多尺度随机系统的渐近行为

李楠楠　解龙杰 *

(江苏师范大学, 数学与统计学院, 徐州, 221000)

摘　要　本文总结具有不规则系数的多尺度随机系统渐近行为的最新进展, 介绍平均化原理、正态偏差及扩散逼近; 特别地, 对于由布朗噪音驱动的经典随机 Langevin 方程与由 Lévy 噪音驱动的随机 Langevin 方程, 介绍其 Smoluchowski-Kramers 逼近. 与经典的布朗噪音驱动的方程不同, 即便是摩擦常数依赖于物体的位置, 在由 Lévy 噪音驱动的随机 Langevin 方程的极限方程中仍不会出现由噪音诱导的新漂移项.

关键词　扩散逼近　Smoluchowski-Kramers 逼近　正态偏差　多尺度系统　平均化原理

Asymptotic Behavior of the Multiscale Stochastic Systems

Li Nannan　Xie Longjie*

(School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, Jiangsu 221000, China)

Abstract　This paper summarizes the recent progress on the limiting behavior of multiscale stochastic systems with irregular coefficients, focuses on presenting the averaging principle, the normal deviation and the diffusion approximation, in particular, the Smoluchowski-Kramers approximations for the classical stochastic Langevin equation driven by Brownian noise and the Langevin equation driven by Lévy noise. Unlike the classical equation driven by Brownian noise, there is no noise induced drift in the limit equation of the Langevin euqation driven by Lévy noise even if the friction is state dependent.

Key words　Diffusion approximation　Smoluchowski-Kramers approximation　Normal deviations　Multiscale system　Averaging principle

doi: 10.3969/j.issn.1006-8074.2022.02.005

1　引言

自然界中的许多现象都具有多尺度特征或多尺度效应, 并且, 人们对这些现象的观察及测量也常在不同尺度（分辨率）下进行 (参见 [3, 13, 16, 21, 24, 26, 30]). 数学上, 多尺度系统能够很好地将这些现象的本质特征反映出来. 近年来, 多尺度问题及相关理论被广泛应用于海洋大气、复合材料、生命科学和金融等领域, 其研究热度不断上升.

多尺度系统往往比较复杂, 处理起来也相对困难. 平均化方法是处理多尺度问题的一个强有力的工具, 它通过构造一个“平均化方程”来简化原系统, 使化简后的方程不再涉及尺度的分离. 本文首先介绍快慢随机系统的平均化原理的强、弱收敛, 特别侧重于噪音所带来的正则化作用. 然后,

国家自然科学基金项目 (No. 12071186) 和江苏省研究生科研与实践创新项目 (No. KYCX21-2612) 资助

通信作者: 解龙杰, 教授, 博士; E-mail: longjiexie@jsnu.edu.cn

收稿日期: 2021 年 11 月 19 日

介绍原系统与其平均化方程的波动估计. 这两个方面分别对应于或类似于概率论中的大数定律与中心极限定理. 而对于更一般的多尺度随机系统, 我们将介绍其扩散逼近, 这一问题与偏微分方程中的同质化有紧密联系. 最后, 我们考虑一类特殊的系统: 随机 Langevin 方程, 分别给出由布朗噪音驱动和 Lévy 噪音驱动的随机 Langevin 方程的 Smoluchowski-Kramers 逼近.

2 平均化原理

考虑如下的随机微分方程:

$$\mathrm{d}Y_t^\varepsilon = F(X_{t/\varepsilon}, Y_t^\varepsilon)\mathrm{d}t + \mathrm{d}W_t, \quad Y_0^\varepsilon = y \in \mathbb{R}^d, \tag{2.1}$$

其中, $d \geqslant 1$ 为空间维数, $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ 为可测函数（称为漂移系数）, W_t 为定义在某概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的 d 维标准布朗运动, $X = (X_t)_{t\geqslant 0}$ 为一个给定的遍历马氏过程, 参数 $0 < \varepsilon \ll 1$ 代表了时间尺度的分离: Y_t^ε 以正常的时间尺度 t 变化, 而 X 以时间尺度 t/ε 变化. 因此, 在方程 (2.1) 中, Y_t^ε 通常称为慢变量, 用来代表实际中可以观测到的量, 也是人们更关心的量; 而 X 称为快变量, 通常用来描述系统所处的快速变化的外界环境或者其他相关的影响因素.

因为方程 (2.1) 中涉及到两个时间尺度, 所以直接研究 (2.1) 比较复杂. 我们希望寻找一个简单的方程来近似地代替该方程. 由于在大多实际问题中, $\varepsilon > 0$ 非常小, 因此问题即转化为研究该系统当 $\varepsilon \to 0$ 时的渐近行为. 对于方程 (2.1), 可以比较直观地推测出其 $\varepsilon \to 0$ 时的极限: 假设 X 的唯一不变测度为 $\mu(\mathrm{d}x)$, 则形式上, 当 $\varepsilon \to 0$ 时, $X_{t/\varepsilon}$ 将收敛到其不变测度 $\mu(\mathrm{d}x)$, 代入到方程 (2.1) 中, 即可得 Y_t^ε 将收敛于 $\bar{Y}_t$, 其中 $\bar{Y}_t$ 满足

$$\mathrm{d}\bar{Y}_t = \bar{F}(\bar{Y}_t)\mathrm{d}t + \mathrm{d}W_t, \quad \bar{Y}_0 = y \in \mathbb{R}^d, \tag{2.2}$$

这里新的漂移系数为

$$\bar{F}(y) := \int_{\mathbb{R}^d} F(x, y)\mu(\mathrm{d}x). \tag{2.3}$$

这一结果称为平均化原理, 方程 (2.2) 称为 (2.1) 的平均化方程, $\bar{F}$ 称为平均化系数.

一般地, 平均化原理的证明需要假设原方程的系数满足一定的正则性条件. 经典的平均化原理由前苏联数学家 Bogoliubov (参见 [6]) 对确定性的常微分方程首先建立. 这一理论后来被 Khasminskii [17] 推广到随机微分方程. 在随机的情形, 从概率的角度来说, Y_t^ε 可以按照多种方式收敛到 $\bar{Y}_t$. 这里我们主要介绍两种:

(i) 强收敛 (p-阶矩收敛): 对任意的 $p \geqslant 1$,

$$\lim_{\varepsilon\to 0} \mathbb{E}|Y_t^\varepsilon - \bar{Y}_t|^p = 0;$$

(ii) 弱收敛 (依分布收敛): 对任意的 $\varphi \in C_b(\mathbb{R}^d)$,

$$\lim_{\varepsilon\to 0} \left|\mathbb{E}\varphi(Y_t^\varepsilon) - \mathbb{E}\varphi(\bar{Y}_t)\right| = 0.$$

从结果上看, 强收敛可以直接推出弱收敛. 但强收敛往往需要更强的假设条件, 并且强收敛的收敛速度也慢于弱收敛. 在下面的定理中我们会具体介绍.

方程 (2.1) 相对来说仍然比较简单: 其快变量 X 不依赖于慢变量 Y. 然而, 在描述一个复杂系统时, 快、慢变量会完全偶合在一起. 因此, 我们需要考虑如下更一般的随机微分方程:

$$\begin{cases}\mathrm{d}X_t^\varepsilon=\varepsilon^{-1}b(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+\varepsilon^{-1/2}\sigma(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}W_t^1, & X_0^\varepsilon=x\in\mathbb{R}^{d_1},\\ \mathrm{d}Y_t^\varepsilon=F(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+G(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}W_t^2, & Y_0^\varepsilon=y\in\mathbb{R}^{d_2}.\end{cases}\tag{2.4}$$

相比于 (2.1), 方程 (2.4) 由可乘噪音驱动, 并且其快变量 X 会依赖于慢变量 Y. 此时, 仍然像之前那样直接推测其平均化方程不再可行. 直观上, 推导方程 (2.4) 的平均化方程可分为两步:

(1) 当观察快变量时, 慢变量几乎保持不动;

(2) 当观察慢变量时, 快变量几乎已经达到其平稳状态.

具体地, 我们先来看快变量. 观察快变量合适的方式是对其做时间变换. 令 $\tilde{X}_t^\varepsilon:=X_{t\varepsilon}^\varepsilon$(以时间尺度 $t\varepsilon$ 来观察快变量), 则可以验证 $\tilde{X}_t^\varepsilon$ 满足如下方程:

$$\mathrm{d}\tilde{X}_t^\varepsilon=b(\tilde{X}_t^\varepsilon,Y_{\varepsilon t}^\varepsilon)\mathrm{d}t+\sigma(\tilde{X}_t^\varepsilon,Y_{\varepsilon t}^\varepsilon)\mathrm{d}\tilde{W}_t^1,\quad \tilde{X}_0^\varepsilon=x\in\mathbb{R}^{d_1},$$

其中, $\tilde{W}_t^1:=\varepsilon^{-1/2}W_{\varepsilon t}^1$ 为新的标准布朗运动. 因为我们关心的是 $\varepsilon\to0$ 时整个系统的行为, 而当 ε 很小时, $Y_{\varepsilon t}^\varepsilon$ 会非常接近其初值 y. 所以我们自然地需要考虑满足如下凝固方程的辅助过程 X_t^y: 对固定的 $y\in\mathbb{R}^{d_2}$,

$$\mathrm{d}X_t^y=b(X_t^y,y)\mathrm{d}t+\sigma(X_t^y,y)\mathrm{d}W_t^1,\quad X_0^y=x\in\mathbb{R}^{d_1}.\tag{2.5}$$

尺度变换后的过程 $\tilde{X}_t^\varepsilon$ 将渐近同分布于 X_t^y. 在一定的假设条件下, 过程 X_t^y 存在唯一的不变测度 $\mu^y(\mathrm{d}x)$ (这里的 $\mu^y(\mathrm{d}x)$ 即相当于方程 (2.1) 中的 $\mu(\mathrm{d}x)$, 即快变量的平稳状态). 对慢方程中的快变量关于此测度取平均, 则可得慢变量 Y_t^ε 将收敛到 $\bar{Y}_t$, 其中, $\bar{Y}_t$ 满足如下的平均化方程:

$$\mathrm{d}\bar{Y}_t=\bar{F}(\bar{Y}_t)\mathrm{d}t+\bar{G}(\bar{Y}_t)\mathrm{d}W_t^2,\quad \bar{Y}_0=y\in\mathbb{R}^{d_2},\tag{2.6}$$

这里, 平均化漂移系数为

$$\bar{F}(y):=\int_{\mathbb{R}^{d_1}}F(x,y)\mu^y(\mathrm{d}x),\tag{2.7}$$

平均化的扩散系数为

$$\bar{G}(y):=\sqrt{\int_{\mathbb{R}^{d_1}}G(x,y)G(x,y)^*\mu^y(\mathrm{d}x)}.$$

注 1 需要指出的是, 尽管平均化方程 (2.2) 和 (2.6) 中平均化系数的形式类似, 但方程 (2.4) 的研究远比 (2.1) 复杂. 实际上, 对于强收敛, 通常需要证明平均化后的系数满足 Lipschitz 条件以保证平均化方程的强适定性. 对于 (2.3), 我们只需要假设原方程 (2.1) 中的漂移系数 $F(x,y)$ 关于 y 变量满足 Lipschitz 连续性的条件, 则容易验证 $\bar{F}$ 也满足 Lipschitz 条件. 但 (2.7) 中 $\bar{F}$ 的 Lipschitz 连续性的验证要困难的多: 其涉及到不变测度 $\mu^y(\mathrm{d}x)$ 关于参数 y 的正则性的研究.

目前, 随机系统的平均化原理已经得到非常深入地研究. 关于由布朗运动驱动的随机微分方程的平均化原理, 可参见 [11, 18, 20, 25, 35]; 关于由 Lévy 噪音驱动的随机微分方程的平均化原理, 可参见 [4, 31, 34]; 关于随机偏微分方程的情形, 可参见 [7, 10] 及其中的参考文献. 下面, 我们分别介

绍两个具体的平均化原理强、弱收敛的结果, 并给出其与经典结果的比较. 这里, 我们主要关心不规则系数的情形, 从而体现出噪音对系统的正则化作用. 在此之前, 我们做一些基本的假设:

(Hσ) 扩散系数 $a=\sigma\sigma^*$ 一致非退化, 即存在常数 $\lambda>1$, 使得对任意的 $x\in\mathbb{R}^{d_1}$ 和 $y\in\mathbb{R}^{d_2}$, 有

$$\lambda^{-1}|\xi|^2\leqslant a^{ij}(x,y)\xi_i\xi_j\leqslant\lambda|\xi|^2,\ \ \forall\xi\in\mathbb{R}^{d_1}.$$

(H$_G$) 扩散系数 $\mathcal{G}=GG^*$ 一致非退化, 即存在常数 $\lambda>1$, 使得对任意的 $x\in\mathbb{R}^{d_1}$ 和 $y\in\mathbb{R}^{d_2}$, 有

$$\lambda^{-1}|\xi|^2\leqslant \mathcal{G}^{ij}(x,y)\xi_i\xi_j\leqslant\lambda|\xi|^2,\ \ \forall\xi\in\mathbb{R}^{d_2}.$$

(Hb) 漂移系数 b 满足如下的 Lyapunov 条件:

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

假设 **(Hσ)** 和 **(Hb)** 保证凝固方程 (2.5) 存在唯一的不变测度 (参见 [29, 36]). 同时, 假设 **(H$_G$)** 将保证我们可以在不规则系数的条件下研究方程 (2.4) 的平均化原理.

我们有如下结果, 参见 [33, Theorem 2.1]. 为简单起见, 以下所有结果的叙述中我们都假设系数有界.

定理 1 (强收敛) 假设 **(Hσ)**, **(H$_G$)**, **(Hb)** 成立, 并且

$$G(x,y)\equiv G(y).\tag{2.8}$$

如果 $\sigma\in C_b^{1,1}$, $G\in C_b^1$ 且 $b,F\in C_b^{\delta,\alpha}$, 其中 $\delta,\alpha>0$, 则对任意的 $T>0$, 有

$$\sup_{t\in[0,T]}\mathbb{E}|Y_t^\varepsilon-\bar{Y}_t|^2\leqslant C_T\,\varepsilon^{\alpha\wedge1},\tag{2.9}$$

其中, $C_T>0$ 为不依赖于 δ 和 ε 的常数.

注 2 (i) 假设 (2.8) 是必要的, 因为当慢方程中的扩散系数 G 依赖于 x(快变量) 时, 强收敛不一定成立 (参见 [22]).

(ii) 虽然我们只假设了方程 (2.4) 中的漂移系数为 Hölder 连续的, 但在 **(Hσ)** 和 **(H$_G$)** 的条件下, 方程 (2.4) 仍存在唯一的强解. 同时, 也可以证明 $\bar{F}\in C_b^\alpha$, 从而平均化方程 (2.6) 同样存在唯一的强解, 这体现了噪音的正则化作用. 关于不规则系数下随机微分方程强适定性的更多介绍, 可参见 [40, 41].

(iii) 经典的平均化原理的结果都需要假设系数满足 Lipschitz 连续或者局部 Lipschitz 连续的条件, 并且, 在系数足够正则的情况下, 强收敛的最优收敛速度为 $\sqrt{\varepsilon}$. 而估计 (2.9) 表明, 当方程中的系数关于 y 变量（慢变量）为 α-Hölder 连续时, 强收敛的收敛速度为 $\varepsilon^{\alpha/2}$. 特别地, 其收敛速度不依赖于方程中系数关于快变量的正则性. 这与直观吻合: 在取极限的过程中, 快变量被完全平均化, 其不再出现在极限方程中.

关于方程 (2.4) 平均化原理的弱收敛, 我们有如下结果 (参见 [32, Theorem 2.3]).

定理 2 (弱收敛) 假设 **(Hσ)**, **(H$_G$)**, **(Hb)** 成立. 如果 $\sigma, b, F, G \in C_b^{\delta,\alpha}$, 其中 $\delta, \alpha > 0$, 则对任意的 $T > 0$ 及 $\varphi \in C_b^{2+\alpha}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]} \Big|\mathbb{E}[\varphi(Y_t^\varepsilon)] - \mathbb{E}[\varphi(Y_t)]\Big| \leqslant C_T\, \varepsilon^{(\alpha/2)\wedge 1}, \tag{2.10}$$

其中, $C_T > 0$ 为不依赖于 δ 的常数.

注 3 (i) 对于平均化原理的弱收敛, 我们只需要保证方程的弱适定性. 由于噪音的非退化性, 在上述系数 Hölder 连续的假设条件下, 原方程 (2.4) 及平均化方程 (2.1) 存在唯一的弱解.

(ii) 估计 (2.10) 表明, 当原方程的系数关于慢变量为 α-Hölder 连续时, 弱收敛的速度为 $\alpha/2$. 特别地, 当 $\alpha = 2$ 时, 得最优的弱收敛速度为 1. 与强收敛的情形类似, 上述的弱收敛也不依赖于系统关于快变量的正则性.

3 正态偏差

随机系统的平均化原理可以理解成一种泛函大数定律. 其意义在于, 我们可以利用平均化方程 (2.6) 近似地代替原方程 (2.4), 而平均化方程不涉及多尺度变量, 从而更加简单. 但极限方程 (2.6) 只有在 $\varepsilon \to 0$ 时才成立. 在实际应用中, 虽然 ε 很小, 但其永远大于 0, 从而, Y_t^ε 会与其平均 $\bar{Y}_t$ 有一定的偏差. 要研究其波动情况, 自然地需要研究中心极限定理, 即考虑标准化的过程

$$Z_t^\varepsilon := \frac{Y_t^\varepsilon - \bar{Y}_t}{\sqrt{\varepsilon}}$$

当 $\varepsilon \to 0$ 时的渐近行为.

为了更清楚地了解问题的困难所在, 我们先考虑 (2.4) 中 $G = I$ (单位矩阵) 的情况. 此时,

$$Y_t^\varepsilon - \bar{Y}_t = \int_0^t F(X_s^\varepsilon, Y_s^\varepsilon) - \bar{F}(\bar{Y}_s)\mathrm{d}s.$$

从而 Z_t^ε 可以进一步写为

$$Z_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}}\int_0^t F(X_s^\varepsilon, Y_s^\varepsilon) - \bar{F}(\bar{Y}_s)\mathrm{d}s.$$

与平均化原理不同, Z_t^ε 的极限并没有特别直观的推测. 实际上, 可以证明当 $\varepsilon \to 0$ 时, 过程 Z_t^ε 将依分布收敛到满足如下线性随机微分方程的 Ornstein-Uhlenbeck 型过程 $\bar{Z}_t$:

$$\mathrm{d}\bar{Z}_t = \nabla_y \bar{F}(\bar{Y}_t)\bar{Z}_t\mathrm{d}t + \zeta(\bar{Y}_t)\mathrm{d}\tilde{W}_t,$$

其中, $\tilde{W}_t$ 为一新的标准布朗运动, 而新的扩散系数为

$$\zeta(y) := \sqrt{\int_0^\infty\!\!\int_{\mathbb{R}^{d_1}} \mathbb{E}\big[F(X_t^y(x), y) - \bar{F}(y)\big]\big[F(x,y) - \bar{F}(y)\big]^* \mu^y(\mathrm{d}x)\mathrm{d}t}. \tag{3.11}$$

对于一般的随机系统, 我们有如下结果 (参见 [33, Theorem 2.3]). 关于多尺度随机偏微分方程正态偏差的研究, 可参见 [8, 37].

定理 3 (正态偏差) 假设 **(Hσ)**, **(H$_G$)**, **(Hb)** 及 (2.8) 成立. 如果 $\sigma, b, F, G \in C_b^{\delta,1+\vartheta}$, 其中 $\delta, \vartheta > 0$, 则对任意的 $T > 0$ 及 $\varphi \in C_b^{2+\vartheta}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]} \Big|\mathbb{E}[\varphi(Z_t^\varepsilon)] - \mathbb{E}[\varphi(\bar{Z}_t)]\Big| \leqslant C_T\, \varepsilon^{(\vartheta\wedge 1)/2}, \tag{3.12}$$

其中, $\bar{Z}_t$ 满足方程

$$\mathrm{d}\bar{Z}_t = \nabla_y \bar{F}(\bar{Y}_t)\bar{Z}_t \mathrm{d}t + \nabla G(\bar{Y}_t)\mathrm{d}W_t^2 + \zeta(\bar{Y}_t)\mathrm{d}\tilde{W}_t,$$

$C_T > 0$ 为不依赖于 δ 和 ε 的常数.

注 4 特别地, 估计 (3.12) 说明, 当 $\vartheta = 1$ 时中心极限定理的最优收敛速度为 $\sqrt{\varepsilon}$.

形式上, 由正态偏差的结果可得到 Y_t^ε 有如下的渐近展开:

$$Y_t^\varepsilon \overset{\mathcal{D}}{\approx} \bar{Y}_t + \sqrt{\varepsilon}\bar{Z}_t,$$

这里, $\overset{\mathcal{D}}{\approx}$ 表示依概率分布的渐近相等. 上述渐近展开在物理中被称为 Van Kampen 逼近 (参见 [1]), 可用于对原方程 (2.4) 提供更高阶的近似逼近.

一个有趣的问题是: 能否找到一个分布不依赖于 ε 的过程 $\mathcal{Z}_t^\varepsilon$, 使得 $\mathcal{Z}_t^\varepsilon$ 以 ε 阶强收敛到 Y_t^ε, 即对任意的 $p \geqslant 1$, 存在常数 $C_p > 0$, 使得

$$\mathbb{E}|Y_t^\varepsilon - \mathcal{Z}_t^\varepsilon|^p \leqslant C_p\, \varepsilon^p?$$

这一问题至今没有得到解答. 需要指出的是, 取 $\mathcal{Z}_t^\varepsilon := \bar{Y}_t + \sqrt{\varepsilon}\bar{Z}_t$ 并不可行. 因为上述的正态偏差结果仅仅是依分布收敛, 而强收敛不成立. Hasselmann 从物理的直观角度出发, 猜测了如下的逼近过程:

$$\mathrm{d}\tilde{Y}_t^\varepsilon = \bar{F}(\tilde{Y}_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\,\zeta(\tilde{Y}_t^\varepsilon)\mathrm{d}W_t,$$

其中, 扩散系数 ζ 由 (3.11) 定义. 但理论上一直没有能够给出证明. 实际上, 直到 2004 年, Bakhtin 和 Kiffer [2] 才证明了

$$\mathbb{E}|Y_t^\varepsilon - \tilde{Y}_t^\varepsilon| \leqslant C_1 \varepsilon^{(\delta+1)/2},$$

其中, $0 < \delta < (18 + 8d)^{-1}$. 虽然得到了高于 $\sqrt{\varepsilon}$ 的收敛速度, 但这个 δ 仍不是最优的.

4 扩散逼近

考虑如下多尺度随机微分方程:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \alpha_\varepsilon^{-2} b(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \beta_\varepsilon^{-1} c(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \alpha_\varepsilon^{-1}\sigma(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}W_t^1, \\ \mathrm{d}Y_t^\varepsilon = F(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \gamma_\varepsilon^{-1} H(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + G(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}W_t^2, \\ X_0^\varepsilon = x \in \mathbb{R}^{d_1}, \quad Y_0^\varepsilon = y \in \mathbb{R}^{d_2}, \end{cases} \tag{4.13}$$

其中, 当 $\varepsilon \to 0$ 时, 参数 $\alpha_\varepsilon, \beta_\varepsilon, \gamma_\varepsilon \to 0$. 特别地, 当 $c = H \equiv 0$, $\alpha_\varepsilon = \sqrt{\varepsilon}$ 时, 方程 (4.13) 即为快慢方程 (2.4). 相比于 (2.4), 方程 (4.13) 有如下两个主要特点:

(i) 快方程中存在两个不同的时间尺度, 分别由 α_ε 和 β_ε 刻画;
(ii) 即便在慢方程中, 也存在着一个快速变化的项, 由 $\gamma_\varepsilon^{-1}H$ 刻画.

方程 (4.13) 的极限行为与偏微分方程中的同质化理论有密切联系 (参见 [14, 19, 23]). 当 $\alpha_\varepsilon=\beta_\varepsilon=\gamma_\varepsilon=\sqrt{\varepsilon}$, 且 $(X_t^\varepsilon,Y_t^\varepsilon)$ 的状态空间为紧集时, 方程 (4.13) 首先由 Papanicolaou, Stroock 和 Varadhan[27] 进行了研究. 而当 $c\equiv 0$, $\alpha_\varepsilon=\gamma_\varepsilon=\sqrt{\varepsilon}$, $(X_t^\varepsilon,Y_t^\varepsilon)$ 的状态空间为全空间时, Pardoux 和 Veretennikov 在一系列文章[28, 29] 中对 (4.13) 的极限进行了研究. 需要指出的是, 在全空间上 (4.13) 的研究要比紧集上困难的多, 其关键在于求解全空间上带参数的 Poisson 方程, 并研究解的正则性.

对于方程 (4.13), 根据时间尺度收敛到 0 的速度不同, 其极限方程也不同. 具体需要分以下四种情况:

$$\begin{cases} \text{情况 1} \quad \lim\limits_{\varepsilon\to 0}\dfrac{\alpha_\varepsilon}{\gamma_\varepsilon}=0 \quad \text{且} \quad \lim\limits_{\varepsilon\to 0}\dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}=0; \\ \text{情况 2} \quad \lim\limits_{\varepsilon\to 0}\dfrac{\alpha_\varepsilon}{\gamma_\varepsilon}=0 \quad \text{且} \quad \alpha_\varepsilon^2=\beta_\varepsilon\gamma_\varepsilon; \\ \text{情况 3} \quad \alpha_\varepsilon=\gamma_\varepsilon \quad \text{且} \quad \lim\limits_{\varepsilon\to 0}\dfrac{\alpha_\varepsilon}{\beta_\varepsilon}=0; \\ \text{情况 4} \quad \alpha_\varepsilon=\beta_\varepsilon=\gamma_\varepsilon. \end{cases} \tag{4.14}$$

当 α_ε 和 α_ε^2 收敛到 0 的速度分别比 γ_ε 和 $\beta_\varepsilon\gamma_\varepsilon$ 快时 (情况 1), 方程 (4.13) 与方程 (2.4) 的平均化方程完全一致, 也就是说, c 和 H 变化的速度不够快, 其作用并没有体现在极限方程中; 当 α_ε 收敛到 0 的速度比 γ_ε 快而 α_ε^2 和 $\beta_\varepsilon\gamma_\varepsilon$ 同阶时 (情况 2), 系数 c 的同质化作用将体现在极限方程中; 当 α_ε 和 γ_ε 同阶而 α_ε 收敛到 0 的速度比 β_ε 快时 (情况 3), 系数 H 的同质化作用将得以体现; 最后, 当所有的参数都同阶时 (情况 4), c 和 H 的同质化作用将同时体现在极限方程中.

为了介绍 (4.13) 的极限方程, 我们需要引入如下的 Poisson 方程:

$$\mathcal{L}_0(x,y)\Phi(x,y)=-H(x,y), \quad x\in\mathbb{R}^{d_1}, \tag{4.15}$$

其中, $y\in\mathbb{R}^{d_2}$ 为参数,

$$\mathcal{L}_0:=\mathcal{L}_0(x,y):=\sum_{i,j=1}^{d_1}a^{ij}(x,y)\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^{d_1}b^i(x,y)\frac{\partial}{\partial x_i},$$

$a(x,y):=\sigma\sigma^*(x,y)$. 需要指出的是, 虽然紧集上的 Poisson 方程已经有了很深入的研究, 但全空间上带参数的 Poisson 方程 (4.15) 的研究却十分困难 (见 [32]). 特别地, 为了保证 (4.15) 解的存在唯一性, 需要假设 H 满足如下条件:

$$\int_{\mathbb{R}^{d_1}}H(x,y)\mu^y(\mathrm{d}x)=0, \quad \forall y\in\mathbb{R}^{d_2}. \tag{4.16}$$

在一定的正则性条件下, 方程 (4.15) 存在唯一的解 $\Phi(x,y)$. 对应于 (4.14) 中的情况 1–4, 我们分别

定义平均化的漂移系数

$$\begin{aligned}
\hat{F}_1(y) &:= \int_{\mathbb{R}^{d_1}} F(x,y)\mu^y(\mathrm{d}x);\\
\hat{F}_2(y) &:= \int_{\mathbb{R}^{d_1}} \big[F(x,y)+c(x,y)\cdot\nabla_x\Phi(x,y)\big]\mu^y(\mathrm{d}x);\\
\hat{F}_3(y) &:= \int_{\mathbb{R}^{d_1}} \big[F(x,y)+H(t,x,y)\cdot\nabla_y\Phi(x,y)\big]\mu^y(\mathrm{d}x);\\
\hat{F}_4(y) &:= \int_{\mathbb{R}^{d_1}} \big[F(x,y)+c(x,y)\cdot\nabla_x\Phi(x,y)\\
&\qquad\qquad + H(x,y)\cdot\nabla_y\Phi(x,y)\big]\mu^y(\mathrm{d}x),
\end{aligned}$$

以及平均化的扩散系数

$$\begin{aligned}
\hat{G}_1(y)=\hat{G}_2(y) &:= \sqrt{\int_{\mathbb{R}^{d_1}} GG^*(x,y)\mu^y(\mathrm{d}x)};\\
\hat{G}_3(y)=\hat{G}_4(y) &:= \sqrt{\int_{\mathbb{R}^{d_1}} \Big[GG^*(x,y)+H(x,y)\Phi^*(x,y)\Big]\mu^y(\mathrm{d}x)}.
\end{aligned}$$

方程 (4.13) 的极限 $\hat{Y}_t^k(k=1,\cdots,4)$ 将分别满足如下的随机微分方程:

$$\mathrm{d}\hat{Y}_t^k=\hat{F}_k(t,\hat{Y}_t^k)\mathrm{d}t+\hat{G}_k(t,\hat{Y}_t^k)\mathrm{d}W_t^2,\quad \hat{Y}_0^k=y. \tag{4.17}$$

我们有如下主要结果 (参见 [32, Theorem 2.3]). 关于随机偏微分方程的扩散逼近, 可参见 [38].

定理 4 假设 **(H$_\sigma$), (H$_b$), (H$_G$)** 及 (4.16) 成立, $T>0$ 且 $\delta\in(0,1]$.

(i) 情况 1 如果 $b,\sigma,F,H,G\in C_b^{\delta,\vartheta}$, 其中 $\vartheta\in(0,2]$, $c\in L^\infty$ 并且 $\lim_{\varepsilon\to 0}\alpha_\varepsilon^\vartheta/\gamma_\varepsilon=0$, 则对任意的 $\varphi\in C_b^{2+\vartheta}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Y_t^\varepsilon)]-\mathbb{E}[\varphi(\hat{Y}_t^1)]\Big|\leqslant C_T\left(\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon^2}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\right);$$

(ii) 情况 2 如果 $b,\sigma,F,H,G,c\in C_b^{\delta,\vartheta}$, 其中 $\vartheta\in(0,2]$ 并且 $\lim_{\varepsilon\to 0}\alpha_\varepsilon^\vartheta/\gamma_\varepsilon=0$, 则对任意的 $\varphi\in C_b^{2+\vartheta}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Y_t^\varepsilon)]-\mathbb{E}[\varphi(\hat{Y}_t^2)]\Big|\leqslant C_T\left(\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon^2}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\right);$$

(iii) 情况 3 如果 $b,\sigma,H\in C_b^{\delta,1+\vartheta}$, $F,G\in C_p^{\delta,\vartheta}$ $\vartheta\in(0,1]$, 其中 $\vartheta\in(0,1]$ 且 $c\in L^\infty$, 则对任意的 $\varphi\in C_b^{2+\vartheta}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Y_t^\varepsilon)]-\mathbb{E}[\varphi(\hat{Y}_t^3)]\Big|\leqslant C_T\left(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon}{\beta_\varepsilon}\right);$$

(iv) 情况 4 如果 $b,\sigma,H\in C_b^{\delta,1+\vartheta}$ 且 $F,G,c\in C_p^{\delta,\vartheta}$, 其中 $\vartheta\in(0,1]$, 则对任意的 $\varphi\in C_b^{2+\vartheta}(\mathbb{R}^{d_2})$, 有

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Y_t^\varepsilon)]-\mathbb{E}[\varphi(\hat{Y}_t^4)]\Big|\leqslant C_T\,\alpha_\varepsilon^\vartheta,$$

其中, 对 $k=1,\cdots,4$, $\hat{Y}_t^k$ 为随机微分方程 (4.17) 的唯一弱解, $C_T>0$ 为不依赖于 δ,ε 的常数.

注 5 上述结论中, 对应于情况 3 和 4 的结论表明: 即使原系统中 $G \equiv 0$ (慢方程中没有噪音项), 其极限方程中仍然会出现可乘的布朗噪音项. 这主要是由慢方程中快速变化的 H 项的同质化导致的.

5 Smoluchowski-Kramers 逼近

5.1 布朗噪音驱动

考虑一个质量为 ε 的小粒子在受外力、正比于速度的摩擦力及噪音作用下的运动. 令 X_t^ε 表示 t 时刻粒子的位置, 则根据牛顿第二定律, X_t^ε 满足如下的随机 Langevin 方程 (参见 [15]):

$$\varepsilon \ddot{X}_t^\varepsilon = F(X_t^\varepsilon) - \gamma \dot{X}_t^\varepsilon + \sigma(X_t^\varepsilon)\dot{W}_t, \quad X_0^\varepsilon = x \in \mathbb{R}^d, \tag{5.18}$$

其中, $F(x): \mathbb{R}^d \to \mathbb{R}^d$ 表示外力项, $\dot{W}_t$ 为标准高斯白噪声, 矩阵函数 $\sigma(x)$ 代表噪声的强度, $\gamma > 0$ 为摩擦常数. 在一定的假设条件下, 当 ε 收敛到 0 时 (即相比于惯性, 摩擦力起主导作用), X_t^ε 将 $L^2(\Omega)$ 收敛到 X_t, 其中 X_t 满足

$$\mathrm{d}X_t = \frac{F(X_t)}{\gamma}\mathrm{d}t + \frac{\sigma(X_t)}{\gamma}\mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^d. \tag{5.19}$$

这一结果被称为 Smoluchowski-Kramers 逼近. 其意义在于: 对于小粒子, 我们可以用一阶方程 (5.19) 来近似代替二阶牛顿方程 (5.18) 来描述其运动.

然而, 当摩擦常数依赖于物体的位置时, 极限方程的形式将会不同. 具体地, 考虑如下随机系统:

$$\varepsilon \ddot{X}_t^\varepsilon = F(X_t^\varepsilon) - \gamma(X_t^\varepsilon) \dot{X}_t^\varepsilon + \sigma(X_t^\varepsilon)\dot{W}_t, \quad X_0^\varepsilon = x \in \mathbb{R}^d, \tag{5.20}$$

其中, $\gamma(x)$ 为 $d \times d$ 矩阵值函数. 此时, 当 $\varepsilon \to 0$ 时 X_t^ε 将 $L^2(\Omega)$ 收敛到 X_t, 其中 X_t 满足

$$\begin{aligned} \mathrm{d}X_t = {} & \Big[\gamma^{-1}(X_t)F(X_t) + S(X_t)\Big]\mathrm{d}t \\ & + \gamma^{-1}(X_t)\sigma(X_t)\mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^d, \end{aligned} \tag{5.21}$$

这里的 $S(x): \mathbb{R}^d \to \mathbb{R}^d$ 被称为由噪音诱导的漂移系数, 其第 i 个分量为

$$S_i(x) := \sum_{j,k=1}^d \frac{\partial}{\partial x_k}\Big[(\gamma^{-1})_{ij}(x)\Big] M_{jk}(x), \quad i = 1, \cdots, d,$$

而 $M(x)$ 满足 Lyapunov 方程

$$M(x)\gamma^*(x) + \gamma(x)M(x) = \Sigma(x) := \sigma(x)\sigma^*(x).$$

特别地, 如果 $\gamma(x)$ 和 $\Sigma(x)$ 可交换, 即 $\gamma(x)\Sigma(x) = \Sigma(x)\gamma(x)$, 则有 $M(x) = \gamma^{-1}(x)\Sigma(x)/2$.

关于 Langevin 方程的 Smoluchowski-Kramers 逼近, 目前已经有很多的研究结果, 关于随机微分方程的情形可参见 [5, 12, 15], 而关于随机偏微分方程的情形可见 [9] 及其中的参考文献. 实际上, Smoluchowski-Kramers 逼近的问题可以转化为扩散逼近问题. 如果我们定义速度过程

$$V_t^\varepsilon := \sqrt{\varepsilon}\dot{X}_t^\varepsilon,$$

则方程 (5.21) 可改写为:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} V_t^\varepsilon \mathrm{d}t, \\ \mathrm{d}V_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} F(X_t^\varepsilon)\mathrm{d}t - \frac{1}{\varepsilon}\gamma(X_t^\varepsilon)V_t^\varepsilon \mathrm{d}t + \frac{1}{\sqrt{\varepsilon}}\sigma(X_t^\varepsilon)\mathrm{d}W_t. \end{cases}$$

注意到这只是方程 (4.13) 的一个特殊情况. 但区别在于, 扩散逼近一般只能得到弱收敛 (依分布收敛), 而 Smoluchowski-Kramers 逼近的特殊之处在于, 这里可以得到强收敛. 我们假设:

(**A**) 矩阵 $\gamma(x)$ 和 $\sigma(x)$ 一致非退化, 即存在常数 $\lambda > 1$, 使得对任意的 $x \in \mathbb{R}^d$, 有

$$\lambda^{-1}|\xi|^2 \leqslant |H(x)\xi|^2 \wedge |\gamma(x)\xi|^2 \wedge |\sigma(x)\xi|^2 \leqslant \lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

我们介绍如下结果 (参见 [39]).

定理 5 假设 (**A**) 成立. 若 $\gamma \in C_b^{1+\beta}, F \in C_b^{\beta,\delta}$ 且 $\sigma \in C_b^1$, 其中 $0 < \delta, \beta \leqslant 1$, 则对任意的 $T > 0$ 及 $q \geqslant 1$, 有

$$\sup_{t\in[0,T]} \mathbb{E}|X_t^\varepsilon - X_t|^q \leqslant C_T\, \varepsilon^{q\beta/2},$$

其中, X_t 为随机微分方程 (5.21) 的唯一强解, $C_T > 0$ 为不依赖于 ε 和 δ 的常数.

5.2 Lévy 噪音驱动

下面, 我们考虑由 Lévy 噪音驱动的随机 Langevin 方程的 Smoluchowski-Kramers 逼近. 为简单起见, 我们考虑一维空间的情形:

$$\varepsilon \ddot{X}_t^\varepsilon = F(X_t^\varepsilon) - \gamma(X_t^\varepsilon)\dot{X}_t^\varepsilon + \sigma(X_t^\varepsilon)\dot{L}_t, \quad X_0^\varepsilon = x \in \mathbb{R}, \tag{5.22}$$

其中, L_t 为旋转不变的 α-稳定过程, $\alpha \in (1,2)$. 与之前类似, 我们引入速度过程

$$V_t^\varepsilon := \varepsilon^{1-1/\alpha}\dot{X}_t^\varepsilon.$$

则方程 (5.22) 可改写为

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \frac{1}{\varepsilon^{1-1/\alpha}} V_t^\varepsilon \mathrm{d}t, \\ \mathrm{d}V_t^\varepsilon = \frac{1}{\varepsilon^{1/\alpha}} F(X_t^\varepsilon)\mathrm{d}t - \frac{1}{\varepsilon}\gamma(X_t^\varepsilon)V_t^\varepsilon \mathrm{d}t + \frac{1}{\varepsilon^{1/\alpha}}\sigma(X_t^\varepsilon)\mathrm{d}L_t. \end{cases}$$

有趣的是, 与布朗噪音驱动的随机 Langevin 方程不同, 即使 (5.22) 中的摩擦常数依赖于物体的位置, 其极限方程中仍不会出现新的由噪音诱导的漂移系数.

我们将证明如下结果. 为了更清楚地介绍证明方法及关键技巧, 我们假设系数充分光滑且有界.

定理 6 假设 $\sigma(x), \gamma(x)$ 非退化. 则对任意的 $T > 0$, 有

$$\sup_{t\in[0,T]} \mathbb{E}|X_t^\varepsilon - X_t| \leqslant C_T\, \varepsilon^{2/\alpha - 1}, \tag{5.23}$$

其中, X_t 满足如下随机微分方程:

$$\mathrm{d}X_t = \gamma^{-1}(X_t)F(X_t)\mathrm{d}t + \gamma^{-1}(X_t)\sigma(X_t)\mathrm{d}L_t, \quad X_0 = x \in \mathbb{R}, \tag{5.24}$$

$C_T > 0$ 为不依赖于 ε 的常数.

证明 为简便起见, 我们定义算子

$$\mathcal{L}_0 := \mathcal{L}_0(x,v) := |\sigma(x)|^\alpha \cdot \Delta_v^{\alpha/2} - \gamma(x)v \cdot \nabla_v$$

及函数

$$\Psi(x,v) := \frac{v}{\gamma(x)}.$$

可以验证

$$\mathcal{L}_0(x,v)\Psi(x,v) = -v. \tag{5.25}$$

对 $\Psi(X_t^\varepsilon, V_t^\varepsilon)$ 用 Itô 公式, 有

$$\begin{aligned}\Psi(X_t^\varepsilon, V_t^\varepsilon) = \Psi(x,v) &+ \int_0^t \Big[\frac{1}{\varepsilon^{1-1/\alpha}} V_s^\varepsilon \cdot \nabla_x \Psi(X_s^\varepsilon, V_s^\varepsilon) \\ &+ \frac{1}{\varepsilon^{1/\alpha}} F(X_s^\varepsilon)\nabla_v \Psi(X_s^\varepsilon, V_s^\varepsilon) + \frac{1}{\varepsilon}\mathcal{L}_0 \Psi(X_s^\varepsilon, V_s^\varepsilon)\Big]\mathrm{d}s \\ &+ \int_0^t \int_{\mathbb{R}^d} \big[\Psi(X_s^\varepsilon, V_s^\varepsilon + \varepsilon^{-1/\alpha}\sigma(X_s^\varepsilon)z) - \Psi(X_s^\varepsilon, V_s^\varepsilon)\big]\tilde{N}(\mathrm{d}s, \mathrm{d}z).\end{aligned}$$

上式两边同时乘以 $\varepsilon^{1/\alpha}$, 并利用公式 (5.25) 可得

$$\begin{aligned}\frac{1}{\varepsilon^{1-1/\alpha}} \int_0^t V_s^\varepsilon \mathrm{d}s &= \varepsilon^{1/\alpha}[\Psi(x,v) - \Psi(X_t^\varepsilon, V_t^\varepsilon)] \\ &+ \int_0^t F(X_s^\varepsilon)\nabla_v \Psi(X_s^\varepsilon, V_s^\varepsilon)\mathrm{d}s + \varepsilon^{2/\alpha-1}\int_0^t V_s^\varepsilon \cdot \nabla_x \Psi(X_s^\varepsilon, V_s^\varepsilon)\mathrm{d}s \\ &+ \varepsilon^{1/\alpha}\int_0^t \int_{\mathbb{R}^d} \big[\Psi(X_s^\varepsilon, V_s^\varepsilon + \varepsilon^{-1/\alpha}\sigma(X_s^\varepsilon)z) - \Psi(X_s^\varepsilon, V_s^\varepsilon)\big]\tilde{N}(\mathrm{d}s, \mathrm{d}z).\end{aligned}$$

注意到, 由 $\Psi(x,v)$ 的定义我们有

$$F(x)\nabla_v \Psi(x,v) = \frac{F(x)}{\gamma(x)}, \quad v \cdot \nabla_x \Psi(x,v) = \frac{v^2}{\gamma^2(x)}\gamma'(x),$$

且

$$\varepsilon^{1/\alpha}\big[\Psi(x, v + \varepsilon^{-1/\alpha}\sigma(x)z) - \Psi(x,v)\big] = \frac{\sigma(x)}{\gamma(x)} \cdot z.$$

从而, 进一步有

$$\begin{aligned}X_t^\varepsilon = \varepsilon^{1/\alpha}[\Psi(x,v) - \Psi(X_t^\varepsilon, V_t^\varepsilon)] + \int_0^t \frac{F(X_s^\varepsilon)}{\gamma(X_s^\varepsilon)}\mathrm{d}s + \int_0^t \frac{\sigma(X_s^\varepsilon)}{\gamma(X_s^\varepsilon)}\mathrm{d}L_s \\ + \varepsilon^{2/\alpha-1}\int_0^t \frac{(V_s^\varepsilon)^2\gamma'(X_s^\varepsilon)}{\gamma^2(X_s^\varepsilon)}\mathrm{d}s.\end{aligned}$$

此式与 (5.24) 相减并取期望, 可得

$$\begin{aligned}\mathbb{E}|X_t^\varepsilon - X_t| \leqslant{} & \varepsilon^{1/\alpha}\mathbb{E}\big|\Psi(x,v) - \Psi(X_t^\varepsilon, V_t^\varepsilon)\big| + \mathbb{E}\left|\int_0^t \frac{F(X_s^\varepsilon)}{\gamma(X_s^\varepsilon)} - \frac{F(X_s)}{\gamma(X_s)}\mathrm{d}s\right| \\ & + \mathbb{E}\left|\int_0^t \frac{\sigma(X_s^\varepsilon)}{\gamma(X_s^\varepsilon)} - \frac{\sigma(X_s)}{\gamma(X_s)}\mathrm{d}L_s\right| + \varepsilon^{2/\alpha-1}\mathbb{E}\left|\int_0^t \frac{(V_s^\varepsilon)^2\gamma'(X_s^\varepsilon)}{\gamma^2(X_s^\varepsilon)}\mathrm{d}s\right|.\end{aligned}$$

令 $\tilde{V}_t^\varepsilon := V_{\varepsilon t}^\varepsilon$, 则可验证

$$\mathrm{d}\tilde{V}_t^\varepsilon = \varepsilon^{1-1/\alpha}F(X_{\varepsilon t}^\varepsilon)\mathrm{d}t - \gamma(X_{\varepsilon t}^\varepsilon)\tilde{V}_t^\varepsilon\mathrm{d}t + \sigma(X_{\varepsilon t}^\varepsilon)\mathrm{d}\tilde{L}_t,$$

其中, $\tilde{L}_t$ 为一新的旋转不变的 α-稳定过程. 由 [40, Lemma 7.1], 有

$$\sup_{t\geqslant 0,\varepsilon>0} \mathbb{E}|\tilde{V}_t^\varepsilon|^2 < \infty,$$

进一步有

$$\sup_{\varepsilon>0}\mathbb{E}|V_t^\varepsilon|^2 < \infty,$$

从而可得

$$\mathbb{E}|X_t^\varepsilon - X_t| \leqslant C_0\varepsilon^{2/\alpha-1} + C_1\mathbb{E}\left(\int_0^t |X_t^\varepsilon - X_t|\mathrm{d}s\right).$$

由 Gronwall 不等式, 即得 (5.23).

参考文献

[1] Arnold L. Hasselmann's Program Revisited: The Analysis of Stochasticity in Deterministic Climate Models[M]. Birkhäuser Basel, 2001.

[2] Bakhtin V, Kifer Y. Diffusion approximation for slow motion in fully coupled averaging[J]. Probability Theory & Related Fields, 2004, 129(2): 157-181.

[3] Ball, Karen, Kurtz, et al. Asymptotic analysis of multiscale approximations to reaction networks[J]. Annals of Applied Probability, 2006, 16(4): 1925-1961.

[4] Bao J, Yin G, Yuan C. Two-time-scale stochastic partial differential equations driven by α-stable noises: Averaging principles[J]. Bernoulli, 2016, 23(1): 645-669.

[5] Birrell J, Hottovy S, Volpe G, et al. Small Mass Limit of a Langevin Equation on a Manifold[J]. Annales Henri Poincare, 2016: 1-49.

[6] Bogoliubov N N, Mitropolsky Y A. Asymptotic Methods in the Theory of Nonlinear Oscillations[M]. New York: Gordon and Breach Inc, 1961.

[7] Bréhier, Charles-Edouard. Analysis of an HMM time-discretization scheme for a system of stochastic PDEs[J]. Siam Journal on Numerical Analysis, 2012, 51(2): 1185-1210.

[8] Cerrai S. Normal deviations from the averaged motion for some reaction-diffusion equations with fast oscillating perturbation[J]. Journal De Mathématiques Pures Et Appliquées, 2009, 91(6): 614-647.

[9] Cerrai S, Freidlin M. On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom[J]. Probability Theory & Related Fields, 2005, 135(3): 363–394.

[10] Cerrai S, Freidlin M. Averaging principle for a class of stochastic reaction-diffusion equations[J]. Probability Theory & Related Fields, 2009, 144(s 1–2): 137–177.

[11] Weinan E, Di L, Vanden-Eijnden E. Analysis of multiscale methods for stochastic differential equations[J]. Communications on Pure and Applied Mathematics, 2005, 58(11): 1544–1585.

[12] Freidlin M, Hu W, Wentzell A. Small mass asymptotic for the motion with vanishing friction[J]. Stochastic Processes & Their Applications, 2013, 123(1): 45–75.

[13] Gao H, Duan J. Dynamics of quasi-geostrophic fluid motion with rapidly oscillating Coriolis force[J]. Nonlinear Analysis Real World Applications, 2003, 4(1): 127–138.

[14] Hairer M, Pardoux T. Fluctuations around a homogenised semilinear random PDE[OL]. arXiv, https://arxiv.org/abs/1911.02865.

[15] Hottovy S, Mcdaniel A, Volpe G, et al. The Smoluchowski-Kramers Limit of Stochastic Differential Equations with Arbitrary State-Dependent Friction[J]. Communications in Mathematical Physics, 2015, 336(3): 1259–1283.

[16] Kang H W, Kurtz T G. Separation of time-scales and model reduction for stochastic reaction networks[J]. Annals of Applied Probability, 2010, 23(2): 164–187.

[17] KKhasminskii R Z. A Limit Theorem for the Solution of Differential Equations with Random Right-Hand Sides[J]. Theory of Probability and Its Applications, 1966, 11(3): 390–406.

[18] Khasminskii R Z, Yin G. On Averaging Principles: An Asymptotic Expansion Approach[J]. SIAM Journal on Mathematical Analysis, 2004, 35(6): 1534–1560.

[19] Khasminskii R Z, Yin G. Limit behavior of two-time-scale diffusions revisited[J]. Journal of Differential Equations, 2005, 212(1): 85–113.

[20] Kifer Y. Averaging in dynamical systems and large deviations[J]. Inventiones Mathematicae, 1992, 110(1): 337–370.

[21] Kifer Y. Averaging and climate models[M]. Progress in Probability book series, Springer, Stochastic Climate Models, 2001.

[22] Kifer Y. L^2 diffusion approximation for slow motion in averaging[J]. Stochastics and Dynamics, 2003, 03(02): 213–246.

[23] Kosygina E, Rezakhanlou F, Varadhan S R S. Stochastic homogenization of Hamilton-Jacobi-Bellman equations[J]. Communications on Pure and Applied Mathematics, 2006, 59: 1489–1521.

[24] Kuehn C. Multiple time scale dynamics[M]. Springer, volume 191 of Applied Mathematical Sciences, 2015.

[25] Li X M. An averaging principle for a completely integrable stochastic Hamiltonian system[J]. Nonlinearity, 2008, 21(4): 803.

[26] Majda A, Timofeyev I, Vanden-Eijnden E. A mathematical framework for stochastic climate models[J]. Communications on Pure and Applied Mathematics, 2001, 54: 891–974.

[27] Papanicolaou G C, Stroock D, Varadhan S R S. Martingale approach to some limit theorems[C].

Proceedings of the 1976 Duke University Conference on Turbulence, Durham, NC, 1976.

[28] Pardoux E, Veretennikov A Y. On the Poisson equation and diffusion approximation I[J]. The Annals of Probability, 2001, 29: 1061-1085.

[29] Pardoux E, Veretennikov A Y. On the Poisson equation and diffusion approximation 2[J]. The Annals of Probability, 2003, 31: 1166-1192.

[30] Pavliotis G A, Stuart A M. Multiscale methods: averaging and homogenization[M]. New York: Springer, 2008.

[31] Pei B , Xu Y. L^p-strong convergence in averaging principle for two time-scales stochastic evolution equations driven by Lévy process[OL]. arXiv, https: //arxiv.org/abs/1511.03438.

[32] Röckner M, Xie L. Diffusion approximation for fully coupled stochastic differential equationS[J]. Annals of Probability, 2021, 49(3): 1205-1236.

[33] Röckner M, Xie L. Averaging principle and normal deviations for multiscale stochastic systems[J]. Communications In Mathematical Physics, 2021, 383: 1889-1937.

[34] Sun X, Xie L, Xie Y. Strong and weak convergence rates for slow-fast stochastic differential equations driven by α-stable process[OL]. Bernoulli, arXiv, 2020, https: //arxiv.org/abs/2004.02595.

[35] Veretennikov A Y. On the averaging principle for systems of stochastic differential equations[J]. Mathematics of the Ussr Sbornik, 1991, 69(1): 271-284.

[36] Veretennikov A Y. On polynomial mixing bounds for stochastic differential equations. Stochastic Processes & Their Applications, 1997, 70: 115-127.

[37] Wang W, Roberts A J. Average and deviation for slow-fast stochastic partial differential equations[J]. Journal of Differential Equations, 2012, 253(5): 1265-1286.

[38] Xie L, Yang L. Diffusion approximation for multi-scale stochastic reaction-diffusion equations[J]. Journal of Differential Equations, 2021, 300: 155-184.

[39] Xie L, Yang L. The Smoluchowski-Kramers limits of stochastic differential equations with irregular coefficients[J]. Stochastic Processes and their Applications, 2022, 150: 91-115.

[40] Xie L, Zhang X. Ergodicity of stochastic differential equations with jumps and singular coefficients[J]. Annales de L'institut Henri Poincare, 2020, 56(1): 175-229.

[41] Zhang X. Stochastic Homeomorphism Flows of SDEs with Singular Drifts and Sobolev Diffusion Coefficients[J]. Electronic Journal of Probability, 2010, 16(13): 1096-1116.