d(x,F_t(x))^2\nu^0(dx).$$

Set $j(t) = J(\nu^t) = e(t) + \frac{v(t)}{2}$ with $e(t) = \mathsf{Ent}(\nu^t)$ and

$$v(t) = \int_{M\times M} V(x,y)\nu^t(dx)\nu^t(dy) = \int_{M\times M} V(F_t(x),F_t(y))\nu^0(dx)\nu^0(dy).$$

Sturm [7] recently proved the beautiful result that

$$\partial^2 e(t) = \int_M \mathsf{Ric}(\dot{F}_t(x),\dot{F}_t(x))\nu^0(dx)$$

where $\partial^2 e(t) := \liminf_{s\to 0} \frac{1}{s^2}(e(t+s) - 2e(t) + e(t-s))$. Clearly

$$\partial^2 v(t) = \int_{M\times M} \mathsf{Hess}_{F_t(x),F_t(y)} V\left((\dot{F}_t(x),\dot{F}_t(y)),(\dot{F}_t(x),\dot{F}_t(y))\right)\nu^0(dx)\nu^0(dy).$$

Hence, under the assumption of Theorem 3,

$$\partial^2 j(t) \geq \frac{K}{2}\int_{M\times M}(||\dot{F}_t(x)||^2 + ||\dot{F}_t(y)||^2)\nu^0(dx)\nu^0(dy) = K d_2^W(\nu^0,\nu^1)^2.$$

In particular, j is strictly convex. It then follows that J (respectively X_V) has a unique minimum (respectively equilibrium). □

Example. Let $M = S^n \subset \mathbb{R}^{n+1}$ be the unit sphere of dimension $n, f : \mathbb{R} \mapsto \mathbb{R}$ a smooth convex function and

$$V(x,y) = f(-||x-y||^2) = g(\langle x,y\rangle)$$

with $g(t) = f(2t-2)$. By invariance of λ under the orthogonal group $O(n+1)$ it is easily seen (see, e.g., Lemma 4.6 of [2]) that $V\lambda$ is a constant map. Hence $\lambda \in X_V^{-1}(0)$ and here, global convergence means convergence to λ.

For all $(x,y) \in M \times M$, $(u,v) \in T_xM \times T_yM$,

$$\mathsf{Hess}_{(x,y)}V((u,v),(u,v)) = g''(\langle x,y\rangle)\left(\langle x,v\rangle + \langle y,u\rangle\right)^2 + g'(\langle x,y\rangle)\left(2\langle u,v\rangle - (||u||^2 + ||v||^2)\langle x,y\rangle\right).$$

Set $t = \langle x,y\rangle$ and assume (without loss of generality) that $||u||^2 + ||v||^2 = 1$. Then $|2\langle u,v\rangle| \le 1$ and the last term on the right-hand side of the preceding equality is bounded below by $-tg'(t) - |g'(t)|$. Therefore the condition of Theorem 3 reads

$$tg'(t) + |g'(t)| < 2(n-1) \tag{1}$$

while Theorem 2 would lead to

$$g^{(k)}(t) \ge 0 \,\forall k \in \mathbb{N},\ |t| \le 1. \tag{2}$$

Remark that condition (1) makes J a displacement-convex function while (2) makes J convex in the usual sense. Of course, none of these conditions is optimal. For instance, suppose that $g(t) = at$. Then (1) reads $|a| < n-1$, and (2) reads $a \ge 0$. On the other hand, this example can be fully analyzed and it was shown in [2] that $\mu_t \to \lambda$ for $a > -(n+1)$ while $\{\mu_t\}$ converges to a "Gaussian" measure with random center, for $a < -(n+1)$.

References

[1] D. Bakry, M. Emery, *Hypercontractivité des semi-groupes de diffusion*, C.R. Acad. Sci. Paris, I, **299** (1984), 775–778.

[2] M. Benaim, M. Ledoux and O. Raimond, *Self-interacting diffusions*, Probab. Theor. Relat. Fields, **122** (2002), 1–41.

[3] M. Benaim and O. Raimond, *Self-interacting diffusions II: Convergence in law*, Annales de l'Institut Henri-Poincaré, **6** (2003), 1043–1055.

[4] M. Benaim and O. Raimond, *Self-interacting diffusions III: Symmetric interactions*, Annals of Probability, **33 (5)** (2005), 1716–1759.

[5] R. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., **11** (2001), 589–608.

[6] I. J. Schoenberg, *Metric spaces and completely monotone functions*, Ann. of Math., **39** (1938), 811–841.

[7] K. T. Sturm, *Convex functionals of probability measures and nonlinear diffusions on manifolds*, J. Math. Pures. Appl., **84** (2005), 149–168.

[8] C. Villani, *Topics in Mass Transportation*, Graduate studies in Mathematics, AMS, 2003.

Michel Benaïm
Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11
Case postale 2, CH-2007 Neuchâtel, Switzerland
e-mail: `michel.benaim@unine.ch`

Olivier Raimond
Laboratoire de Modélisation Stochastique et Statistique, Université Paris Sud, France
e-mail: `Olivier.Raimond@math.u-psud.fr`