Stochastic Processes and their Applications 93 (2001) 57–85

stochastic processes and their applications

www.elsevier.com/locate/spa

Homogenization of random parabolic operator with large potential

Fabien Campillo[a,*], Marina Kleptsyna[b], Andrey Piatnitski[c]

[a] *SYSDYS Project, INRIA/LATP, IMT, 38 rue F. Joliot–Curie, 13451 Marseille cedex 20, France*
[b] *Institute for Information Transmission Problems (IPPI), 19 Bolshoy Karetnii, Moscow 101447, Russia*
[c] *Narvik Institute of Technology, HiN, 8505 Narvik, Norway and Lebedev Physical Institute, Leninski prospect 53, Moscow 117924, Russia*

Received 27 September 1999; received in revised form 19 May 2000; accepted 18 October 2000

Abstract

We study the averaging problem for a divergence form random parabolic operators with a large potential and with coefficients rapidly oscillating both in space and time variables. We assume that the medium possesses the periodic microscopic structure while the dynamics of the system is random and, moreover, diffusive. A parameter α will represent the ratio between space and time microscopic length scales. A parameter β will represent the effect of the potential term. The relation between α and β is of great importance. In a trivial case the presence of the potential term will be "neglectable". If not, the problem will have a meaning if a balance between these two parameters is achieved, then the averaging results hold while the structure of the limit problem depends crucially on α (with three limit cases: one classical and two given under martingale problems form). These results show that the presence of stochastic dynamics might change essentially the limit behavior of solutions.

Keywords: Random operators; Homogenization; Averaging

1. Introduction

We study the averaging problem for a random parabolic operators with symmetric elliptic part in the presence of a large potential, all the coefficients being rapidly oscillating functions both in space and time variables.

The homogenization problems for various random structures are widely discussed in the physical and mathematical literature, see, for example, Jikov et al. (1994) and its bibliography.

In multidimensional case the first rigorous results for the divergence form random elliptic operators with stochastically homogeneous coefficients were obtained in Kozlov (1980). Then, another approach was developed in Papanicolaou and Varadhan (1982).

* Corresponding author. Fax: +33-4-91054681.
E-mail addresses: campillo@sophia.inria.fr (F. Campillo), asarin@ippi.ac.msk.su (M. Kleptsyna), andrey@sci.lpi.msk.su (A. Piatnitski).

0304-4149/01/$ - see front matter
PII: S0304-4149(00)00095-8

Later, many other random structures were investigated, among them the random porous media (see, for instance, Hornung, 1997), the equations with lower-order terms in Avellaneda and Majda (1991,1994) Fannjiang and Papanicolaou (1996,1997), systems of equations, nonlinear models and others. An efficient method of homogenization of random structures was proposed in Bourgeat et al. (1994) where the technique developed earlier for periodic micro-structures (Allaire, 1992) was generalized to the random case.

Currently, there are several mathematical approaches which allow to examine homogenization problems in random media, but in all the studied models the randomness in spatial variables and the presence of a group of transformation preserving some probability measure, are supposed.

In our model, we assume that the medium possesses the periodic microscopic structure while the dynamics of the system is random and, moreover, diffusive. The equations without potential were previously considered in Kleptsyna and Piatnitski (1997).

The problems of this type appear, for instance, when one studies the macroscopic behavior of micro-inhomogeneous locally periodic dissipative media whose properties change randomly in time, or the effect of random action on locally periodic structures with large dissipation.

The corresponding parabolic operator is of the form:

$$u_t(t,x) - \operatorname{div}\left[a\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right)\nabla u(t,x)\right] - \varepsilon^{-\beta} c\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) u(t,x), \quad x \in \mathbb{R}^n,$$

where ε is a small positive parameter, α, $\beta > 0$, the coefficients $a^{ij}(z,y)$ and $c(z,y)$ are periodic in the first argument, and ξ_t is a diffusion process with values in $\mathbb{R}^d$ solving the following stochastic differential equation:

$$\mathrm{d}\xi_t = B(\xi_t)\,\mathrm{d}t + \sigma(\xi_t)\,\mathrm{d}W_t.$$

The parameter α represents the ratio between space and time microscopic length scales; in the “self-similar” case $\alpha = 2$, a coupling between space and time averaging appears.

Regarding the process ξ_s, we suppose good mixing and localization properties. Our approach requires a sufficiently fast decay of the density of the invariant measure of ξ_s at the infinity; for instance, a condition of the Khasminski type $B(y)y/|y| < -c$, $c > 0$, is sufficient. In fact, quoting a new work Pardoux and Veretennikov (2001), we will assume a weaker condition:

$$B(y)\frac{y}{|y|} < -c|y|^\mu$$

for some $\mu > -1$ and $c > 0$. The exact assumptions on the regularity of coefficients and the process ξ are given in Section 2.2.

The relation between α and β is of great importance. If $\beta < \alpha/2 \wedge 1$ then the presence of the potential $c(z,y)$ is neglectable in a proper sense. If, on the contrary, $\beta > \alpha/2 \wedge 1$ then, in general, the family of solutions is not compact. The only exception is the case where $\alpha > 2$ and where the mean value of $c(z,y)$ in z is equal to zero for all y. In this case, the proper choice is $\beta = \alpha/2$ instead of $\beta = 1$.

We will show that for $\beta = \alpha/2 \wedge 1$ under the natural regularity assumptions, the averaging results hold while the structure of the limit problem depends crucially on whether $\alpha = 2$, or $\alpha < 2$ or $\alpha > 2$.

If $\alpha > 2$ then the family of solutions of corresponding Cauchy problems converges in probability in a proper functional space to the solution of the Cauchy problem for parabolic operator with constant nonrandom coefficients. This result looks like classical homogenization result with the only difference that we obtain convergence in probability. The almost sure convergence is an open question.

If $\alpha \leqslant 2$ then the family of measures generated by the solutions, converges weakly to the unique solution of the limit martingale problem which involves the one-dimensional Brownian motion. The formula for the coefficients of the limit problem are rather different in the cases $\alpha = 2$ and $\alpha < 2$.

These results show that the presence of stochastic dynamics might change essentially the limit behavior of solutions. It is interesting to note that in a particular case $c(z, y) \equiv 0$ the limit equation is always deterministic Kleptsyna and Piatnitski (1997).

In Section 2 the precise setting of the problem is given and some auxiliary statements are quoted.

The main results of the paper are formulated in Section 3.

The next section is devoted to the proof of the main statements. It should be noted that, in general, the expectation of the norm of u^ε does not admit uniform in ε upper bound. Thus, we cannot apply the standard technique in order to obtain weak compactness results. Instead, we decompose u^ε into the product:

$$u^\varepsilon(t,x) = \exp\left(\frac{1}{\varepsilon^{1\wedge\alpha/2}} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \,\mathrm{d}s \right) v^\varepsilon(t,x)$$

and introduce in this way new unknown function v^ε; $\langle \cdot \rangle$ stands for the mean value of a periodic function.

For the family of functions v^ε we obtain a priori estimates, prove the convergence of v^ε in probability to a deterministic limit in a suitable functional space and find the auxiliary homogenized equation for the limit function. To this end we introduce a family of correctors involving the solutions of proper Poisson equations and vanishing in a suitable sense as $\varepsilon \to 0$, and, with the help of Itô's calculus, reduce the problem to studying the limiting behavior of a family of semi-martingales. Then, the technique developed in Viot (1976) and Bouc and Pardoux (1984) can be applied.

The description of the limit distribution of

$$\exp\left(\frac{1}{\varepsilon^{1\wedge\alpha/2}} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \,\mathrm{d}s \right)$$

is due to Pardoux and Veretennikov (2001), where the weak convergence of the integrals

$$\frac{1}{\varepsilon^{\alpha/2}} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \,\mathrm{d}s$$

has been proved.

Finally, the passage to the limit in the product relies on the deterministic nature of the limit of v^ε.

The analysis used in this paper is essentially based on the properties of the density on the invariant measure of ξ_t and of a solution of a Poisson equation stated in the product of $\mathbb{R}^d$ and the torus $\mathbb{T}^n$. Although the statements that characterize the said properties are of the same nature as those of Pardoux and Veretennikov (2001) and can be obtained by similar methods, they are not presented in the cited works and thus are to be proved. The appendix deals with a number of technical extensions of the results from Pardoux and Veretennikov (2001).

2. The setup

We consider the asymptotic behavior of the solution of the following Cauchy problem as $\varepsilon \downarrow 0$:

$$u_t^\varepsilon(t,x) = \operatorname{div}\left[a\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right)\nabla u^\varepsilon(t,x)\right] + \frac{1}{\varepsilon^{1\wedge\alpha/2}} c\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) u^\varepsilon(t,x), \tag{1}$$

$$u^\varepsilon(0,x) = u_0(x),\ x\in\mathbb{R}^n,\ t\in[0,T], \tag{2}$$

where $\alpha>0$ is a parameter and $T>0$ is fixed.

The coefficients $a(z,y)$ and $c(z,y)$ are periodic in z (or z belongs to the unit torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$), and $\{\xi_t\}_{t\geqslant 0}$ is stationary ergodic diffusion process, with values in $\mathbb{R}^d$, given by

$$\mathrm{d}\xi_t = B(\xi_t)\,\mathrm{d}t + \sigma(\xi_t)\,\mathrm{d}W_t. \tag{3}$$

Let us introduce the following operators:

- the infinitesimal generator of the diffusion process $\{\xi_t\}$:

$$\mathscr{L}g(y) = \sum_{1\leqslant k,l\leqslant d} q^{kl}(y) g_{y_k y_l}(y) + \sum_{1\leqslant k\leqslant d} B^k(y) g_{y_k}(y) \tag{4}$$

  with $q = \frac{1}{2}\sigma\sigma^*$,
- and

$$\mathscr{A}^\varepsilon h(x) = \operatorname{div}\left(a\left(\frac{x}{\varepsilon}, y\right)\nabla h(x)\right). \tag{5}$$

  $\mathscr{A}$ will denote $\mathscr{A}^\varepsilon$ for $\varepsilon = 1$.

Note that, applied to a function $f(z,y)$, $\mathscr{L}$ acts on the function $y\mapsto f(z,y)$ for z fixed, and $\mathscr{A}^\varepsilon$ acts on the function $z\mapsto f(z,y)$ for y fixed.

2.1. Notations

- In $\mathbb{R}^n$, $x{\cdot}x'$ will denote the scalar product and $|\cdot|$ the corresponding norm.
- In the space $L^2(\mathbb{R}^n)$, $(\cdot\,,\cdot)$ will denote the inner product, and $\|\cdot\|$ the norm.
- For a function $(z,y)\mapsto f(z,y)$, we use the following notations:

$$\overline{\langle f(\cdot\,,\cdot)\rangle} = \int_{\mathbb{T}_n}\int_{\mathbb{R}^d} f(z,y)\rho(y)\,\mathrm{d}y\,\mathrm{d}z,$$

$$\langle f(\cdot\,,y)\rangle = \int_{\mathbb{T}_n} f(z,y)\,\mathrm{d}z,$$

$$\overline{f(z,\cdot)} = \int_{\mathbb{R}^d} f(z,y)\rho(y)\,\mathrm{d}y;$$

here and in what follows $\rho(y)$ stands for the density of invariant measure of process ξ_s. The question of the existence of invariant measure and the properties of $\rho(y)$ will be discussed later.

- For a function or process $(t,x)\mapsto u(t,x)$, $u(t)$ will denote an application $x\mapsto u(t,x)$. Hence $\|u(t)\|$ is $(\int_{\mathbb{R}^n}|u(t,x)|^2\,\mathrm{d}x)^{1/2}$. This notation is also used for $u^\varepsilon(t,x)$ and for the gradient $\nabla u^\varepsilon(t,x)$. We use, as well, the contracted notation:

$$c^\varepsilon = c(\tfrac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}),\quad a^\varepsilon = a(\tfrac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}),\quad a^{\varepsilon,ij} = a^{ij}(\tfrac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}),$$

and for a generic function $g(z,y):\ g^\varepsilon = g(x/\varepsilon,\xi_{t/\varepsilon^\alpha})$.

2.2. Hypotheses

In this section, we provide the precise assumptions on the coefficient of (1) and on the process ξ_t.

Hypothesis 2.1. The coefficients a, c, and q as well as their derivatives are uniformly bounded: there exists $C_1>0$ such that for all $(z,y)\in\mathbb{T}^n\times\mathbb{R}^d$

$$|a^{ij}(z,y)| + |\nabla_z a^{ij}(z,y)| + |\nabla_y a^{ij}(z,y)| \leqslant C_1,$$

$$|c(z,y)| + |\nabla_z c(z,y)| + |\nabla_y c(z,y)| \leqslant C_1,$$

$$|q^{kl}(y)| + |\nabla_y q^{ij}(y)| \leqslant C_1,$$

for all $1\leqslant i,\ j\leqslant n,\ 1\leqslant k,\ l\leqslant d$ [1]. The vector function B as well as its derivatives satisfy polynomial growth condition:

$$|B(y)| + |\nabla_y B(y)| \leqslant C_1(1+|y|)^{\mu_1}$$

for some $\mu_1\geqslant 0$.

Hypothesis 2.2. Operators $\mathscr{L}$ and $\mathscr{A}$ are uniformly elliptic: there exists a constant $C_2>0$ such that for all $(z,y)\in\mathbb{T}^n\times\mathbb{R}^d$

$$C_2|z'|^2 \leqslant (a(z,y)z')\cdot z' \quad \forall z'\in\mathbb{R}^n,$$

$$C_2|y'|^2 \leqslant (q(z,y)y')\cdot y' \quad \forall y'\in\mathbb{R}^d.$$

Hypothesis 2.3. There exist constants $\mu>-1$, $R_1>0$ and $C_3>0$ such that

$$B(y)\cdot\frac{y}{|y|} \leqslant -C_3|y|^\mu \quad \forall y \text{ s.t. } |y|>R_1.$$

Under Hypotheses 2.1, 2.2, and 2.3, process $\{\xi_t\}$ admits the unique invariant measure with smooth density $\rho(y)$ given by

$$\mathscr{L}^*\rho = 0 \text{ on } \mathbb{R}^d \quad\text{and}\quad \int_{\mathbb{R}^d}\rho(y)\,\mathrm{d}y = 1. \tag{6}$$

[1] Notations: ∇_z, ∇_y are the space gradient with respect to z and y, respectively. When there is no ambiguity about the argument of the function we use the notation ∇, for example $\nabla u^\varepsilon(t,x)$ is $\nabla_x u^\varepsilon(t,x)$.

Moreover, the density $\rho(\cdot)$ decays faster than any negative power of $|y|$ as $|y|\to\infty$ (see Lemma A.1 in Appendix). In fact, the following bound holds:

$$\rho(y) \leqslant c_1 \exp(-c|y|^{1+\mu}), \quad c>0.$$

Hypothesis 2.4. We will suppose that $\overline{\langle c(\cdot,\cdot)\rangle}=0$.

This hypothesis is, in fact, not a restriction. Indeed, considering the new unknown function $\tilde{u}^{\varepsilon}(t,x)=\mathrm{e}^{-t\overline{\langle c\rangle}/\varepsilon}u^{\varepsilon}(t,x)$, one can always achieve the relation $\overline{\langle c\rangle}=0$.

2.3. Auxiliary results

Let $\{(\eta_t,\xi_t)\}$ be the diffusion process associated to the infinitesimal generator $\mathscr{A}+\mathscr{L}$, and let $L^2_\rho(\mathbb{T}^n\times\mathbb{R}^d)$ denote the weighted space with the norm:

$$\|f\|^2_\rho = \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} f(z,y)^2\rho(y)\,\mathrm{d}y\,\mathrm{d}z.$$

Also, we introduce the spaces:

$$\bar{L}^2_\rho(\mathbb{T}^n\times\mathbb{R}^d) = \{f\in L^2_\rho(\mathbb{T}^n\times\mathbb{R}^d); \overline{\langle f(\cdot,\cdot)\rangle}=0\},$$

$$\bar{H}^1_\rho(\mathbb{T}^n\times\mathbb{R}^d) = \{f\in \bar{L}^2_\rho(\mathbb{T}^n\times\mathbb{R}^d); |\nabla_x f|+|\nabla_z f|\in L^2_\rho(\mathbb{T}^n\times\mathbb{R}^d)\}.$$

The next statement is a generalization of [16, Theorem 1]. It is proved in Appendix.

Lemma 2.5. *Let $f\in\bar{L}^2_\rho(\mathbb{T}^n\times\mathbb{R}^d)$, and assume that*

$$|f(z,y)| \leqslant C_5(1+|y|^p) \quad \forall (z,y)\in\mathbb{T}^n\times\mathbb{R}^d$$

for some constants $C_5>0$ and $p\in\mathbb{N}$. Then the equation

$$(\mathscr{A}+\mathscr{L})u(z,y)=f(z,y) \tag{7}$$

does have a unique solution $u\in\bar{H}^1_\rho(\mathbb{T}^n\times\mathbb{R}^d)$ and the estimate

$$|u(z,y)| \leqslant C_6(1+|y|^{p_1}) \quad \forall (z,y)\in\mathbb{T}^n\times\mathbb{R}^d$$

holds; moreover, p_1 depends only on p and μ and the constant C_6 depends only on C_5 and p and μ (we assume that the dimensions are fixed).

If, in addition, there exists $N>0$ such that for all n_1, $n_2\in\mathbb{N}$ with $n_1+n_2\leqslant N$ we have

$$|\partial_z^{n_1}\partial_y^{n_2} f(z,y)| \leqslant C_5(1+|y|^p) \quad \forall (z,y)\in\mathbb{T}^n\times\mathbb{R}^d$$

then

$$|\partial_z^{n_1}\partial_y^{n_2} u(z,y)| \leqslant C_6(1+|y|^{p_1}) \quad \forall (z,y)\in\mathbb{T}^n\times\mathbb{R}^d.$$

Applying the technique developed by Pardoux and Veretennikov (2001, Proposition 2), leads to the following statement (proved in the Appendix):

Proposition 2.6. *For any fixed $T>0$, $p>0$, $\alpha>0$, and $\beta>0$:*

$$\lim_{\varepsilon\downarrow 0} E\left(\sup_{t\leqslant T}\varepsilon^\beta|\xi_{t/\varepsilon^\alpha}|^p\right)=0.$$

3. Main results

Here we formulate the main results of the paper; the proof will be given in the following section.

It should be noted that for $\alpha \leqslant 2$, we obtain the weak convergence of the law of $u^\varepsilon(t,x)$ to the nontrivial limit law which solves a proper martingale problem, while for $\alpha > 2$, the limit law is a Dirac measure concentrated on the solution of the Cauchy problem for the limit deterministic parabolic equation with constant coefficients.

Let $L^2_w(\mathbb{R}^n)$ denote the space $L^2(\mathbb{R}^n)$ endowed with the weak topology. Define

$$\Omega_T = L^2_w((0,T); H^1(\mathbb{R}^n)) \cap \mathscr{C}([0,T]; L^2_w(\mathbb{R}^n)) \tag{8}$$

endowed with supremum of the topology of uniform convergence over the space $\mathscr{C}([0,T]; L^2_w(\mathbb{R}^n))$ and weak topology over the space $L^2((0,T); H^1(\mathbb{R}^n))$. Ω_T is a Lusin and regular space; denote by $\mathscr{F}$ its Borel σ-field.

For any $\varepsilon > 0$, let Q^ε be the Radon probability measure on $(\Omega_T, \mathscr{F})$ which is defined by the law of $\{u^\varepsilon(t);\ 0 \leqslant t \leqslant T\}$. The asymptotic behavior of u^ε, as $\varepsilon \downarrow 0$, depends on whether $\alpha < 2$, $\alpha = 2$, or $\alpha > 2$.

Theorem 3.1. *Let* $\alpha < 2$, *then under Hypotheses* 2.1–2.4, *the law* Q^ε *of the solution* u^ε *of Eqs.* (1)–(2) *converges, as* $\varepsilon \downarrow 0$, *in space* Ω_T *to the law* $\hat{Q}$ *of the solution* $\hat{u}$ *of the* SPDE

$$\mathrm{d}\hat{u}(t,x) = [\operatorname{div}(\hat{a}\nabla\hat{u}(t,x)) + \hat{c}\hat{u}(t,x)]\,\mathrm{d}t + \lambda\hat{u}(t,x)\,\mathrm{d}\hat{W}_t \tag{9}$$

with initial condition $\hat{u}(0,x) = u_0(x)$, *where* $(t,x) \in [0,T] \times \mathbb{R}^n$, $\{\hat{W}_t\}$ *is standard Brownian motion in* $\mathbb{R}$ *and*

$$\hat{a} = \overline{\langle a(I + \nabla_z \Psi)\rangle},$$

$$\hat{c} = \overline{\langle Gc\rangle} \text{ which is also equal to } \overline{q\nabla G\cdot\nabla G},$$

$$\lambda^2 = 2\overline{q\nabla G\cdot\nabla G},$$

and the functions $G \in \bar{H}^1_\rho(\mathbb{R}^d)$ *and* $\Psi^i \in \bar{H}^1_\rho(\mathbb{T}^n \times \mathbb{R}^d)$ *are the solutions of the equations*:

$$\mathscr{L}G(y) = -\langle c(\cdot\,, y)\rangle, \tag{10}$$

$$\mathscr{A}\Psi^i(z,y) = -\sum_{j=1}^{n} a^{ij}_{z_j}(z,y), \tag{11}$$

for $(z,y) \in \mathbb{T}^n \times \mathbb{R}^d$ *and* $1 \leqslant i \leqslant n$.

Theorem 3.2. *For* $\alpha = 2$, *under Hypotheses* 2.1–2.4, *the law* Q^ε *of the solution* u^ε *of Eqs.* (1)–(2) *converges, as* $\varepsilon \downarrow 0$, *in space* Ω_T *to the law* $\hat{Q}$ *of the solution* $\hat{u}$ *of the* SPDE

$$\mathrm{d}\hat{u}(t,x) = [\operatorname{div}(\hat{a}\nabla\hat{u}(t,x)) - \hat{b}\cdot\nabla\hat{u}(t,x) + \hat{c}\hat{u}(t,x)]\,\mathrm{d}t + \lambda\hat{u}(t,x)\,\mathrm{d}\hat{W}_t, \tag{12}$$

with initial condition $\hat{u}(0,x)=u_0(x)$, *where* $(t,x)\in[0,T]\times\mathbb{R}^n$, $\{\hat{W}_t\}$ *is a standard Brownian motion in* $\mathbb{R}$, *and*

$$\hat{a}=\overline{\langle a(I+\nabla_z\Psi)\rangle},$$

$$\hat{b}=\overline{\langle \Psi c+a\nabla_z G\rangle},$$

$$\hat{c}=\overline{\langle Gc\rangle},$$

$$\lambda^2=2\overline{q\langle\nabla_y G\rangle\cdot\langle\nabla_y G\rangle},$$

and the functions G, $\Psi^j\in\bar{H}^1_\rho(\mathbb{T}^n\times\mathbb{R}^d)$ *are the solutions of the equations*

$$(\mathscr{A}+\mathscr{L})G(z,y)=-c(z,y), \tag{13}$$

$$(\mathscr{A}+\mathscr{L})\Psi^j(z,y)=-\sum_{i=1}^{n}a^{ij}_{z_i}(z,y) \tag{14}$$

for $(z,y)\in\mathbb{T}^n\times\mathbb{R}^d$ *and* $1\leqslant j\leqslant n$.

Theorem 3.3. *For* $\alpha>2$, *under Hypotheses* 2.1–2.4, *the solution* u^ε *of Eqs.* (1)–(2) *converges in probability in the space* Ω_T *to the solution* $\hat{u}$ *of the following limit Cauchy problem*:

$$\hat{u}_t(t,x)=\operatorname{div}(\hat{a}\nabla\hat{u}(t,x))+\hat{c}\hat{u}(t,x),\quad \hat{u}(0,x)=u_0(x) \tag{15}$$

with $(t,x)\in[0,T]\times\mathbb{R}^n$ *and*

$$\hat{a}=\langle\bar{a}(I+\nabla_z\Psi)\rangle,\quad \hat{c}=\langle G\bar{c}\rangle,$$

where the functions G, $\Psi^i\in\bar{H}^1(\mathbb{T}^n)$ *are solutions of equations*

$$\bar{\mathscr{A}}G(z)=-\overline{c(z,\cdot)}, \tag{16}$$

$$\bar{\mathscr{A}}\Psi^i(z)=-\sum_{j=1}^{n}\overline{a^{ij}_{z_j}(z,\cdot)} \tag{17}$$

for $z\in\mathbb{T}^n$ *and* $1\leqslant i\leqslant n$, *where the operator* $\bar{\mathscr{A}}$ *is defined by*

$$\bar{\mathscr{A}}f(z)=\operatorname{div}(\overline{a(z,\cdot)}\nabla f(z)). \tag{18}$$

From which we deduce the following

Corollary 3.4. *For* $\alpha>2$, *under Hypotheses* 2.1–2.4, *we have*

$$P\text{-}\lim_{\varepsilon\downarrow 0}\|u^\varepsilon-\hat{u}\|_{L^2((0,T)\times\mathbb{R}^n)}=0 \tag{19}$$

where u^ε (*resp.* $\hat{u}$) *is the solution of Eqs.* (1)–(2) (*resp.* (15)).

Comparison with systems without potential or without "*noise input*": It is interesting to compare the limit problems (9), (12), (15) above with the limit problems for the equation without potential:

$$u_t(t,x)=\operatorname{div}\left(a\left(\frac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}\right)\nabla u(t,x)\right) \tag{20}$$

and the equation “without noise input”:

$$u_t(t,x) = \operatorname{div}\left(a\left(\frac{x}{\varepsilon}\right)\nabla u(t,x)\right) + \frac{1}{\varepsilon}c\left(\frac{x}{\varepsilon}\right)u(t,x). \tag{21}$$

According to Kleptsyna and Piatnitski (1997), the absence of the potential in (20) always leads to the deterministic form of homogenized problem:

$$\overline{u}_t(t,x) = \sum_{1 \leqslant i,j \leqslant n} \bar{a}^{ij}\frac{\partial^2}{\partial x_i \partial x_j}\overline{u}(t,x),$$

this operator involves neither stochastic nor lower-order terms. The limit problem for (21) takes the following form:

$$\overline{u}_t(t,x) = \sum_{1 \leqslant i,j \leqslant n} \bar{a}^{ij}\frac{\partial^2}{\partial x_i \partial x_j}\overline{u}(t,x) + \bar{c}\,\overline{u}(t,x)$$

where $\bar{c}$ is the so-called “strange term”.

Comment on the limiting Eq. (12): The appearance of the first-order term $\hat{b}\cdot\nabla$ in the drift part of the limit problem (12) is of special interest. It should be shown that this first-order term is not necessarily null.

To this end, let us propose an example with $n=1$ and $d=2$. We take $q^{kl}=\delta_{kl}$ in (4) and choose

$$B(y) = -2\begin{pmatrix} y_1 + \nu y_2 \\ y_2 - \nu y_1 \end{pmatrix}$$

where $\nu > 0$. One can easily check that the density $\rho(y)$ of the invariant measure of ξ_t is given, under this choice of q^{kl} and $B(y)$, by the formula

$$\rho(y) = c\mathrm{e}^{-|y|^2},$$

here c is a normalization constant. We consider the one-dimensional case (w.r.t. z):

$$\hat{b} = \int_{\mathbb{T}^1}\int_{\mathbb{R}^2}[\Psi(z,y)c(z,y) + a(z,y)G_z(z,y)]\rho(y)\,\mathrm{d}z\,\mathrm{d}y.$$

Integrating by parts and taking into account Eqs. (13) and (14), we obtain

$$\hat{b} = \int_{\mathbb{T}^1}\int_{\mathbb{R}^2}(\Psi\mathscr{L}G - G\mathscr{L}\Psi)(z,y)\rho(y)\,\mathrm{d}z\,\mathrm{d}y. \tag{22}$$

here we also used the divergent form and the symmetry of the operator $\mathscr{A}$.

Finally, in view of (6), after integrating by parts, we get

$$\begin{aligned}\hat{b} &= -\int_{\mathbb{T}^1}\int_{\mathbb{R}^2}[(\nabla_y\Psi)\cdot G(\nabla_y\rho - B\rho)](z,y)\,\mathrm{d}z\,\mathrm{d}y \\ &= -\nu\int_{\mathbb{T}^1}\int_{\mathbb{R}^2}(\nabla_y\Psi)(z,y)\cdot G(z,y)\begin{pmatrix} y_2 \\ y_1\end{pmatrix}\mathrm{d}z\,\mathrm{d}y.\end{aligned} \tag{23}$$

Denote $k(y) = (\mathscr{L}-1)\sin(y_1)$ and construct the operator $\mathscr{A}$ to be a small perturbation of operator with constant coefficient:

$$a(z,y) = 1 + \aleph d(z,y) \tag{24}$$

where $\aleph$ is a small parameter and $d(z,y)=\sin(z)k(y)$. It is easy to see that the function $\Psi^1(z,y)=\cos(z)\sin(y_1)$ solves the following auxiliary problem:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}z^2}+\mathscr{L}\right)\Psi^1(z,y)=-\cos(z)k(y).$$

Now, substituting the function $\aleph\Psi^1(z,y)$ in Eq. (14) which in our particular case reads

$$\left(\frac{\mathrm{d}}{\mathrm{d}z}a(z,y)\frac{\mathrm{d}}{\mathrm{d}z}+\mathscr{L}\right)\Psi(z,y)=-\aleph\frac{\mathrm{d}}{\mathrm{d}z}d(z,y)=-\aleph\cos(z)k(y)$$

and taking into account the coerciveness of the latter problem, we get

$$\|\Psi-\aleph\Psi^1\|=\mathrm{O}(\aleph^2)\quad\text{as } \aleph\downarrow 0.$$

Thus, it suffices to show that the integral

$$\int_{\mathbb{T}^1}\int_{\mathbb{R}^2}(\nabla_y\Psi^1)(z,y)\cdot G(z,y)\begin{pmatrix}y_2\\ y_1\end{pmatrix}\mathrm{e}^{-|y|^2}\,\mathrm{d}z\,\mathrm{d}y$$

$$=\int_{\mathbb{T}^1}\int_{\mathbb{R}^2}\cos(z)y_2\cos(y_1)G(z,y)\mathrm{e}^{-|y|^2}\,\mathrm{d}z\,\mathrm{d}y$$

is not equal to zero under the proper choice of $G(z,y)$.

To do so one can construct a smooth nonnegative function $G(z,y)$ with a small compact support concentrated near the point (z^0,y^0) such that

$$y_2^0\cos(z^0)\cos(y_1^0)\mathrm{e}^{-|y^0|^2}\neq 0,$$

say $(z^0,y^0)=(\pi//4,(\pi//4,\pi//4))$.

4. Proof of the main results

4.1. Decomposition of the solution u^ε: Auxiliary homogenization problem

The expectation of the L^2 or $L^2((0,T);H^1(\mathbb{R}^n))$-norms of u^ε in general does not admit uniform in ε estimates. Thus, in order to obtain the compactness of u^ε we need an approach that does not rely on estimates of the expectation.

In this section we decompose u^ε into the product of the exponent of given function, and of new unknown function $v^\varepsilon(x,t)$:

$$u^\varepsilon(t,x)=\exp\left(\frac{1}{\varepsilon^{1\wedge\alpha/2}}\int_0^t\langle c(\cdot,\xi_{s/\varepsilon^\alpha})\rangle\,\mathrm{d}s\right)v^\varepsilon(t,x). \tag{25}$$

Direct calculations show that the function v^ε satisfies the following equation:

$$v_t^\varepsilon(t,x)=\mathrm{div}\left[a\left(\frac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}\right)\nabla v^\varepsilon(t,x)\right]+\frac{1}{\varepsilon^{1\wedge\alpha/2}}\tilde{c}\left(\frac{x}{\varepsilon},\xi_{t/\varepsilon^\alpha}\right)v^\varepsilon(t,x), \tag{26}$$

$$v^\varepsilon(0,x)=u_0(x),\quad x\in\mathbb{R}^n,\ \ t\in[0,T], \tag{27}$$

where $\tilde{c}(z,y)=c(z,y)-\langle c(\cdot,y)\rangle$.

In the following sections we show that v^ε converges in probability in the space Ω_T to nonrandom function being the solution of Cauchy problem with constant coefficient.

4.2. Tightness result for auxiliary homogenization problem

This section deals with the tightness of the family of distributions associated with solutions v^ε of Problem (26)–(27). We start by obtaining uniform estimates for v^ε.

Proposition 4.1. *The inequalities*
(E1) $\sup_{t\leqslant T}\|v^\varepsilon(t)\|^2\leqslant c$,
(E2) $\int_0^T\|\nabla v^\varepsilon(s)\|^2\,\mathrm{d}s\leqslant c$, *hold uniformly in* $\varepsilon>0$. *For any test function* $\varphi\in\mathscr{C}_0^\infty(\mathbb{R}^n)$ *the inequality*
(E3) $|(\varphi, v^\varepsilon(t)-v^\varepsilon(s))|\leqslant c|t-s|^{1/2}$ *holds uniformly in* $\varepsilon>0$.

Proof. Multiplying Eq. (26) by v^ε and integrating by parts we get

$$\begin{aligned}\tfrac12\|v^\varepsilon\|_t^2 &= -(a^\varepsilon\nabla v^\varepsilon,\nabla v^\varepsilon)+\varepsilon^{-(1\wedge\alpha/2)}(\tilde c^\varepsilon v^\varepsilon, v^\varepsilon)\\ &= -(a^\varepsilon\nabla v^\varepsilon,\nabla v^\varepsilon)+2\varepsilon^\gamma(\kappa^\varepsilon v^\varepsilon,\nabla v^\varepsilon),\end{aligned}$$

where $\gamma=1-(1\wedge\alpha/2)$ and $\kappa(z,y)$ is z-periodic vector function defined by the relation

$$\mathrm{div}_z\,\kappa(z,y)=\tilde c(z,y). \tag{28}$$

In order to construct such a function $\kappa(z,y)$, one considers the equation

$$\Delta_z K(z,y)=\tilde c(z,y),\quad z\in\mathbb{T}^n.$$

Thanks to the relation $\langle\tilde c(\cdot,y)\rangle=0$ this equation is solvable for any $y\in\mathbb{R}^d$. Moreover, since $\tilde c\in\mathscr{C}_b^1(\mathbb{T}^n\times\mathbb{R}^d)$, all the derivatives of $K(z,y)$ in z up to the third order are uniformly bounded. By putting $\kappa(z,y)=\nabla_z K(z,y)$ we obtain the desired representation.

With the help of the Cauchy–Buniakovski inequality after simple transformation we obtain

$$\begin{aligned}\frac12\|v^\varepsilon\|^2 &\leqslant \frac12\|u_0\|^2-\int_0^t(a^\varepsilon\nabla v^\varepsilon,\nabla v^\varepsilon)\,\mathrm{d}s+\delta\int_0^t\|\nabla v^\varepsilon\|^2\,\mathrm{d}s+\frac{c}{\delta}\int_0^t\|v^\varepsilon\|^2\,\mathrm{d}s\\ &\leqslant \frac12\|u_0\|^2-\Lambda\int_0^t\|\nabla v^\varepsilon\|^2\,\mathrm{d}s+c_1\int_0^t\|v^\varepsilon\|^2\,\mathrm{d}s,\end{aligned}$$

for some $\Lambda>0$. Now (E1) and (E2) follow from Gronwall's lemma.

For any test function φ we have

$$\begin{aligned}|(\varphi, v^\varepsilon(t)-v^\varepsilon(s))| &\leqslant \int_s^t(\nabla\varphi, a^\varepsilon\nabla v^\varepsilon)\,\mathrm{d}r+\varepsilon^\gamma\int_s^t|(\kappa^\varepsilon,\nabla(\varphi v^\varepsilon)|\,\mathrm{d}r\\ &\leqslant c|t-s|^{1/2}.\end{aligned}$$

This completes the proof. □

For any $\varepsilon>0$ denote by $\tilde Q^\varepsilon$ a Radon probability measure on $(\Omega_T,\mathscr{F})$ defined as the law of $\{v^\varepsilon(t); 0\leqslant t\leqslant T\}$ in Ω_T. According to Viot (1976) and Bouc and Pardoux 1984, Theorem 2.5), Proposition 4.1 implies the tightness of the family $\{\tilde Q^\varepsilon\}$ in Ω_T.

4.3. Description of homogenized operator for auxiliary problem – convergence in probability

In this section we prove the convergence of v^ε in probability and show that the limit function satisfies parabolic equation with constant coefficients. We study the cases $\alpha>2$, $\alpha=2$ and $\alpha<2$ separately.

Case $\alpha<2$: Let $\hat v(t,x)$ be the solution of the following Cauchy problem:

$$\hat v_t = \mathrm{div}(\hat a \nabla \hat v), \;\; \hat v(x,0) = u_0(x), \tag{29}$$

$\hat a$ has been defined in Theorem 3.1.

Proposition 4.2. *For any test function* $\varphi(x,t)\in C^\infty([0,T];C_0^\infty(R^n))$ *the following limit relation holds*:

$$\lim_{\varepsilon\downarrow 0} E \sup_{t\leqslant T}\left| (\varphi(t),v^\varepsilon(t)) - (\varphi(0),u_0) - \int_0^t (\varphi_s(s),v^\varepsilon(s))\,\mathrm{d}s \right.$$
$$\left. - \int_0^t \left(\sum_{ij} \hat a^{ij}\varphi_{x_i x_j}(s), v^\varepsilon(s)\right)\mathrm{d}s\right| = 0. \tag{30}$$

Proof. Let us introduce a process:

$$H^\varepsilon(t) = (\varphi, v^\varepsilon(t)) + \varepsilon \sum_{k=1}^n (\Psi^k(\cdot/\varepsilon, \xi_{t/\varepsilon^\alpha})\varphi_{x_k}, v^\varepsilon(t)), \tag{31}$$

where the z-periodic functions $\Psi^k \in \bar H^1_\rho(\mathbb{T}^n\times\mathbb{R}^n)$ are defined as the solutions of the following equations:

$$\mathscr{A}\Psi^k(z,y) = -\sum_j a^{kj}_{z_j}(z,y), \quad k=1,\dots,n. \tag{32}$$

From Proposition 4.1, (E1), (E2), Lemma 2.5 and Proposition 2.6 we have

$$\lim_{\varepsilon\to 0} E \sup_{t\leqslant T} |(\varphi, v^\varepsilon) - H^\varepsilon(t)| = 0. \tag{33}$$

Indeed, according to Hypothesis 2.1 the function on the right-hand side of (32) is uniformly bounded in $|y|$. Therefore, by Lemma 2.5 the solution Ψ^k admits a polynomial estimate, and from (E1) and Proposition 2.6 we have

$$\begin{aligned} E \sup_{t\leqslant T} |(\varphi, v^\varepsilon) - H^\varepsilon(t)| &= \varepsilon E \sup_{t\leqslant T} |(\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon(t))| \\ &\leqslant C\varepsilon E \sup_{t\leqslant T} |\Psi^{\varepsilon,k}|\,\|v^\varepsilon\| \\ &\leqslant C\varepsilon E \sup_{t\leqslant T} |\xi_{t/\varepsilon^\alpha}|^m \underset{\varepsilon\downarrow 0}{\longrightarrow} 0. \end{aligned}$$

Now, in order to prove (30), we assume for a while that $a(x,y)$ is three times differentiable in x and y, and that all its derivatives up to the third order admit polynomial estimates as $|y|\uparrow\infty$.

Then, clearly, we have for some $m \geqslant 0$:

$$|\Psi^k(y,z)| + |\nabla_z \Psi^k(y,z)| + |\nabla_y \Psi^k(y,z)| + |\nabla_{yy} \Psi^k(y,z)| \leqslant C(1+|y|)^m. \tag{34}$$

Applying Ito's formula to the process $H^\varepsilon(t)$ and integrating by parts one has

$$\begin{aligned}
\mathrm{d}H^\varepsilon(t) = {} & \sum_{ij} (a^{\varepsilon,ij}\varphi_{x_i x_j}, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{-1}\sum_{ij}(a^{\varepsilon,ij}_{z_j}\varphi_{x_i}, v^\varepsilon)\,\mathrm{d}t \\
& + \varepsilon^{-\alpha/2}(\tilde{c}^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + (\varphi_t, v^\varepsilon)\,\mathrm{d}t \\
& + \sum_k \Bigg[\varepsilon^{1-\alpha}(\mathscr{L}\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{1-\alpha/2}(\nabla_y \Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon)\cdot\sigma^\varepsilon\,\mathrm{d}W_t \\
& \quad + \varepsilon^{-1}((\mathscr{A}\Psi^{\varepsilon,k})\varphi_{x_k}, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{1-\alpha/2}(\Psi^{\varepsilon,k}\tilde{c}^\varepsilon\varphi_{x_k}, v^\varepsilon)\,\mathrm{d}t \\
& \quad + \sum_{ij}((a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_i})\varphi_{x_k x_j}, v^\varepsilon)\,\mathrm{d}t \\
& \quad - \varepsilon\sum_{ij}(a^{\varepsilon,ij}\Psi^{\varepsilon,k}\varphi_{x_k x_i}, v^\varepsilon_{x_j})\,\mathrm{d}t\Bigg].
\end{aligned}$$

The functions Ψ^k satisfy the relation $\langle \Psi^k(\cdot\,,y)\rangle = 0$, thus

$$\langle(\mathscr{L}\Psi^k)(\cdot\,,y)\rangle = \mathscr{L}\langle\Psi^k(\cdot\,,y)\rangle = 0,$$

and in the same way as in (28) we have

$$\mathscr{L}\Psi^k(z,y) = \mathrm{div}_z\,\kappa^k(z,y) \tag{35}$$

with continuous $\kappa^k(z,y)$ of polynomial growth in $|y|$.

Taking into account (32) after simple transformation we get

$$\begin{aligned}
\mathrm{d}H^\varepsilon = {} & \sum_{ij}(a^{\varepsilon,ij}\varphi_{x_i x_j}, v^\varepsilon)\,\mathrm{d}t + \sum_k\sum_{ij}((a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j})\varphi_{x_k x_i}, v^\varepsilon)\,\mathrm{d}t + (\varphi_t, v^\varepsilon)\,\mathrm{d}t \\
& + \varepsilon^{\delta/2}(\kappa^\varepsilon, \nabla(\varphi v^\varepsilon))\,\mathrm{d}t + \varepsilon^\delta(\kappa^{\varepsilon,k}, \nabla(\varphi_{x_k} v^\varepsilon))\,\mathrm{d}t + \varepsilon(\Psi^{\varepsilon,k}\varphi_{t x_k}, v^\varepsilon)\,\mathrm{d}t \\
& + \varepsilon\sum_k\sum_{ij}(a^{\varepsilon,ij}\Psi^{\varepsilon,k}\varphi_{x_k x_i}, v^\varepsilon_{x_j})\,\mathrm{d}t + \varepsilon^{\delta/2}(\tilde{c}\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon)\,\mathrm{d}t \\
& + \varepsilon^{\delta/2}(\nabla_y\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon)\sigma^\varepsilon\,\mathrm{d}W_t \\
= {} & \sum_{ij}(a^{\varepsilon,ij}\varphi_{x_i x_j}, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{\delta/2}(\kappa^\varepsilon, \nabla(\varphi v^\varepsilon))\,\mathrm{d}t \\
& + \sum_k\sum_{ij}((a^{\varepsilon,ij}\Psi^{\varepsilon,k})_{z_i}\varphi_{x_k x_j} + a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j}\varphi_{x_k x_i}, v^\varepsilon)\,\mathrm{d}t \\
& + (\varphi_t, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{\delta/2}\,\mathrm{d}R^\varepsilon(t)
\end{aligned} \tag{36}$$

where $\delta = 2 - \alpha$, and $\lim_{\varepsilon\downarrow 0} E\sup_{t\leqslant T}|R^\varepsilon(t)| \leqslant C$. This limit relation follows immediately from (E1), (E2) and Burkholder–Davis–Gundy inequality. Therefore,

$$(\varphi, v^\varepsilon) - (\varphi, u_0) - \int_0^t (\varphi_s, v^\varepsilon)\,\mathrm{d}s - \int_0^t \sum_{ij} (\hat{a}^{ij}\varphi_{x_ix_j}, v^\varepsilon)\,\mathrm{d}s$$

$$= \int_0^t \left[\sum_{ij} (a^{\varepsilon,ij}\varphi_{x_ix_j}, v^\varepsilon) + \sum_k \sum_{ij} (a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j}\varphi_{x_kx_i}, v^\varepsilon) \right.$$

$$\left. - \sum_{ij} (\hat{a}^{ij}\varphi_{x_ix_j}, v^\varepsilon) \right] \mathrm{d}s + \varepsilon^{\delta/2} R^\varepsilon(t);$$

where we have also added and subtracted the term $\sum_{ij}(\hat{a}^{ij}\varphi_{x_ix_j}, v^\varepsilon)\,\mathrm{d}t$.

It remains to show that

$$P\text{-}\lim_{\varepsilon\downarrow 0} \sup_{t\leqslant T} \left| \int_0^t \left(\sum_{ij} a^{\varepsilon,ij}\varphi_{x_ix_j} + \sum_k\sum_{ij} a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j}\varphi_{x_ix_k} - \sum_{ij}\hat{a}^{ij}\varphi_{x_ix_j}, v^\varepsilon \right) \mathrm{d}s \right| = 0$$

which is the aim of the following Lemma 4.3.

Lemma 4.3. *For any z-periodic function* $(z, y) \mapsto h(z, y)$ *such that* $\overline{\langle h\rangle} = 0$ *and* $|h| \leqslant c(1 + |y|)^\theta$ *for some* $\theta \geqslant 0$, *we have*

$$P\text{-}\lim_{\varepsilon\downarrow 0} \sup_{t\leqslant T} \left| \int_0^t \left(h\left(\frac{\cdot}{\varepsilon}, \xi_{s/\varepsilon^\alpha}\right)\varphi, v^\varepsilon(s) \right) \mathrm{d}s \right| = 0$$

for any test function φ.

Proof. Denote by H a z-periodic vector function $(z, y) \mapsto H(z, y)$ given by

$$\mathrm{div}_z H(z, y) = h(z, y) - \langle h(\cdot, y)\rangle, \quad |H| \leqslant c(1 + |y|)^\theta;$$

for instance, one can put $H(z, y) = \nabla_z \hat{H}(z, y)$ where z-periodic $\hat{H}(z, y)$ solves the equation $\Delta_z \hat{H}(z, y) = h(z, y) - \langle h(\cdot, y)\rangle$.

In view of the identity

$$\int_0^t \left(h\left(\frac{\cdot}{\varepsilon}, \xi_{s/\varepsilon^\alpha}\right)\varphi, v^\varepsilon(s) \right) \mathrm{d}s = \varepsilon \int_0^t \left(H\left(\frac{\cdot}{\varepsilon}, \xi_{s/\varepsilon^\alpha}\right), \nabla(\varphi v^\varepsilon(s)) \right) \mathrm{d}s$$

$$+ \int_0^t \langle h(\cdot, \xi_{s/\varepsilon^\alpha})\rangle(\varphi, v^\varepsilon(s))\,\mathrm{d}s$$

and the estimates

$$E\sup_{t\leqslant T} \left| \int_0^t \sum_i \left(H^i\left(\frac{\cdot}{\varepsilon}, \xi_{s/\varepsilon^\alpha}\right), (\varphi v^\varepsilon(s))_{x_i} \right) \mathrm{d}s \right| \leqslant CE(\|v^\varepsilon\|^2_{L^2((0,T);H^1(\mathbb{R}^n))})^{1/2} \leqslant C_1$$

it is sufficient to obtain the following limit relation:

$$P\text{-}\lim_{\varepsilon\downarrow 0} \sup_{t\leqslant T} \left| \int_0^t \langle h(\cdot, \xi_{s/\varepsilon^\alpha})\rangle(\varphi, v^\varepsilon(s))\,\mathrm{d}s \right| = 0.$$

Note that the family of distributions of the processes (φ, v^ε) is relatively compact in $\mathscr{C}[0,T]$. That is, for any $\gamma>0$, there exist $N>0$ and $f^1,\dots,f^N\in\mathscr{C}([0,T];\mathbb{R}^n)$ such that $P(\mathscr{A}_\gamma)<\gamma$ for all $\varepsilon>0$ with

$$\mathscr{A}_\gamma=\bigcap_{k=1}^{N}\left\{\sup_{0\leqslant t\leqslant T}|(\varphi,v^\varepsilon(t))-f^k(t)|>\gamma\right\}.$$

Hence $\mathscr{A}_\gamma^c$ is of the form $\bigcup_k \mathscr{B}_\gamma^k$, $\mathscr{B}_\gamma^k=\{\sup_{0\leqslant t\leqslant T}|(\varphi,u^\varepsilon(t))-f^k(t)|\leqslant\gamma\}$, so $\mathscr{A}_\gamma^c=\bigcup_k\tilde{\mathscr{B}}_\gamma^k$, where $\tilde{\mathscr{B}}_\gamma^k\subset\mathscr{B}_\gamma^k$ and $\tilde{\mathscr{B}}_\gamma^k\cap\tilde{\mathscr{B}}_\gamma^l=\emptyset$ for $k\neq l$. Then,

$$\begin{aligned}
&E\sup_{t\leqslant T}\left|\int_0^t\langle h(\cdot,\xi_{s/\varepsilon^\alpha})\rangle(\varphi,v^\varepsilon)\,\mathrm{d}s\right|\\
&\quad=E\left[\mathbf{1}_{\mathscr{A}_\gamma}\sup_{t\leqslant T}\left|\int_0^t\langle h(\cdot,\xi_{s/\varepsilon^\alpha})\rangle(\varphi,v^\varepsilon)\,\mathrm{d}s\right|\right]\\
&\qquad+E\left[\mathbf{1}_{\mathscr{A}_\gamma^c}\sup_{t\leqslant T}\left|\int_0^t\langle h(\cdot,\xi_{s/\varepsilon^\alpha})\rangle(\varphi,v^\varepsilon)\,\mathrm{d}s\right|\right]\\
&\quad\leqslant C\sqrt{\gamma}+\sum_k E\left[\mathbf{1}_{\tilde{\mathscr{B}}_\gamma^k}\sup_{t\leqslant T}\left|\int_0^t\langle h(\cdot,\xi_{s/\varepsilon^\alpha})\rangle(\varphi,v^\varepsilon)\,\mathrm{d}s\right|\right]\\
&\quad\leqslant C_1\sqrt{\gamma}+\sum_k E\left[\sup_{t\leqslant T}\left|\int_0^t\langle h(\cdot,\xi_{s/\varepsilon^\alpha})\rangle f^k(s)\,\mathrm{d}s\right|\right]
\end{aligned}$$

and the required statement follows from the averaging principle of Liptser and Shiryaev (1989, Theorem 9.6.1) if we pass to the limit on the right-hand side first as $\varepsilon\downarrow 0$ and then as $\gamma\downarrow 0$.

This completes the proof of Lemma 4.3. □

Proof of Proposition 4.2 (Conclusion)**.** Lemma 4.3 proves Proposition 4.2 in the smooth case.

For a general matrix-valued function $a(z,y)$ satisfying Hypotheses 2.1 and 2.2, Relation (30) can be achieved by approximation of $a(z,y)$ with smooth functions. Denote

$$a(z,y,\theta)=\frac{1}{\theta^{n+d}}\int_{\mathbb{T}^n}\int_{\mathbb{R}^d}\varphi_0\left(\frac{(z,y)-(z',y')}{\theta}\right)a(z',y')\,\mathrm{d}z'\,\mathrm{d}y',\ \ \theta>0$$

where $\varphi_0\in\mathscr{C}_0^\infty(\mathbb{T}^n\times\mathbb{R}^d)$, $\varphi_0\geqslant 0$, $\iint\varphi_0(z,y)\,\mathrm{d}z\,\mathrm{d}y=1$. From Hypothesis 2.1 we obtain

$$|a(z,y,\theta)-a(z,y)|\leqslant C\theta,$$

$$|a(z,y,\theta)|+|\nabla_y a(z,y,\theta)|+|\nabla_z a(z,y,\theta)|\leqslant C. \tag{37}$$

Higher-order derivatives of $a(z,y,\theta)$ show polynomial growth in $|y|$ although the corresponding estimates are not uniform in $\theta>0$. For each fixed $\theta>0$, Relation (30) holds

$$\lim_{\varepsilon\downarrow 0} E \sup_{0\leqslant t\leqslant T}\left|(\varphi(t),v^\varepsilon(t,\theta))-(\varphi(0),u_0)-\int_0^t(\varphi_s(s),v^\varepsilon(s,\theta))\,\mathrm{d}s\right.$$

$$\left.-\int_0^t\left(\sum_{ij}\hat a^{ij}(\theta)\varphi_{x_ix_j}(s),v^\varepsilon(s,\theta)\right)\mathrm{d}s\right|=0. \tag{38}$$

The difference $v^\varepsilon - v^\varepsilon(\theta)$ satisfies the following equation

$$(v^\varepsilon - v^\varepsilon(\theta))_t = \sum_{ij}\frac{\partial}{\partial x_i}\left(a^{ij}(x/\varepsilon,\xi_{t/\varepsilon^\alpha})\frac{\partial}{\partial x_j}(v^\varepsilon - v^\varepsilon(\theta))\right)$$

$$+\sum_{ij}\frac{\partial}{\partial x_i}\left([a^{ij}(x/\varepsilon,\xi_{t/\varepsilon^\alpha})-a^{ij}(x/\varepsilon,\xi_{t/\varepsilon^\alpha},\theta)]\frac{\partial}{\partial x_j}v^\varepsilon(\theta)\right), \tag{39}$$

$$(v^\varepsilon - v^\varepsilon(\theta))|_{t=0}=0.$$

Multiplying the latter equation by $v^\varepsilon - v^\varepsilon(\theta)$, we have after integrating by parts:

$$\int_{\mathbb{R}^n}((v^\varepsilon(t)-v^\varepsilon(t,\theta))^2$$

$$+\int_0^t\int_{\mathbb{R}^n}\sum_{ij}a^{\varepsilon,ij}\left(\frac{\partial}{\partial x_i}(v^\varepsilon(s)-v^\varepsilon(s,\theta))\right)\left(\frac{\partial}{\partial x_j}(v^\varepsilon(s)-v^\varepsilon(s,\theta))\right)\mathrm{d}x\,\mathrm{d}s$$

$$=-\int_0^t\int_{\mathbb{R}^n}\sum_{ij}(a^{\varepsilon,ij}-a^{\varepsilon,ij}(\theta))\left(\frac{\partial}{\partial x_i}v^\varepsilon(s,\theta)\right)\left(\frac{\partial}{\partial x_j}(v^\varepsilon(s)-v^\varepsilon(s,\theta))\right)\mathrm{d}x\,\mathrm{d}s$$

$$\leqslant C\theta\|v^\varepsilon(\theta)\|_{L^2(0,T;H^1(\mathbb{R}^n))}\|v^\varepsilon-v^\varepsilon(\theta)\|_{L^2(0,T;H^1(\mathbb{R}^n))}$$

$$\leqslant C\theta.$$

This results in the following uniform upper bound:

$$\|v^\varepsilon - v^\varepsilon(\theta)\|_{L^\infty(0,T;L^2(\mathbb{R}^n))}+\|v^\varepsilon - v^\varepsilon(\theta)\|_{L^2(0,T;H^1(\mathbb{R}^n))}\leqslant C\theta. \tag{40}$$

Now we want to estimate the difference $\hat a - \hat a(\theta)$. To this end we consider the corresponding auxiliary equations $\mathscr{A}\Psi^k = -\sum_i a^{ik}_{z_i}$ and $\mathscr{A}(\theta)\Psi^k(\theta) = -\sum_i a^{ik}_{z_i}(\theta)$. Subtracting we find

$$\mathscr{A}(\Psi^k - \Psi^k(\theta)) = \sum_i (a^{ik}_{z_i}(\theta) - a^{ik}_{z_i}) + (\mathscr{A}(\theta)-\mathscr{A})\Psi^k(\theta). \tag{41}$$

Again, multiplying by $\Psi^k - \Psi^k(\theta)$ and integrating by parts, one has

$$\int_{\mathbb{T}^n}\sum_{ij}a^{ij}\frac{\partial}{\partial z_i}(\Psi^k - \Psi^k(\theta))\frac{\partial}{\partial z_j}(\Psi^k - \Psi^k(\theta))\,\mathrm{d}z$$

$$=-\int_{\mathbb{T}^n}\sum_i(a^{ik}(\theta)-a^{ik})\frac{\partial}{\partial z_i}(\Psi^k - \Psi^k(\theta))\,\mathrm{d}z$$

$$-\int_{\mathbb{T}^n}\sum_{ij}(a^{ij}(\theta)-a^{ij})\frac{\partial}{\partial z_i}\Psi^k(\theta)\frac{\partial}{\partial z_j}(\Psi^k - \Psi^k(\theta))\,\mathrm{d}z$$

$$\leqslant C\theta\|\Psi^k - \Psi^k(\theta)\|_{H^1(\mathbb{T}^n)} + C\theta\|\Psi^k - \Psi^k(\theta)\|_{H^1(\mathbb{T}^n)}\|\Psi^k(\theta)\|_{H^1(\mathbb{T}^n)}$$

$$\leqslant C\theta.$$

Therefore, $\|\nabla\Psi^k - \nabla\Psi^k(\theta)\|_{L^2(\mathbb{T}^n)} \leqslant C\sqrt{\theta}$ uniformly in $y \in \mathbb{R}^d$. Hence,

$$|\hat{a} - \hat{a}(\theta)| = \left| \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} [(I + \nabla\Psi)a - (I + \nabla\Psi(\theta))a(\theta)]p(y)\,\mathrm{d}y\,\mathrm{d}z \right| \leqslant C\sqrt{\theta}. \quad (42)$$

It remains to note that (38), (40), and (42) imply (30). Proposition 4.2 is completely proved. □

It follows from Proposition 4.2 that v^ε converges in probability to the solution $\hat{v}$ of (29). In order to show this, let us introduce a bounded continuous functional $\Phi_\varphi(u)$ on the space Ω_T:

$$\Phi_\varphi(u) = 1 \wedge \sup_{t\leqslant T} \left| (\varphi(t), u(t)) - (\varphi(0), u_0) - \int_0^t ((\varphi_s(s), u(s))\,\mathrm{d}s \right.$$
$$\left. - \int_0^t \left(\sum_{ij} \hat{a}^{ij}\varphi_{x_ix_j}(s), u(s) \right) \mathrm{d}s \right|.$$

From (30), we get $\lim_{\varepsilon\downarrow 0} E\Phi_\varphi(v^\varepsilon) = 0$, so for any limiting point Q of the family $\tilde{Q}^\varepsilon$, we obtain $E^Q\Phi_\varphi(u) = 0$ and therefore any limiting measure Q is concentrated on the weak solution of the deterministic equation (29). Thus, the uniqueness of the solution of the latter problem implies the desired convergence in probability. □

Case $\alpha = 2$: We follow the same scheme as above: let $\hat{v}(x,t)$ be the solution of the Cauchy problem

$$\hat{v}_t = \operatorname{div}(\hat{a}\nabla\hat{v}) - \hat{b}\nabla\hat{v} + \hat{\tilde{c}}\hat{v}, \ \ \hat{v}(x,0) = u_0(x), \quad (43)$$

where $\hat{a}$ and $\hat{b}$ were defined in Theorem 3.1, and $\hat{\tilde{c}} = \overline{\langle \tilde{G}\tilde{c}\rangle}$.

We are going to prove that any limit point of the family $\{\tilde{Q}^\varepsilon\}$ is the δ-type measure concentrated on $\hat{v}$.

Proposition 4.4. *For any test function $\varphi(t,x) \in C^\infty([0,T]; C_0^\infty(R^n))$ the following limit relation holds*:

$$\lim_{\varepsilon\downarrow 0} E \sup_{t\leqslant T} \left| (\varphi(t), v^\varepsilon(t)) - (\varphi(0), u_0) - \int_0^t (\varphi_s(s), v^\varepsilon(s))\,\mathrm{d}s \right.$$
$$\left. - \int_0^t \left(\left[\sum_{ij} \hat{a}^{ij}\varphi_{x_ix_j}(s) + \hat{b}\nabla\varphi(s) + \hat{\tilde{c}}\varphi(s) \right], v^\varepsilon(s) \right) \mathrm{d}s \right| = 0.$$

Proof. Consider the auxiliary process:

$$H^\varepsilon(t) = (\varphi, v^\varepsilon(t)) + \varepsilon \sum_k \left(\Psi^k\left(\frac{\cdot}{\varepsilon}, \xi_{t/\varepsilon^2}\right)\varphi_{x_k}, v^\varepsilon(t) \right) + \varepsilon\left(\tilde{G}\left(\frac{\cdot}{\varepsilon}, \xi_{t/\varepsilon^2}\right)\varphi, v^\varepsilon(t) \right)$$

where Ψ^k is defined by (14) and $\tilde{G}(z,y)$ satisfies the equation

$$(\mathscr{A} + \mathscr{L})\tilde{G}(z,y) = -\tilde{c}(z,y). \quad (44)$$

It follows from Proposition 4.1, (E1), (E2), Lemma 2.5 and Proposition 2.6 that:

$$\lim_{\varepsilon\downarrow 0} E \sup_{t\leqslant T} |H^\varepsilon(t) - (\varphi, v^\varepsilon(t))| = 0.$$

Applying Ito's formula to $H^\varepsilon(t)$ gives

$$\begin{aligned}
\mathrm{d}H^\varepsilon = {} & \sum_{ij} [(a^{\varepsilon,ij}\varphi_{x_ix_j}, v^\varepsilon) + \varepsilon^{-1}(a^{\varepsilon,ij}_{z_i}\varphi_{x_j}, v^\varepsilon)]\,\mathrm{d}t + \varepsilon^{-1}(\tilde{c}^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t \\
& + (\nabla_y \tilde{G}^\varepsilon\varphi + \sum_k \nabla_y \Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon)\cdot\sigma^\varepsilon\,\mathrm{d}W_t \\
& + \Bigg[\varepsilon^{-1}\sum_k(\mathscr{L}\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon) + \sum_k(\Psi^{\varepsilon,k}\tilde{c}^\varepsilon\varphi_{x_k}, v^\varepsilon) + (\varphi_t, v^\varepsilon) \\
& \quad + \varepsilon^{-1}\sum_k(\mathscr{A}\Psi^{\varepsilon,k}\,\varphi_{x_k}, v^\varepsilon) + \varepsilon\sum_k(\Psi^{\varepsilon,k}\,\varphi_{tx_k}, v^\varepsilon) \\
& \quad + \sum_{ijk}((a^{\varepsilon,ij}\Psi^{\varepsilon,k})_{z_i}\varphi_{x_kx_j} + a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j}\varphi_{x_kx_i}, v^\varepsilon) \\
& \quad + \varepsilon\sum_{ijk}(a^{\varepsilon,ij}\Psi^{\varepsilon,k}\varphi_{x_ix_jx_k}, v^\varepsilon) \\
& \quad + \varepsilon^{-1}(\mathscr{L}\tilde{G}^\varepsilon\varphi, v^\varepsilon) + (\tilde{G}^\varepsilon\tilde{c}^\varepsilon\varphi, v^\varepsilon) \\
& \quad + \varepsilon^{-1}(\mathscr{A}\tilde{G}^\varepsilon\varphi, v^\varepsilon) + \sum_{ij}((a^{\varepsilon,ij}\tilde{G}^\varepsilon)_{z_i}\varphi_{x_j} + a^{\varepsilon,ij}\tilde{G}^\varepsilon_{z_j}\varphi_{x_i}, v^\varepsilon) \\
& \quad + \varepsilon\sum_{ij}(a^{\varepsilon,ij}\tilde{G}^\varepsilon\varphi_{x_ix_j}, v^\varepsilon)\Bigg]\,\mathrm{d}t\,.
\end{aligned}$$

Taking into account Eqs. (14) and (44) we simplify the expression on the right-hand side as follows:

$$\begin{aligned}
\mathrm{d}H^\varepsilon = {} & \Bigg[\sum_{ij}(a^{\varepsilon,ij}\,\varphi_{x_ix_j}, v^\varepsilon) + \sum_{ijk}((a^{\varepsilon,ij}\Psi^{\varepsilon,k})_{z_i}\varphi_{x_kx_j} + a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_j}\varphi_{x_kx_i}, v^\varepsilon) \\
& + \sum_k(\Psi^{\varepsilon,k}\tilde{c}^\varepsilon\varphi_{x_k}, v^\varepsilon) + (\tilde{G}^\varepsilon\tilde{c}^\varepsilon\varphi, v^\varepsilon) \\
& + \sum_{ij}((a^{\varepsilon,ij}\tilde{G}^\varepsilon)_{z_i}\,\varphi_{x_j} + a^{\varepsilon,ij}\tilde{G}^\varepsilon_{z_j}\varphi_{x_i}, v^\varepsilon) + (\varphi_t, v^\varepsilon)\Bigg]\,\mathrm{d}t \\
& + \left(\nabla_y\tilde{G}^\varepsilon\varphi + \sum_k \nabla_y\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\cdot\sigma^\varepsilon\,\mathrm{d}W_t + \varepsilon\,\mathrm{d}R^\varepsilon(t)
\end{aligned} \tag{45}$$

where $E\sup_{t\leqslant T}|R^\varepsilon(t)|\leqslant C$. Denote by $M^\varepsilon(t)$ the stochastic term on the right-hand side of the latter formula:

$$M^\varepsilon(t)=\int_0^t\left(\nabla_y\tilde G\left(\frac{\cdot}{\varepsilon},\xi_{s/\varepsilon^2}\right)\varphi+\sum_k\nabla_y\Psi^{\varepsilon,k}\,\varphi_{x_k},v^\varepsilon(s)\right)\cdot\sigma(\xi_{s/\varepsilon^2})\varepsilon\,\mathrm{d}W_{s/\varepsilon^2},$$

clearly $M^\varepsilon(t)$ is $\mathscr{F}^\varepsilon_t$-adapted square integrable martingale.

Proposition 4.5. *The following limit relation holds true*:

$$\lim_{\varepsilon\to 0}E\sup_{t\leqslant T}|M^\varepsilon(t)|=0.$$

Proof. It is easy to see that the operator $\mathscr{L}$ does commute with averaging in variable z. Thus, taking the mean value in z in Eqs. (14) and (44), we find $\langle\nabla_y\Psi(\cdot,y)\rangle=0$ and $\langle\nabla_y\tilde G(\cdot,y)\rangle=0$ for all y. Therefore,

$$E\sup_{t\leqslant T}\left|\int_0^t\left[\sum_k(\nabla_y\Psi^{\varepsilon,k}\varphi_{x_k},v^\varepsilon(s))+(\nabla_y\tilde G^\varepsilon,\varphi v^\varepsilon(s))\right]\cdot\sigma(\xi_{s/\varepsilon^2})\,\mathrm{d}W_s\right|$$

$$=\varepsilon E\sup_{t\leqslant T}\left|\int_0^t\left[\sum_k(H^{\varepsilon,k},\nabla(\varphi_{x_k}v^\varepsilon(s)))+(\tilde H^\varepsilon,\nabla(\varphi v^\varepsilon(s)))\right]\cdot\sigma(\xi_{s/\varepsilon^2})\,\mathrm{d}W_s\right|$$

where $H^k(z,y)$ and $\tilde H(z,y)$ are given by

$$\mathrm{div}_z\,H^k(z,y)=\nabla_y\Psi^k(z,y),$$

$$\mathrm{div}_z\,\tilde H(z,y)=\nabla_y\tilde G(z,y).$$

By virtue of (E1), (E2), and the Burkholder–Davis–Gundy inequality we have:

$$\lim_{\varepsilon\downarrow 0}E\sup_{t\leqslant T}\left|\sum_k\int_0^t(\nabla_y\Psi^{\varepsilon,k}\varphi_{x_k},v^\varepsilon(s))\cdot\sigma(\xi_{s/\varepsilon^2})\,\mathrm{d}W_s\right|=0,$$

$$\lim_{\varepsilon\downarrow 0}E\sup_{t\leqslant T}\left|\int_0^t(\nabla_y\tilde G^\varepsilon\varphi,v^\varepsilon(s))\cdot\sigma(\xi_{s/\varepsilon^2})\,\mathrm{d}W_s\right|=0.\qquad\square$$

Passing to the limit on the right-hand side of (45) by means of Proposition 4.5 and Lemma 4.3, we complete the proof of Proposition 4.4.

Now, the convergence of v^ε in probability can be derived in the same way as in the case $\alpha<2$.

Case $\alpha>2$: The approach used in this case is quite similar to that of the preceding cases so we consider it briefly. We introduce a function $\hat v(x,t)$ to satisfy the Cauchy problem

$$\hat v_t=\mathrm{div}(\hat a\nabla\hat v)+\hat c\hat v,\ \ \hat v(x,0)=u_0(x),\tag{46}$$

with $\hat a$ and $\hat c$ defined in Theorem 3.3. The proof of the fact that v^ε converges to $\hat v$ in probability relies on the following.

Proposition 4.6. *For any test function* $\varphi(x,t)\in C^{\infty}([0,T];\ C_0^{\infty}(R^n))$ *the following limit relation holds*:

$$\lim_{\varepsilon\downarrow 0} E \sup_{t\leqslant T}\left|(\varphi(t),v^{\varepsilon}(t))-(\varphi(0),u_0)-\int_0^t(\varphi_s(s),v^{\varepsilon}(s))\ \mathrm{d}s\right.$$
$$\left.-\int_0^t\left(\sum_{ij}\hat{a}^{ij}\varphi_{x_ix_j}(s)+\hat{c}\varphi(s),v^{\varepsilon}(s)\right)\mathrm{d}s\right|=0. \qquad (47)$$

Proof. We define the functions $G\in\bar{H}^1(\mathbb{T}^n)$, $g\in\bar{H}^1(\mathbb{T}^n\times\mathbb{R}^d)$, to be solutions of the system of equations:

$$\bar{\mathscr{A}}G(z)=-\overline{c(z,\cdot)},$$

$$\mathscr{L}g(z,y)=-[\tilde{c}(z,y)-\overline{c(z,\cdot)}+\mathscr{A}G(z)-\bar{\mathscr{A}}G(z)], \qquad (48)$$

and the functions $\Psi^k\in\bar{H}^1(\mathbb{T}^n)$ and $\Phi^k\in\bar{H}^1_{\rho}(\mathbb{T}^n\times\mathbb{R}^d)$, $k=1,\dots,n$, to satisfy the system

$$\bar{\mathscr{A}}\Psi^k(z)=-\sum_i\overline{a^{ik}_{z_i}(z,\cdot)},$$
$$\mathscr{L}\Phi^k(z,y)=-\left[(\mathscr{A}-\bar{\mathscr{A}})\Psi^k(z)+\sum_i(a^{ik}_{z_i}(z,y)-\overline{a^{ik}_{z_i}(z,\cdot)})\right] \qquad (49)$$

$k=1,\dots,n$, $\delta=\alpha-2>0$. Applying Ito's formula to the expression

$$H^{\varepsilon}=(\varphi,v^{\varepsilon})+\varepsilon\left(\sum_k\Psi^{\varepsilon,k}\varphi_{x_k},v^{\varepsilon}\right)+\varepsilon^{1+\delta}\left(\sum_k\Phi^{\varepsilon,k}\varphi_{x_k},v^{\varepsilon}\right)$$
$$+\varepsilon(G^{\varepsilon}\varphi,v^{\varepsilon})+\varepsilon^{1+\delta}(g^{\varepsilon}\varphi,v^{\varepsilon})$$

we obtain after simple transformations:

$$\mathrm{d}H^{\varepsilon}=(\varphi_t,v^{\varepsilon})\,\mathrm{d}t+\varepsilon^{-1}\left(\sum_{ij}a^{\varepsilon,ij}_{z_i}\varphi_{x_j},v^{\varepsilon}\right)\mathrm{d}t+\varepsilon^{-1}(\tilde{c}^{\varepsilon}\varphi,v^{\varepsilon})\,\mathrm{d}t$$
$$+\left(\sum_{ij}a^{\varepsilon,ij}\varphi_{x_ix_j},v^{\varepsilon}\right)\mathrm{d}t+\left(\sum_{ij}a^{\varepsilon,ij}\varphi_{x_i},v^{\varepsilon}_{x_j}\right)\mathrm{d}t$$
$$+\varepsilon^{-1}\left(\sum_k\mathscr{A}\Psi^{\varepsilon,k}\varphi_{x_k},v^{\varepsilon}\right)\mathrm{d}t+\left(\sum_k\sum_{ij}a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_i}\varphi_{x_kx_j},v^{\varepsilon}\right)\mathrm{d}t$$
$$+\varepsilon\left(\sum_k\sum_{ij}a^{\varepsilon,ij}\Psi^{\varepsilon,k}\varphi_{x_kx_i},v^{\varepsilon}_{x_j}\right)\mathrm{d}t+\varepsilon\left(\sum_k\Psi^{\varepsilon,k}\,\varphi_{tx_k},v^{\varepsilon}\right)\mathrm{d}t$$
$$+\left(\sum_k\tilde{c}^{\varepsilon}\Psi^{\varepsilon,k}\varphi_{x_k},v^{\varepsilon}\right)\mathrm{d}t$$

$$+\,\varepsilon^{\delta-1}\left(\sum_k \mathscr{A}\Phi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\mathrm{d}t + \varepsilon^{\delta}\left(\sum_k\sum_{ij} a^{\varepsilon,ij}\Phi^{\varepsilon,k}_{z_j}\varphi_{x_kx_j}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,\varepsilon^{\delta+1}\left(\sum_k\sum_{ij} a^{\varepsilon,ij}\Phi^{\varepsilon,k}\varphi_{x_kx_i}, v^\varepsilon_{x_j}\right)\mathrm{d}t + \varepsilon^{\delta+1}\left(\sum_k \Phi^{\varepsilon,k}\varphi_{tx_k}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,\varepsilon^{\delta}\left(\sum_k \tilde{c}^\varepsilon\Phi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\mathrm{d}t + \varepsilon^{-1}\left(\sum_k (\mathscr{L}\Phi)^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,\varepsilon^{\delta/2}\left(\sum_k \nabla_y\Phi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\mathrm{d}W_t$$

$$+\,\varepsilon^{-1}(\mathscr{A}G^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + \left(\sum_{ij} a^{\varepsilon,ij}G^\varepsilon_{z_i}\varphi_{x_kx_j}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,\varepsilon\left(\sum_{ij} a^{\varepsilon,ij}G^\varepsilon\varphi_{x_kx_i}, v^\varepsilon_{x_j}\right)\mathrm{d}t + \varepsilon(G^\varepsilon\varphi_t, v^\varepsilon)\,\mathrm{d}t + (\tilde{c}^\varepsilon G^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t$$

$$+\,\varepsilon^{\delta-1}((\mathscr{A}g)^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{\delta}\left(\sum_k\sum_{ij} a^{\varepsilon,ij}g^\varepsilon_{z_i}\varphi_{x_kx_j}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,\varepsilon^{\delta+1}\left(\sum_k\sum_{ij} a^{\varepsilon,ij}g^\varepsilon\varphi_{x_kx_i}, v^\varepsilon_{x_j}\right)\mathrm{d}t + \varepsilon^{\delta+1}(g^\varepsilon\varphi_t, v^\varepsilon)\,\mathrm{d}t$$

$$+\,\varepsilon^{\delta}(\tilde{c}^\varepsilon g^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{-1}((\mathscr{L}g)^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{\delta/2}(\nabla_y g^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}W_t.$$

According to (48) and (49) the following terms on the right-hand side are mutually cancelled:

$$\varepsilon^{-1}\left(\sum_{ij} a^{\varepsilon,ij}_{z_i}\varphi_{x_j}, v^\varepsilon\right) + \varepsilon^{-1}\left(\sum_k \mathscr{A}\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right) + \varepsilon^{-1}\left(\sum_k \mathscr{L}\Phi^{\varepsilon,k}\ \varphi_{x_k}, v^\varepsilon\right) = 0,$$

$$\varepsilon^{-1}(\tilde{c}^\varepsilon\varphi, v^\varepsilon) + \varepsilon^{-1}(\mathscr{A}G^\varepsilon\varphi, v^\varepsilon) + \varepsilon^{-1}(\mathscr{L}g^\varepsilon\varphi, v^\varepsilon) = 0,$$

and we obtain

$$\mathrm{d}H^\varepsilon = (\varphi_t, v^\varepsilon)\,\mathrm{d}t + \left(\sum_{ij} a^{\varepsilon,ij}\varphi_{x_ix_j}, v^\varepsilon\right)\mathrm{d}t + \left(\sum_k\sum_{ij} a^{\varepsilon,ij}\Psi^{\varepsilon,k}_{z_i}\varphi_{x_kx_j}, v^\varepsilon\right)\mathrm{d}t$$

$$+\left(\sum_k \tilde{c}^\varepsilon\Psi^{\varepsilon,k}\varphi_{x_k}, v^\varepsilon\right)\mathrm{d}t + \left(\sum_{ij} a^{\varepsilon,ij}G^\varepsilon_{z_i}\varphi_{x_j}, v^\varepsilon\right)\mathrm{d}t$$

$$+\,(\tilde{c}^\varepsilon G^\varepsilon\varphi, v^\varepsilon)\,\mathrm{d}t + \varepsilon^{\delta/2\wedge 1}\,\mathrm{d}\tilde{R}^\varepsilon$$

where $E\sup_{t\leqslant T}|\tilde{R}^\varepsilon(t)| \leqslant C$.

In the same way as above one can check that (33) holds. The passage to the limit is straightforward by virtue of Lemma 4.3.

It remains to show that the first-order terms in the limit equation are equal to zero or, equivalently, that

$$\langle \bar{c}\Psi^k + \bar{a}\nabla G \rangle = 0.$$

Indeed, Definitions (48) of G and (49) of Ψ^k lead to

$$\begin{aligned}\langle \bar{c}\Psi^{\varepsilon,k} + \bar{a}\nabla G \rangle &= \langle -(\bar{\mathscr{A}}G)\Psi^{\varepsilon,k} + \bar{a}\,\nabla G \rangle \\ &= \langle G(\bar{\mathscr{A}}\Psi^{\varepsilon,k} + \nabla \bar{a}) \rangle \\ &= 0. \qquad \square\end{aligned}$$

4.4. Convergence in law of the solution of initial problem – description of the limit distribution

In this section we establish the convergence of the family $\{u^\varepsilon\}$ in law. We start by studying the limit distribution of the exponent from (25). According to Pardoux and Veretennikov (2001, Theorem 3) the family of integrals

$$J_t^{\varepsilon,\alpha} = \frac{1}{\varepsilon^{\alpha/2}} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \, \mathrm{d}s$$

converges in law in $\mathscr{C}[0,T]$ as $\varepsilon \to 0$, to $\lambda \hat{W}_t$, where $\hat{W}_t$ is a standard one-dimensional Brownian motion and λ is defined as follows:

$$\lambda^2 = 2\overline{q\nabla G{\cdot}\nabla G}, \qquad \mathscr{L}G = -\langle c(\cdot\,, y) \rangle.$$

In fact, this definition coincides with that of Theorems 3.1 and 3.2. In particular, the above result implies that for $\alpha > 2$

$$P\text{-}\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \, \mathrm{d}s = 0$$

and, therefore,

$$P\text{-}\lim_{\varepsilon \to 0} \exp\left(\frac{1}{\varepsilon} \int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha}) \rangle \, \mathrm{d}s \right) = 1. \tag{50}$$

For $\alpha \leqslant 2$ the family $\exp(J_t^{\varepsilon,\alpha})$ converges in law to the distribution of $\exp(\lambda \hat{W}_t)$.

In order to pass to the limit in the product $\exp(J_t^{\varepsilon,\alpha})\, v^\varepsilon$, we represent it in the form

$$\exp(J_t^{\varepsilon,\alpha}) v^\varepsilon = \exp(J_t^{\varepsilon,\alpha}) \hat{v} + \exp(J_t^{\varepsilon,\alpha})(v^\varepsilon - \hat{v}). \tag{51}$$

It is easy to see that the map $v(\cdot) \mapsto v(\cdot)\hat{v}$ is a continuous map from $\mathscr{C}[0,T]$ to Ω_T. Thus, the first term on the right-hand side of (51) converges in law to $\exp(\lambda \hat{W}_t)\hat{v}$ in Ω_T as $\varepsilon \to 0$.

The second term converges in probability to zero. Indeed, by Prokhorov's theorem (Billingsley, 1968), for any $\gamma > 0$ there is a compact subset K^γ of $\mathscr{C}[0,T]$ such that

$$P\{\exp(J_\cdot^{\varepsilon,\alpha}) \in K^\gamma\} \geqslant 1 - \gamma.$$

For a finite γ-net $f_1(t), f_2(t), \dots, f_N(t)$ in K^γ, we construct mutually nonintersecting sets $K_1^\gamma, K_2^\gamma, \dots, K_N^\gamma$ such that K_l^γ belongs to γ-neighborhood of $f_l(\cdot)$, $l = 1, 2, \dots, N$, and $K^\gamma = \bigcup_l K_l^\gamma$. Now one can rewrite the expression studied as follows:

$$\begin{aligned}\exp(J_t^{\varepsilon,\alpha})(v^\varepsilon - \hat{v}) = {} & \mathbf{1}_{\exp(J_t^{\varepsilon,\alpha}) \notin K^\gamma} \exp(J_t^{\varepsilon,\alpha})(v^\varepsilon - \hat{v}) \\ & + \sum_l \mathbf{1}_{\exp(J_t^{\varepsilon,\alpha}) \in K_l^\gamma} f_l(t)\,(v^\varepsilon - \hat{v}) \\ & + \sum_l \mathbf{1}_{\exp(J_t^{\varepsilon,\alpha}) \in K_l^\gamma} [\exp(J_t^{\varepsilon,\alpha}) - f_l(t)]\,(v^\varepsilon - \hat{v}). \end{aligned} \tag{52}$$

It remains to notice that by (E1) and (E2)

$$\begin{aligned} & \|\mathbf{1}_{\exp(J_t^{\varepsilon,\alpha}) \in K_l^\gamma} [\exp(J_t^{\varepsilon,\alpha}) - f_l(t)]\,(v^\varepsilon - \hat{v})\|_{L^\infty(0,T;L^2(R^n))} \\ & \quad + \|\mathbf{1}_{\exp(J_t^{\varepsilon,\alpha}) \in K_l^\gamma} [\exp(J_t^{\varepsilon,\alpha}) - f_l(t)]\,(v^\varepsilon - \hat{v})\|_{L^2(0,T;H^1(R^n))} \leqslant c\,\gamma, \end{aligned}$$

and pass to the limit in (52) first as $\varepsilon \to 0$ and then as $\gamma \to 0$.

Finally, in order to show that the limit distributions obtained satisfy the limit SPDEs (9) or (12) one can apply Ito's formula to the product $\exp(\lambda \hat{W}_t)\,\hat{v}(t,x)$ and use the auxiliary homogenized equations (29) or (43), respectively.

With evident simplifications, one can use (50) to pass to the limit in the product:

$$u^\varepsilon(t,x) = \exp\left(\frac{1}{\varepsilon}\int_0^t \langle c(\cdot\,, \xi_{s/\varepsilon^\alpha})\rangle\,\mathrm{d}s\right)\, v^\varepsilon(t,x)$$

in the case $\alpha > 2$, and to show that the functions v^ε converge in probability in Ω_T to the solution $\hat{v}$ of problem (46).

Acknowledgements

During the review process the referees made a number of very important remarks and suggestions. We express our deep gratitude to them.

Appendix A

In this appendix we prove the properties of the process ξ_t, of its invariant density, and of a solution to the Poisson equation (7) which were formulated throughout the paper, in particular Lemma 2.5 and Proposition 2.6.

Lemma A.1. *Under Hypotheses* 2.1–2.3, *the process* ξ_t *governed by the operator* $\mathscr{L}$ *possesses a unique invariant measure whose density decays at the infinity faster than any negative power of* $|y|$.

Proof. The existence and uniqueness of the invariant measure of ξ_t have been obtained in Veretennikov (1997, Theorem 6) and Pardoux and Veretennikov (2001, Proposition 1). The density $\rho(y)$ of the invariant measure satisfies the stationary Kolmogorov equation:

$$\mathscr{L}^* \rho = 0 \tag{A.1}$$

that is

$$\begin{aligned}
&\sum_{ij} \frac{\partial^2}{\partial y_i \, \partial y_j} (q_{ij}(y)\, \rho(y)) - \sum_i \frac{\partial}{\partial y_i} (B_i(y)\, \rho(y)) \\
&\quad = \sum_{ij} \frac{\partial}{\partial y_i} \left(q_{ij}(y) \frac{\partial}{\partial y_j} \rho(y) \right) + \sum_{ij} \frac{\partial}{\partial y_i} \left(\left(\frac{\partial}{\partial y_j} q_{ij}(y) \right) \rho(y) \right) \\
&\qquad - \sum_i \frac{\partial}{\partial y_i} (B_i(y)\, \rho(y)) \\
&\quad = 0.
\end{aligned}$$

Our regularity assumptions of Hypotheses 2.1 and 2.2 imply (see, for instance, Gilbarg and Trudinger, 1994, Theorem 8.24) the Hölder continuity of $\rho(y)$ and the following Harnack inequality (cf. Gilbarg and Trudinger, 1994, Theorem 8.20): in a ball

$$Q_{y_0} = \left\{ y; \frac{|y - y_0|}{(1 + |y_0|)^{\mu_1}} \leqslant 1 \right\}$$

we have

$$\frac{\max_{y \in Q_{y_0}} \rho(y)}{\min_{y \in Q_{y_0}} \rho(y)} \leqslant C. \tag{A.2}$$

Indeed, if we make a change of variables

$$y' = (1 + |y_0|)^{\mu_1} \, y = \theta(y_0) \, y$$

then in the coordinates y', Eq. (A.1) reads

$$\begin{aligned}
&\sum_{ij} \frac{\partial}{\partial y'_i} \left(q_{ij} \left(\frac{y'}{\theta(y_0)} \right) \frac{\partial}{\partial y'_j} \rho \left(\frac{y'}{\theta(y_0)} \right) \right) \\
&\quad + \sum_{ij} \frac{\partial}{\partial y'_i} \left(\left(\frac{\partial}{\partial y'_j} q_{ij} \left(\frac{y'}{\theta(y_0)} \right) \right) \rho \left(\frac{y'}{\theta(y_0)} \right) \right) \\
&\quad - \frac{1}{\theta(y_0)} \sum_i \frac{\partial}{\partial y'_i} \left(B_i \left(\frac{y'}{\theta(y_0)} \right) \rho \left(\frac{y'}{\theta(y_0)} \right) \right) = 0.
\end{aligned}$$

In the ball $\{y'; |y' - y_0|/\theta(y_0) \leqslant 2\}$, the coefficients of the latter equation admit an upper bound uniformly in y_0. Therefore, in the smaller ball

$$\frac{1}{\theta(y_0)} Q_{y_0} = \{y'; |y' - y_0|/\theta(y_0) \leqslant 1\}$$

we have

$$\frac{\max_{y' \in \frac{1}{\theta(y_0)} Q_{y_0}} \rho(y'/\theta_0)}{\min_{y' \in \frac{1}{\theta(y_0)} Q_{y_0}} \rho(y'/\theta_0)} \leqslant C$$

and (54) follows. Proposition 1 of Pardoux and Veretennikov (2001) states that, for any $m>0$

$$\int_{Q_{y_0}} (1+|y|^m)\,\rho(y)\,\mathrm{d}y \leqslant C(m). \tag{A.3}$$

From (A.2) and (A.3) we obtain

$$\max_{y\in Q_{y_0}} \rho(y) \leqslant C_1(m)\,\frac{(1+|y_0|)^{\mu_1}}{(1+|y_0|)^m} = C_1(m)\,(1+|y_0|)^{\mu_1-m}.$$

Finally, the desired statement follows from the fact that m is an arbitrary positive number. □

Proof of Lemma 2.5. The statement of Lemma 2.5 is similar to that of Theorem 1 of Pardoux and Veretennikov (2001), but in our case the diffusion process under consideration takes on values on the product of the torus $\mathbb{T}^n$ and the whole space $\mathbb{R}^d$ while, in the cited paper, a process with values in $\mathbb{R}^d$ is studied.

A simple analysis of the proof of Theorem 1 of Pardoux and Veretennikov (2001) shows that this statement relies crucially on the estimate of the variance of the difference between invariant measure of a diffusion process and the law of this process issued from a given point at time 0.

Denote by $v^{x,y}(\mathrm{d}x',\mathrm{d}y')$ the distribution of the diffusion process (η_t^x,ξ_t^y) governed by the operator $\mathscr{A}+\mathscr{L}$ with the initial conditions $\eta_0^x=x$, $\xi_0^y=y$, $x\in\mathbb{T}^n$, $y\in\mathbb{R}^d$. Then we have

$$v^{x,y}(\mathrm{d}x',\mathrm{d}y') = p(t,x,y,x',y')\,\mathrm{d}x'\,\mathrm{d}y'$$

where a positive continuous density $p(t,x,y,x',y')$ exists for all $t>0$.

Our aim is to obtain an estimate

$$\mathrm{var}(v_t^{x,y}-v_{\mathrm{inv}}) \leqslant (1+t)^{-\kappa}\,(1+|y|)^m \tag{A.4}$$

with arbitrary $m>0$, $0<\kappa<m/2$; here v_{inv} is the invariant measure of (η_t,ξ_t):

$$v_{\mathrm{inv}}(\mathrm{d}x',\mathrm{d}y') = \rho(y')\,\mathrm{d}x'\,\mathrm{d}y'. \tag{A.5}$$

Denote by $\tilde{v}_t^y(\mathrm{d}y') = \tilde{p}(t,y,y')\,\mathrm{d}y'$ the distribution of ξ_t^y, $t>0$. Since the coefficients of $\mathscr{L}$ do not depend on x, we have an evident relation

$$\int_{\mathbb{T}^n} p(t,x,y,x',y')\,\mathrm{d}x' = \tilde{p}(t,y,y'). \tag{A.6}$$

We want to estimate

$$\begin{aligned}\mathrm{var}(v_t^{x,y}-v_{\mathrm{inv}}) &= \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} |p(t,x,y,x',y')-\rho(y')|\,\mathrm{d}x'\,\mathrm{d}y' \\ &\leqslant \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} |p(t,x,y,x',y')-\tilde{p}(t,y,y')|\,\mathrm{d}x'\,\mathrm{d}y' \\ &\quad + \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} |\tilde{p}(t,y,y')-\rho(y')|\,\mathrm{d}x'\,\mathrm{d}y'\end{aligned}$$

$$= \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} |p(t,x,y,x',y') - \tilde{p}(t,y,y')|\,\mathrm{d}x'\,\mathrm{d}y'$$

$$+ \int_{\mathbb{R}^d} |\tilde{p}(t,y,y') - \rho(y')|\,\mathrm{d}y'.$$

The second integral of the right-hand side has been estimated in Veretennikov (1997, Theorem 6), and Pardoux and Veretennikov (2001, Eq. (6)).

We are going to show that the first one admits the upper bound:

$$\mathrm{var}(v_t^{x,y} - \tilde{v}_t^y) = \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} |p(t,x,y,x',y') - \tilde{p}(t,y,y')|\,\mathrm{d}x'\,\mathrm{d}y' \leqslant C\,\mathrm{e}^{-\gamma t}, \quad t>0. \tag{A.7}$$

Denote by $\tilde{\mathscr{Q}}(y,\mathrm{d}\varphi(\cdot))$ the law of $\xi_\cdot^y$ in the space $(\mathscr{C}[0,T])^d$ of the continuous trajectories in $\mathbb{R}^d$, and by $\mathscr{Q}(t,x,y,\mathrm{d}x',\mathrm{d}\varphi(\cdot))$ the joint distribution of $(\eta_t^x,\{\xi_\cdot^y\})$ in $\mathbb{T}^n \times (\mathscr{C}[0,T])^d$. For any trajectory $\xi_\cdot^y$, the process $\eta_\cdot^x$ is governed by the operator

$$\frac{\partial}{\partial x_i'} a_{ij}(x',\xi_t^y)\frac{\partial}{\partial x_j'}.$$

Thus

$$\mathscr{Q}(t,x,y,\mathrm{d}x',\mathrm{d}\varphi) = g(t,x,x',\varphi)\,\mathrm{d}x'\,\tilde{\mathscr{Q}}(y,\mathrm{d}\varphi) \tag{A.8}$$

where the function $g(t,x,x',\varphi)$ satisfies the following equation:

$$\frac{\partial}{\partial t} g(t,x,x',\varphi(\cdot)) = \sum_{ij}\frac{\partial}{\partial x_i'}\left(a_{ij}(x',\varphi(t))\frac{\partial}{\partial x_j'} g(t,x,x',\varphi(\cdot))\right), \quad g|_{t=0} = \delta(x'-x). \tag{A.9}$$

Now, we need the following lemma:

Lemma A.2. *Uniformly in* x *and* $\varphi(\cdot)$ *the following estimate holds*:

$$\int_{\mathbb{T}^n} |g(t,x,x',\varphi(\cdot)) - 1|\,\mathrm{d}x' \leqslant C\,\mathrm{e}^{-\gamma t}, \quad \gamma>0. \tag{A.10}$$

Proof. By the Harnack inequality (Trudinger, 1968, Theorem 5.1):

$$\frac{\max_{x'} g(1,x')}{\min_{x'} g(1,x')} \leqslant C \tag{A.11}$$

where the constant C only depends on the upper and lower ellipticity constant of $\{a_{ij}\}$; for brevity here and afterwards we omit the arguments x and $\varphi(\cdot)$ of g.

In view of evident relation:

$$\int_{\mathbb{T}^n} g(t,x')\,\mathrm{d}x' = 1 \tag{A.12}$$

Eq. (A.11) implies the upper bound

$$\|g(1,\cdot)\|_{\mathscr{C}(\mathbb{T}^n)} \leqslant C.$$

We want to show that $|g(t,x')-1|$ decays exponentially as $t\uparrow\infty$. By (A.12), we have $\int_{\mathbb{T}^n}(g(t,x')-1)\,\mathrm{d}x'=0$; thus it is sufficient to show that

$$\operatorname*{osc}_{x'\in\mathbb{T}^n} g(t_0+1,x') \leqslant C \operatorname*{osc}_{x'\in\mathbb{T}^n} g(t_0,x') \quad \text{with } 0<C<1, \tag{A.13}$$

for any $t_0>1$. Making an appropriate linear transformation $k_1\,g+k_2$ and considering the fact that (A.13) is invariant with respect to such a transformation, we may assume, without loss of generality, that $\max_{x'\in\mathbb{T}^n} g(t_0,x')=1$, and $\min_{x'\in\mathbb{T}^n} g(t_0,x')=-1$. Denote $g^+(t_0,x')=g(t_0,x')\vee 0$, $g^-(t_0,x')=-(g(t_0,x')\wedge 0)$, and consider the following two problems:

$$\frac{\partial}{\partial t}\tilde g^{\pm}(t,x')=\sum_{ij}\frac{\partial}{\partial x'_i}\left(a_{ij}(x',\varphi(t))\frac{\partial}{\partial x'_j}\tilde g^{\pm}(t,x')\right), \quad x'\in\mathbb{T}^n,\ t\geqslant t_0,$$

$$\tilde g^{\pm}(t_0,x')=g^{\pm}(t_0,x').$$

Clearly

$$g(t,x')=\tilde g^{+}(t,x')-\tilde g^{-}(t,x'). \tag{A.14}$$

According to the maximum principle:

$$0\leqslant \tilde g^{\pm}(t,x')\leqslant 1,\ t\geqslant t_0. \tag{A.15}$$

Again, by the Harnack inequality one has

$$\max_{x'\in\mathbb{T}^n}\tilde g^{\pm}(t,x')\leqslant C\min_{x'\in\mathbb{T}^n}\tilde g^{\pm}(t,x'). \tag{A.16}$$

From (A.14)–(A.16) we derive

$$\begin{aligned}\max_{x'\in\mathbb{T}^n} g(t_0+1,x') &= \max_{x'\in\mathbb{T}^n}(\tilde g^{+}(t_0+1,x')-\tilde g^{-}(t_0+1,x'))\\ &\leqslant \max_{x'\in\mathbb{T}^n}\tilde g^{+}(t_0+1,x')-\min_{x'\in\mathbb{T}^n}\tilde g^{-}(t_0+1,x')\\ &\leqslant \max_{x'\in\mathbb{T}^n}\tilde g^{+}(t_0+1,x')-\frac{1}{C}\max_{x'\in\mathbb{T}^n}\tilde g^{-}(t_0+1,x').\end{aligned}$$

Similarly,

$$\min_{x'\in\mathbb{T}^n} g(t_0+1,x')\geqslant \frac{1}{C}\max_{x'\in\mathbb{T}^n}\tilde g^{+}(t_0+1,x')-\max_{x'\in\mathbb{T}^n}\tilde g^{-}(t_0+1,x').$$

Subtracting the two last inequalities leads to

$$\begin{aligned}\operatorname*{osc}_{x'\in\mathbb{T}^n} g(t_0+1,x') &\leqslant \left(1-\frac{1}{C}\right)\left(\max_{x'\in\mathbb{T}^n}\tilde g^{+}(t_0+1,x')-\max_{x'\in\mathbb{T}^n}\tilde g^{-}(t_0+1,x')\right)\\ &\leqslant \left(1-\frac{1}{C}\right)\operatorname*{osc}_{x'\in\mathbb{T}^n} g(t_0+1,x').\end{aligned}$$

It remains to put $\lambda=\log(1-1/C)$. $\square$

We can, now, return to the proof of Lemma 2.5.

By virtue of Lemma A.2 we obtain the following upper bound:

$$\begin{aligned}
&\operatorname{var}(\mathscr{Q}(t,x,y,\mathrm{d}x',\mathrm{d}\varphi) - \tilde{\mathscr{Q}}(y,\mathrm{d}\varphi)\,\mathrm{d}x') \\
&\quad \leqslant \int_{(\mathscr{C}[0,T])^d} \int_{\mathbb{T}^n} |g(t,x,x',\varphi(\cdot)) - 1|\tilde{\mathscr{Q}}(y,\mathrm{d}\varphi)\,\mathrm{d}x' \\
&\quad \leqslant C\,\mathrm{e}^{-\gamma t} \int_{(\mathscr{C}[0,T])^d} \tilde{\mathscr{Q}}(y,\mathrm{d}\varphi) \\
&\quad = C\,\mathrm{e}^{-\gamma t}
\end{aligned}$$

which, in turn, implies the estimate

$$\operatorname{var}(v_t^{x,y}(\mathrm{d}x',\mathrm{d}y') - v_t^{y}(\mathrm{d}y')\,\mathrm{d}x') \leqslant \operatorname{var}(\mathscr{Q}(t,x,y,\mathrm{d}x',\mathrm{d}\varphi) - \tilde{\mathscr{Q}}(y,\mathrm{d}\varphi)\,\mathrm{d}x') \leqslant C\,\mathrm{e}^{-\gamma t}. \tag{A.17}$$

In order to justify the first inequality here it suffices to note that if we reduce in the second variance the collection of test subsets and consider instead of all Borel subsets of $(\mathscr{C}[0,T])^d$ only the collection of cylindrical subsets of the form

$$\{\varphi(\cdot)\,;\,,\varphi(t)\in G\}$$

where G is a Borel set in $\mathbb{R}^d$, then we will obtain the first variance in (A.17). Estimate (A.4) is now straightforward. Now, let $f(x,y)$ satisfy the estimate

$$|f(x,y)| \leqslant C\,(1+|y|)^{\beta}$$

with some β. From the inequality of item A of the proof of Pardoux and Veretennikov (2001, Theorem 1), it follows that the function

$$u(x,y) = \int_0^{\infty} \mathrm{d}t \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} v_t^{x,y}(\mathrm{d}x',\mathrm{d}y')\,f(x',y')$$

is well-defined and

$$|u(x,y)| \leqslant C(m)\,(1+|y|)^m$$

for any $m > 2 \vee \beta$. Moreover,

$$\lim_{N\uparrow\infty} \sup_{x,y} \left((1+|y|)^{-m} \left| u(x,y) - \int_0^N \mathrm{d}t \int_{\mathbb{T}^n}\int_{\mathbb{R}^d} v_t^{x,y}(\mathrm{d}x',\mathrm{d}y')\,f(x',y') \right| \right) = 0.$$

Furthermore, items C and D of the said proof of Pardoux and Veretennikov (2001) only rely on probabilistic representation of solutions of Cauchy problem for a parabolic operator and Dirichlet problem for an elliptic operator, and on local regularity properties of these equations. Thus, all the arguments apply in our case and we have

$$(\mathscr{A} + \mathscr{L}^*)u(x,y) = f(x,y),$$

which ends the proof of Lemma 2.5. □

Proof of Proposition 2.6. As was shown in Pardoux and Veretennikov (2001, proof of Proposition 2)

$$E \sup_{0\leqslant s\leqslant t} |\xi_s^x|^p \leqslant C(p)\,(1+|x|)^{p-1/2}\sqrt{t}$$

for any $x \in \mathbb{R}^d$, $p>0$, $t>0$. By introducing $t = T/\varepsilon^\alpha$ with $T>0$ fixed, we get

$$E \sup_{0 \leqslant s \leqslant T} |\xi^x_{s/\varepsilon^\alpha}|^p \leqslant \sqrt{T}\, C(p)(1+|x|)^{p-1/2}\,\varepsilon^{-\alpha/2}.$$

Multiplying by ε^α gives

$$E\varepsilon^\alpha \sup_{0 \leqslant s \leqslant T} |\xi^x_{s/\varepsilon^\alpha}|^p \leqslant \sqrt{T}\, C(p)(1+|x|)^{p-1/2}\,\varepsilon^{\alpha/2}.$$

For the stationary process ξ_s distributed with the invariant law μ_{inv} we find

$$E_{\mu_{\text{inv}}}\left(\varepsilon^\alpha \sup_{0 \leqslant s \leqslant T} |\xi_{s/\varepsilon^\alpha}|^p\right) \leqslant \varepsilon^{\alpha/2}\sqrt{T}\, C(p)\int_{\mathbb{R}^d}(1+|x|)^{p-1/2}\,\mu_{\text{inv}}(\mathrm{d}x)$$
$$\leqslant \sqrt{T}\, C_1(p)\,\varepsilon^{\alpha/2}.$$

Here Proposition 2.5 has also been used. This implies the required statement for $\beta \geqslant \alpha$. In case $\beta < \alpha$ we have by Hölder inequality

$$E_{\mu_{\text{inv}}}\left(\varepsilon^\beta \sup_{0 \leqslant s \leqslant T} |\xi_{s/\varepsilon^\alpha}|^p\right) \leqslant \left(E_{\mu_{\text{inv}}}\left(\varepsilon^\alpha \sup_{0 \leqslant s \leqslant T} |\xi_{s/\varepsilon^\alpha}|^{\frac{\alpha}{\beta}p}\right)\right)^{\beta/\alpha} \underset{\varepsilon\downarrow 0}{\longrightarrow} 0. \qquad \square$$

References

Allaire, G., 1992. Homogenization and two-scale convergence. SIAM J. Math. Anal. 23, 1482–1518.

Avellaneda, M., Majda, A.J., 1991. An integral representation and bounds on the effective diffusivity in passive advection with laminar and turbulent flows. Comm. Math. Phy. 138, 339–391.

Avellaneda, M., Majda, A.J., 1994. Simple examples with features of renormalization for turbulent transport. Philos. Trans. Roy. Soc. London Ser. A 346, 205–233.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Bouc, R., Pardoux, É., 1984. Asymptotic analysis of PDEs with wide-band noise disturbances, and expansion of the moments. Stochastic Anal. Appl. 2 (4), 369–422.

Bourgeat, A., Luckhaus, S., Mikelić, A., 1994. A rigorous result for a double porosity model of immiscible two-phase flow. Comptes Rendus à l'Académie des Sciences 320 (I), 1289–1294.

Fannjiang, A., Papanicolaou, G., 1996. Diffusion in turbulence. Probab. Theory Related Fields 105, 279–334.

Fannjiang, A., Papanicolaou, G., 1997. Convection enhanced diffusion for random flows. J. Statist. Phys. 88, 1033–1076.

Gilbarg, D., Trudinger, N., 1994. Elliptic Partial Differential Equations of Second Order. Springer, Berlin.

Hornung, U. (Ed.), 1997. Homogenization and Porous Media, Volume 6 of IAM. Springer, Berlin.

Jikov, V.V., Kozlov, S.M., Oleinik, O.A., 1994. Homogenization of Differential Operators and Integral Functionals. Springer, Berlin.

Kleptsyna, M.L., Piatnitski, A.L., 1997. Homogenization of random parabolic operators. GAKUTO International Series, Mathematical Sciences Applications 9, 241–255.

Kozlov, S.M., 1980. Averaging of random operators. Math. USSR-Sb. 37 (2), 167–180.

Liptser, R.Sh., Shiryaev, A.N., 1989. Theory of Martingales. Mathematics and its Applications (Soviet Series). Kluwer, Dordrecht.

Papanicolaou, G.C., Varadhan, S.R.S., 1982. Diffusions with random coefficients. In: Kallianpur, G., Krishnaiah, P.R., Ghosh, J.K. (Eds.), Statistics and Probability: An Essay in Honor of C.R. Rao. North-Holland, Amsterdam, pp. 547–552.

Pardoux, É., Veretennikov, A.Yu., 2001. On Poisson equation and diffusion approximation. Part I. Ann. Probab., to appear.

Trudinger, N.S., 1968. Pointwise estimates and quasilinear parabolic equations. Comm. Pure Appl. Math. 21, 205–226.

Veretennikov, A.Yu., 1997. On polynomial mixing bounds for stochastic differential equations. Stochastic Process. Appl. 70 (1), 115–127.

Viot, M., 1976. Solutions Faibles d'Equations aux Dérivées Partielles Stochastiques Non Linéaires. Thèse, Université Paris VI.