On ergodic measures for McKean–Vlasov stochastic equations 2

A. Yu. Veretennikov *

December 15, 2003

4 Introduction 2

This is a continuation of the preprint [24]. All numbers of equations, sections, theorems are continued, too; however, one can read this text without any look at the first part, because everything - assumptions, etc. - which is required has been repeated here. The references are identical except for additional [24] and [25] which have got last two additional numbers in the reference list. Remind that we consider the McKean-Vlasov equation in $\mathbb{R}^d$,

$$dX_t = b[X_t, \mu_t]\,dt + \,dW_t,\ X_0 = x_0 \in \mathbb{R}^d, \tag{25}$$

where $b[x, \mu] := \int b(x, y)\,\mu(dy)$ for any measure μ (this is a notation convention), with locally Borel functions $b(\cdot, \cdot) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, and d-dimensional Wiener process W_t. Here μ_t is the marginal distribution of X_t. Related equations which provide other descriptions of the problem are the nonlinear equation for measures,

$$\partial_t \mu_t = L^*(\mu_t)\mu_t, \tag{26}$$

with

$$L(\mu) = \Delta/2 + b[x, \mu]\partial_x,$$

and the approximation N-particle equation,

$$dX_t^{i,N} = \frac{1}{N}\sum_{j=1}^{N} b(X_t^{i,N}, X_t^{j,N})\,dt + \,dW_t^i,\ X_0^{i,N} = x_0,\ 1 \le i \le N, \tag{27}$$

with d-dimensional independent Wiener processes W_t^i. It is known that under reasonable assumptions the process $X^{i,N}$ converges weakly to the solution of the McKean-Vlasov equation with the same W^i (see [19], [1], [13], [24] et al.),

$$d\bar{X}_t^i = b[\bar{X}_t^i, \mu_t^i]\,dt + \,dW_t^i,\ \bar{X}_0 = x_0 \in \mathbb{R}^d, \tag{28}$$

*School of Mathematics, University of Leeds (UK) & Institute of Information Transmission Problems (Russia); email: veretenn@maths.leeds.ac.uk. This work has been done for the programme “Interaction and Growth in Complex Stochastic Systems” while visiting Isaac Newton Institute for Mathematical Sciences and Clare Hall, University of Cambridge.

where μ_t^i is the law of $\bar{X}_t^i$ (given initial data). Under assumptions below (practically the same as in Theorem 2 from [24]), the law μ_t^i actually does not depend on i because the solution of the equation (25) is unique in law. In this part of the paper we investigate the Euler approximations for the N-particle equation (27), namely,

$$dX_t^{i,N,h} = \frac{1}{N}\sum_{j=1}^{N} b(X_{\kappa_h(t)}^{i,N,h}, X_{\kappa_h(t)}^{j,N,h})\, dt + dW_t^i,\ X_0^{i,N} = x_0,\ 1 \le i \le N, \quad (29)$$

with $\kappa_h(t) := [t/h]h$, where $[a]$ denotes the integer part of the value a, that is, the maximal integer not exceeding a. In terms of the first part, we only consider the "case (2°)" here; its assumptions seem to be more reasonable. The only change in assumptions is a more restrictive boundedness condition (30) on the drift b instead of the linear growth condition in the first part. This is done for simplicity: most probably, everything could be redone under linear growth assumption. However, this requires the analogue of the Lemma 1 from [24] which is correct, but, as we said, for the purposes of establishing the bound (38) below can be simply replaced by (30).

The next section contains the main result, and the section 6 – its proof. The calculus is rather close to that in [24] (although a new feature is of course an approximation scheme). Hence, the most essential is, to the author's mind, not the calculus but a new version of recurrence conditions which imply mixing, see Proposition 2. Implicitly, this method was already used in [23]. In the framework of approximations this is joint result with S. A. Klokov.

5 Main approximation result

We always assume that the function $b(x, \cdot)$ satisfies all *assumptions of existence,* as we said above, a little bit more restrictive than in the first part, namely, it is bounded,

$$\sup_{x,y} |b(x, y)| \le C < \infty, \quad (30)$$

and continuous with respect to the second variable y for any x. Remind the *main assumptions for ergodicity and uniqueness*:

- Coefficient b is decomposed into two parts,

  $$b(x, y) = b_0(x) + b_1(x, y),$$

  where the first part is responsible for the "enviromnent", while the second for the interaction itself. Next assumptions concern both whole b, and separately b_0 and b_1.

- b: recurrence–1

  $$\lim_{|x|\to\infty} \sup_y \langle b(x, y), x\rangle = -\infty, \quad (31)$$

- b: recurrence–2

  $$\lim_{|x|\to\infty} \sup_y \langle b(x, y), x\rangle \le -r < 0, \quad (32)$$

- b_0: attraction to zero which grows at least linearly with distance (= one-sided Lipschitz condition), for any x, x',

$$\sup_y \langle b_0(x) - b_0(x'), x - x' \rangle \le -c_0|x - x'|^2 \qquad (c_0 > 0). \tag{33}$$

  The next two assumptions are required if c_0 is any positive. Instead, one can assume (36) which says that c_0 is large enough, along with (37) saying b_1 is Lipschitz. In this version conditions (34-35) are not used, however, we keep them for completeness.

- b_1: anty-symmetry of interactions,

$$b_1(x, x') - b_1(x', x) = 0. \tag{34}$$

- b_1: "attraction" between particles, which *increases with distance*, in a certain non-rigorous sense

$$\langle (x - x') - (\bar{x} - \bar{x}'), b_1(x, x') - b_1(\bar{x}, \bar{x}') \rangle \le 0. \tag{35}$$

- b_0: *large* attraction to zero,

$$c_0 > C^{b_1}_{Lip}, \tag{36}$$

  where $C^{b_1}_{Lip} < \infty$ is the best constant satisfying

$$\max\left(|b_1(x, y) - b_1(x', y)|, |b_1(y, x) - b_1(y, x')|\right) \le C^{b_1}_{Lip}|x - x'|, \tag{37}$$

  for all x, x', y.

Denote $\mu^{i,N,h}_t$ the law of $X^{i,N,h}_t$. In the next Theorem, where we only consider the case (2°) from [24], Theorem 2, and a notation μ_∞ is used. Under assumptions of this Theorem and according to Theorem 2 from [24], there exists a unique stationary measure for equation (25); μ_∞ denotes this very measure.

Theorem 3 *Let conditions (30-33) and (36-37) be satisfied, with either (32) and $r \ge r(d)$ large enough, or, alternatively, (31) and then $r = -\infty$. Then there is a weak limit*

$$\mu^{1,N,h}_{\ell h} \Longrightarrow \mu^{1,N,h}_\infty, \quad \ell \to \infty,$$

moreover,

$$\mu^{1,N,h}_\infty \Longrightarrow \mu_\infty, \quad N, h^{-1} \to \infty.$$

6 Proof

6.1 Auxuliary results

We first formulate the result from [25] concerning convergence for Euler schemes for SDEs *without interactions*. To this end, we temporarily, – for this subsection only, – assume that in equation (25) the function b *does not depend on the variable* y. This assumption is used in Proposition 1 below without a reminder.

Notice that it is not convenient here to use results from [9] for more general approximations, simply because in the present state they do require a certain smoothness which unfortunately depends on dimension; hence, when $N \to \infty$, we actually need $b \in C^\infty$ in order to use these results. On the contrary, Gaussian case does not require any smoothness. We formulate a rather special case of Theorem 1, item 3, from [25], and only the part which concerns convergence to equilibrium, not mixing (which is the main point in the cited work). Notice that assumption (31) coincides exactly with condition (A_F) from the cited paper in our particular case.

Proposition 1 ([25]) *Under assumption (31), or (32) with* $r > (m-1)d/2$ *and* $m > 4$*, for any* h *small enough there exists the invariant measure* μ^h_∞*, and the following bound for* β*-mixing and for marginal distributions* $\mu^h_{\ell h} = \mathcal{L}(X^h_{\ell h})$ *holds true. Then*

$$\|\mu^h_{\ell h} - \mu^h_\infty\|_{TV} \quad \leq \quad C(1+|x|^m)(1+\ell h)^{-k-1},$$

with some $C > 0$*, and any* $k \in (0, (m-2)/2)$ *(any* $k > 0$ *under (31)).*

Here $\|\cdot\|_{TV}$ is a total variation norm. Underline that here all constants C, m, k may be chosen uniformly in $h \leq 1$. Below we will apply this result to the process $\hat{X}^{N,h}_t$. We will also need the following slightly different version of this result.

Proposition 2 *Let the process* X_t *be decomposed into* K *components of dimensions* $d_1, \ldots, d_K$*,* $X_t = (X^1_t, \ldots, X^K_t)$*, and for each component* X^j *the following recurrence condition holds true:*

$$\langle b^j(x), x^j\rangle \leq -r_j, \quad |x^j| \geq M, \quad r_j > (m-1)d_j/2.$$

Then for any h *small enough there exists the invariant measure* μ^h_∞*, and the following bound for* β*-mixing and for marginal distributions* $\mu^h_{\ell h} = \mathcal{L}(X^h_{\ell h})$ *holds true. If* $m > 4$ *and* $r > (m-1)d/2$*, then*

$$\|\mu^h_{\ell h} - \mu^h_\infty\|_{TV} \quad \leq \quad C(1+|x|^m)(1+\ell h)^{-k-1},$$

with some $C > 0$*, and any* $k \in (0, (m-2)/2)$*.*

Proposition 1 is proved in [25]. Proposition 2 follows from the same calculus in addition to the reasoning from [24], 3.3.E.

6.2 Proof of Theorem 3

A. We are going to show that

$$\sup_t E|X_t^{i,N,h} - \bar{X}_t^i|^2 \le C(N^{-1} + h^{1/2}). \tag{38}$$

The statement of the Theorem then follows from this inequality either directly, – i.e., after analysing ergodic properties of the process $X_t^{i,N,h}$, – or from Theorem 2 in [24], which asserts that $\mu_t \Longrightarrow \mu_\infty$ (where μ_t is the law of $\bar{X}_t^1$). We prefer the latter reasoning because the last statement has been already established. However, ergodic properties of $X_t^{i,N,h}$ will be needed anyway.

Throughout the calculus, whenever we compare the values like $X_t^{i,N,h}$ and $X_s^{i,N,h}$ under expectation, we always have in mind the following bound:

$$E|X_t^{i,N,h} - X_s^{i,N,h}|^2 \le 2\left(E|W_t^i - W_s^i|^2 + \|b\|^2_{L_\infty}|t-s|^2\right).$$

We have,

$$d(X_t^{i,N,h} - \bar{X}_t^i)^2 = 2(X_t^{i,N,h} - \bar{X}_t^i)(b[X_{\kappa_h(t)}^{i,N,h}, \hat{\mu}_t^{N,h}] - b[\bar{X}_t^i, \mu_t^1])\,dt$$

$$= 2(X_t^{i,N,h} - \bar{X}_t^i)(b[X_{\kappa_h(t)}^{i,N,h}, \hat{\mu}_t^{N,h}] - b[\bar{X}_t^i, \mu_t^1])\,dt$$

$$+2(X_t^{i,N,h} - \bar{X}_t^i)(b_0(X_{\kappa_h(t)}^{i,N,h}) - b_0(\bar{X}_t^i))\,dt$$

$$\le \left[2(X_t^{i,N,h} - \bar{X}_t^i)(b[X_{\kappa_h(t)}^{i,N,h}, \hat{\mu}_t^{N,h}] - b[\bar{X}_t^i, \mu_t^1]) - 2c_0|X_t^{i,N,h} - \bar{X}_t^i|^2\right]dt.$$

Hence,

$$d\sum_{i=1}^N (X_t^{i,N,h} - \bar{X}_t^i)^2$$

$$= 2\sum_{i=1}^N (X_t^{i,N,h} - \bar{X}_t^i)(b[X_{\kappa_h(t)}^{i,N,h}, \hat{\mu}_t^{N,h}] - b[\bar{X}_t^i, \mu_t^1])\,dt$$

$$\le -2c_0\sum_{i=1}^N |X_t^{i,N,h} - \bar{X}_t^i|^2\,dt$$

$$+2\sum_{i=1}^N (X_t^{i,N,h} - \bar{X}_t^i)(b[X_{\kappa_h(t)}^{i,N,h}, \hat{\mu}_t^{N,h}] - b[X_t^{i,N,h}, \mu_t^1])\,dt.$$

Therefore,

$$E\sum_{i=1}^N (X_t^{i,N,h} - \bar{X}_t^i)^2 - E\sum_{i=1}^N (X_s^{i,N,h} - \bar{X}_s^i)^2$$

$$= 2E\int_s^t \sum_{i=1}^N (X_r^{i,N,h} - \bar{X}_r^i)(b[X_{\kappa_h(r)}^{i,N,h}, \hat{\mu}_r^{N,h}] - b[\bar{X}_r^i, \mu_r^1])\,dr$$

$$\le -2c_0 E\int_s^t \sum_{i=1}^N |X^{i,N,h}_r - \bar X^i_r|^2\, dr$$

$$+2E\int_s^t \sum_{i=1}^N (X^{i,N,h}_r - \bar X^i_r)(b[X^{i,N,h}_{\kappa_h(r)}, \hat\mu^{N,h}_r] - b[\bar X^i_r, \mu^1_r])\, dr.$$

We have,

$$E\sum_{i=1}^N (X^{i,N,h}_r - \bar X^i_r)(b[X^{i,N,h}_{\kappa_h(r)}, \hat\mu^{N,h}_r] - b[\bar X^i_r, \mu^1_r])$$

$$= E\sum_{i=1}^N (X^{i,N,h}_r - \bar X^i_r)\left(\frac{1}{N}\sum_{j=1}^N (b(X^{i,N,h}_{\kappa_h(r)}, X^{j,N,h}_r) - b[\bar X^i_r, \mu^1_r])\right)$$

$$= E\sum_{i=1}^N (X^{i,N,h}_r - \bar X^i_r)\left(\frac{1}{N}\sum_{j=1}^N (b(X^{i,N,h}_{\kappa_h(r)}, X^{j,N,h}_r) - b(\bar X^i_r, \bar X^j_r))\right)$$

$$+E\sum_{i=1}^N (X^{i,N,h}_r - \bar X^i_r)\left(\frac{1}{N}\sum_{j=1}^N (b(\bar X^i_r, \bar X^j_r) - b[\bar X^i_r, \mu^1_r])\right) =: A_1 + A_2.$$

The second term here possesses the bound (see [24]),

$$|A_2| = |E\sum_{i,j=1}^N (X^{i,N,h}_r - \bar X^i_r)\left(\frac{1}{N}(b(\bar X^i_r, \bar X^j_r) - b[\bar X^i_r, \mu^1_r])\right)|$$

$$\le N^{-1}\sum_{i=1}^N \left(E\left(X^{i,N,h}_r - \bar X^i_r\right)^2\right)^{1/2}\left(E\left(\sum_{j=1}^N b(\bar X^i_r, \bar X^j_r) - b[\bar X^i_r, \mu^1_r]\right)^2\right)^{1/2}$$

$$\le CN^{1/2}\left(E|X^{1,N,h}_r - \bar X^1_r|^2\right)^{1/2}. \quad (39)$$

The first term may be estimated as follows,

$$|A_1| \le 2C^{b_1}_{Lip}\alpha^h(r) + CNh^{1/4}\alpha^{1/2}(r),$$

where $\alpha^h(t) := E(X^{1,N,h}_t - \bar X^1_t)^2$. So, one gets ($t > s$),

$$N\alpha^h(t) - N\alpha^h(s) \le -2c_0\int_s^t N\alpha^h(r)\, dr + C(N^{1/2} + Nh^{1/4})\int_s^t (\alpha^h(r))^{1/2}\, dr.$$

or

$$\alpha^h(t) - \alpha^h(s) \le -2c_0\int_s^t \alpha^h(r)\, dr + C(N^{-1/2} + h^{1/4})\int_s^t (\alpha^h(r))^{1/2}\, dr. \quad (40)$$

This implies

$$\alpha^h(t) \le \frac{C^2}{4(c_0 - C^{b_1}_{Lip})^2}\left(N^{-1/2} + h^{1/4}\right)^2$$

(see, e.g., [24]).

B. The following statement follows directly from the bound (38), and it is natural to formulate it here, although it will not be used in the sequel.

Corollary 1 *Under assumptions of the Theorem 3, for any finite number of indices $i_1 < i_2 \ldots < i_k$,*

$$\left(X_t^{i_1,N,h}, \ldots, X_t^{i_k,N,h}\right) \Longrightarrow \left(\bar{X}_t^{i_1}, \ldots, \bar{X}_t^{i_k}\right), \quad N, h^{-1} \to \infty, \tag{41}$$

uniformly with respect to $t \geq 0$, where the random variables in the right hand side are independent.

C. We are going to show that the *homogeneous Markov process* $\hat{X}_{kh}^{N,h} = (X_{kh}^{1,N,h}, \ldots, X_{kh}^{N,N,h})$ converges to an equilibrium measure $\hat{\mu}_\infty^{N,h}$ in total variation as $kh \to \infty$, with a polynomial rate or faster (depending on $r < \infty$ or $r = \infty$). This convergence is uniform in $s < h$ and $h \leq 1$ (see Proposition 1). Then the measure $\hat{\mu}_\infty^{N,h}$ is stationary for the "big" Markov process $\hat{X}_{kh}^{N,h}$, $k = 0, 1, 2, \ldots$ This implies a convergence for projections, too, namely, $\|\mu_{kh}^{1,N,h} - \mu_\infty^{1,N,h}\|_{\mathrm{TV}} \to 0$ as $k \to \infty$.

Of course, the law of $\hat{X}_{kh+T}^{N,h}$ tends in TV topology to some other limiting measure which depends also on T. However, as $T < h$, the difference between all these limits corresponding to different T's becomes negligible when h tends to zero, just because $\sup_{|t-s| \leq h} E|\hat{X}_t^{N,h} - \hat{X}_s^{N,h}|^2 \leq C_N h$.

We firstly show the ergodicity and mixing under an easier condition (31). Due to the Proposition 1 and similarly to [24], it suffices to notice that the mixing condition (see [23]) for the large (Markov diffusion) process $\hat{X}^N \in \mathbb{R}^{dN}$,

$$\lim_{|\hat{x}^N| \to \infty} \langle \hat{x}^N, \hat{b}^N(\hat{x}^N) \rangle = -\infty, \tag{42}$$

follows directly from (31).

D. Now we will show the same under assumption (32). The approach to establishing uniform beta-mixing bounds as well as convergence rate to a (unique) equilibrium measure in total variation for approximation schemes consists of two estimates,

$$\sup_{h \leq 1} \sup_{t \geq 0} E_{\hat{X}_0^h} |\hat{X}_t^{N,h}|^m 1(t < \tau) \leq C(1 + |\hat{X}_0^{N,h}|^m),$$

and

$$\sup_{h \leq 1} E_{\hat{X}_0} \tau^{k+1} \leq C(1 + |\hat{X}_0|^m),$$

with $\tau := \inf(s : |\hat{X}_t^{N,h}| \leq R)$, for R large enough, with the appropriate k.

Due to Proposition 2, one possible technical tool for establishing these two bounds is the inequality

$$\lim_{|\hat{x}| \to \infty} \sum_{i=1}^N \langle \hat{x}^i, \hat{b}^i(\hat{x}) \rangle \, |\hat{x}^i|^{m-2} = -\infty. \tag{43}$$

The latter follows from (32) exactly as in [24].

Acknowledgements

This work was supported by Isaac Newton Institute for Mathematical Sciences, University of Cambridge, and grant INTAS-99-0590.

References

[1] Benachour, S.; Roynette, B.; Talay, D.; Vallois, P. Nonlinear self-stabilizing processes. I: Existence, invariant probability, propagation of chaos. Stochastic Processes Appl. 75, No.2, 173-201 (1998).

[2] Benachour, S.; Roynette, B.; Vallois, P. Nonlinear self-stabilizing processes. II: Convergence to invariant probability. Stochastic Processes Appl. 75, No.2, 203-224 (1998).

[3] Benedetto, D.; Caglioti, E.; Carrillo, J.A.; Pulvirenti, M. A non-Maxwellian steady distribution for one-dimensional granular media. J. Stat. Phys. 91, No.5-6, 979-990 (1998).

[4] Bossy, M., Talay, D. A stochastic particle method for the McKean-Vlasov and the Burgers equation. Math. Comput. 66, No.217, 157-192 (1997).

[5] Dawson, D. A.; Gärtner, J. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. Stochastics 20, 247-308 (1987).

[6] Funaki, T. A certain class of diffusion processes associated with nonlinear parabolic equations. Z. Wahrscheinlichkeitstheor. Verw. Geb. 67, 331-348 (1984).

[7] Kac, M. Foundations of kinetic theory. Proc. 3rd Berkeley Sympos. Math. Statist. Probability 3, 171-197 (1956).

[8] Klokov, S. A. On lower bounds for mixing rates for a class of Markov Processes, submitted; preprint version at http://www.mathpreprints.com/math/Preprint/klokov/20020920/1/

[9] Klokov, S. A., Veretennikov, A. Yu. Polynomial mixing and convergence rate for a class of stochastic difference equations, submitted

[10] Kotelenez, P. A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation. Probab. Theory Relat. Fields 102, No.2, 159-188 (1995).

[11] Krylov, N. V. Controlled diffusion processes. NY et al., Springer (1980).

[12] Landau, L. D., Lifshitz, E. M. Course of theoretical physics, v. 10, Physical kinetics / by E. M. Lifshitz and L. P. Pitaevskii, Oxford, Pergamon (1981).

[13] Malrieu, F. Logarithmic Sobolev inequalities for some nonlinear PDE's http://www.mathpreprints.com/statistics/0106017/

[14] McKean, H.P. A class of Markov processes associated with nonlinear parabolic equations Proc. Natl. Acad. Sci. USA 56, 1907-1911 (1966).

[15] Nagasawa, M. Schrödinger equations and diffusion theory. Basel: Birkhäuser (1993).

[16] Shiga, T. Stationary distribution problem for interacting diffusion systems. Dawson, D. A. (ed.), Measure-valued processes, stochastic partial differential equations, and interacting systems. Providence, RI: American Mathematical Society. CRM Proc. Lect. Notes. 5, 199-211 (1994).

[17] Skorokhod, A. V. Studies in the theory of random processes. Reading, Mass.: Addison-Wesley (1965).

[18] Skorokhod, A. V. Stochastic equations for complex systems. Dordrecht (Netherlands) etc.: D. Reidel Publishing Company (1988).

[19] Sznitman, A.-S. Topics in propagation of chaos. Calcul des probabilités, Ec. d'été, Saint-Flour/Fr. 1989, Lect. Notes Math. 1464, 165-251 (1991).

[20] Tamura, Y. On asymptotic behavior of the solution of a non-linear diffusion equation. J. Fac. Sci., Univ. Tokyo, Sect. I A 31, 195-221 (1984).

[21] Veretennikov, A.Yu. On strong solutions and explicit formulas for solutions of stochastic integral equations. Math. USSR Sb. 39, 387-403 (1981)

[22] Veretennikov, A.Yu. On polynomial mixing bounds for stochastic differential equations. Stochastic Processes Appl. 70, No.1, 115–127 (1997).

[23] Veretennikov, A. Yu. On polynomial mixing and the rate of convergence for stochastic differential and difference equations. (Russian) Teor. Veroyatnost. i Primenen. 44, 1999, no. 2, 312–327; English translation: Theory Probab. Appl. 44, no. 2, 361–374 (1999).

[24] Veretennikov, A. Yu. On ergodic measures for McKean-Vlasov stochastic equations. preprint NI03066-IGS (2003), http://www.newton.cam.ac.uk/preprints/NI03066.pdf

[25] Veretennikov, A. Yu., Klokov, S. A. Mixing rates for Euler scheme for SDEs, submitted