Journal of
Linear and Topological Algebra
Vol. 04*, No.* 02*,* 2015*,* 101- 114

Stochastic averaging for SDEs with Hopf Drift and polynomial diffusion coefficients

M. Alvand*

Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran, 84156-83111.

Received 7 May 2015; Revised 25 September 2015; Accepted 18 October 2015.

Abstract. It is known that a stochastic differential equation (SDE) induces two probabilistic objects, namely a diffusion process and a stochastic flow. While the diffusion process is determined by the infinitesimal mean and variance given by the coefficients of the SDE, this is not the case for the stochastic flow induced by the SDE. In order to characterize the stochastic flow uniquely the infinitesimal covariance given by the coefficients of the SDE is needed in addition. The SDEs we consider here are obtained by a weak perturbation of a rigid rotation by random fields which are white in time. In order to obtain information about the stochastic flow induced by this kind of multiscale SDEs we use averaging for the infinitesimal covariance. The main result here is an explicit determination of the coefficients of the averaged SDE for the case that the diffusion coefficients of the initial SDE are polynomial. To do this we develop a complex version of Cholesky decomposition algorithm.

Keywords: Stochastic differential equation; stochastic flow; stochastic averaging; Cholesky decomposition; system of complex bilinear equations.

2010 AMS Subject Classification: 34C29; 60H10; 60H35; 15A23.

1. Introduction

The method of averaging is used to describe the behavior of a dynamical system driven by a fast varying force, i.e. consider the d-dimensional dynamical system described by

$$\dot{X}_t^\varepsilon = b(X_t^\varepsilon, \tfrac{t}{\varepsilon}), \qquad X_0^\varepsilon = x, \tag{1}$$

*Corresponding author.
E-mail address: m.alvand@math.iut.ac.ir (M. Alvand).

Print ISSN: 2252-0201
Online ISSN: 2345-5934

http://jlta.iauctb.ac.ir

where b is a time-dependent vector field on $\mathbb{R}^d$. Since X_t^ε depends on the initial value x we sometimes denote it by $X_t^{\varepsilon,x}$. For $\varepsilon \to 0$, the behavior of system is determined by the time average of the force. More precisely, if

$$\hat{b}(x) := \lim_{t\to\infty} \frac{1}{t} \int_0^t b(x,t)dt. \tag{2}$$

is a well defined vector field and the convergence is uniform on compact sets of x, then for any $T \in (0,\infty)$ the behavior of (1) in the interval $[0,T]$, as $\varepsilon \to 0$, is described by

$$\dot{\hat{X}}_t = \hat{b}(\hat{X}_t), \qquad \hat{X}_0 = x. \tag{3}$$

For an example of a rigorous statement of the averaging method see Theorem 4.3.6 in [10]. An important example is the periodic case, where $b(x,t) = b(x,t+T)$, $x \in \mathbb{R}^d, t \geq 0$, for some $T > 0$. In this situation the limit in (2) is $\hat{b}(x) = \frac{1}{T}\int_0^T b(x,t)dt$.

Stochastic averaging is now an extension of this idea to stochastic differential equations (SDEs) like

$$dX_t^\varepsilon = b(X_t^\varepsilon, \tfrac{t}{\varepsilon})dt + \sum_{r=1}^n \sigma_r(X_t^\varepsilon, \tfrac{t}{\varepsilon})dW_t^{(r)}, \qquad X_0^\varepsilon = x, \tag{4}$$

where $\{W_\cdot^{(r)} \mid r = 1,\ldots,n\}$ is a familiy of independent standard Wiener processes.

The phrase 'stochastic averaging' is also used, even more frequently, for random differential equations (RDEs). In the literature RDEs and SDEs are distinguished by the characteristics of the noise being colored or white in time, resp. This implies that RDEs and SDEs differ in the differentiablity properties of the solution paths $t \to X_t^\varepsilon(\omega)$ (see chapter 2 of [3]).

Before studying the behavior of (4) it is necessary to specify what is modeled by (4); a diffusion process or a stochastic flow.

Considering the SDE (4) as a description of a diffusion process means to study the one-point motions of (4) the solutions $\{X_t^{\varepsilon,x} \mid t \geq 0\}$, $x \in \mathbb{R}^d$. Two quantities uniquely determine the laws of the one-point motions, namely the infinitesimal mean and the infinitesimal variance of (4). The infinitesimal mean of (4) is the drift vector field $b(x, \frac{t}{\varepsilon})$. The infinitesimal variance for (4) is computed by the formula

$$\alpha(x, \tfrac{t}{\varepsilon}) = \sum_{r=1}^n \sigma_r(x, \tfrac{t}{\varepsilon})\sigma_r^\top(x, \tfrac{t}{\varepsilon}),$$

where '$\top$' stands for transpose. After finding the averaged infinitesimal mean $\hat{b}$ and the averaged infinitesimal variance $\hat{\alpha}$ by a procedure similiar to (2), an SDE of the form

$$d\hat{X}_t = \hat{b}(\hat{X}_t)dt + \sum_{r=1}^{\hat{n}} \hat{\sigma}(\hat{X}_t)dW_t^r, \qquad \hat{X}_0 = x, \tag{5}$$

can be constructed to describe the limiting behavior of (4), as $\varepsilon \to 0$, where the family of diffusion vector fields $\{\hat{\sigma}_r \mid r = 1, \dots, \hat{n}\}$ satisfies

$$\hat{\alpha}(x) = \sum_{r=1}^{\hat{n}} \hat{\sigma}_r(x)\hat{\sigma}_r^\top(x). \tag{6}$$

As described in [7], one solution to (6) is to take $\hat{n} = d$ and σ_r to be the rth column of $\hat{\alpha}^{\frac{1}{2}}$, $r = 1, \dots, d$.

The validity of using the simpler SDE (5) to approximate the laws of the one-point motions of (4) in finite time intervals, has been studied by several authors, see [1] and references there. Some other works show the use of the method in problems like the approximation of the law of $\{\chi_t := \max\{g(X_s) \mid 0 \le s \le t\} \mid t \in [0, T]\}$, where g is some real valued function ([6]), the approximation of the stationary density or its marginals ([8] and [13]) and the study of Lyapunov stability of the deterministic fixed points ([13]). Note that stationary densities and the leading Lyapunov exponent of a deterministic fixed point for (4) depend only on the laws of the one-point motions. Therefore the above way of stochastic averaging is expected to be a suitable method to study them and indeed, this claim is proved in appendix of [4].

Considering the SDE (4) as a description of a stochastic flow, it means to study $\{X_\cdot^{\varepsilon,\cdot}\}$, which is a stochastic process indexed by $[0, T] \times \mathbb{R}^d$. The infinitesimal mean and the infinitesimal covariance of (4) determine the law of its induced stochastic flow uniquely ([5]). The infinitesimal covariance is given by

$$\mathcal{A}(x, y, \tfrac{t}{\varepsilon}) = \sum_{r=1}^{n} \sigma_r(x, \tfrac{t}{\varepsilon})\sigma_r^\top(y, \tfrac{t}{\varepsilon}).$$

Using stochastic averaging to describe the limiting behavior of the stochastic flow induced by (4), as $\varepsilon \to 0$, is needed to do averaging procedure on the infinitesimal covariance. This approach is helpful in studying general stability problems i.e. stability along non-deterministic solutions (see [4]). For general results on the theory of stochastic flows and using stochastic averaging for them see [9]. Particular examples, which show that the consideration of the infinitesimal covariance is necessary, can be found in [11] and [4].

The SDEs we consider here are obtained by a weak perturbation of a rigid rotation in Euclidean plane by random fields which are white in time. This situation is also considered in [4] and it is a framework for the problem of 'stochastic Hopf bifurcation'. In order to obtain information about the stochastic flow induced by this kind of SDEs we use stochastic averaging for the infinitesimal covariance. The main result here is an explicit determination of the coefficients of the averaged SDE for the case that the diffusion coefficients of the initial SDE are polynomial.

The paper is organized as follows. In Section 2 we discuss the problem of determination of the averaged SDE, considering SDEs as stochastic flows, for weak perturbation of a rigid rotation by white noises and possibly deterministic forces. It turns out that the problem reduces to solving a system of bilinear functional equations. In Section 3 we continue the problem with the additional assumption that the diffusion vector fields σ_r are polynomials. Averaging of the infinitesimal covariance is described in Subsection 3.1. In Subsection 3.1 first we see that the system of functional equations in the last of Section 2 is became algebraeic by the aditional assumption of this Section 3. Then we deal with derivation of the averaged SDE, i.e. an SDE consistent with the averaged infinitesimal covariance. The main step is developing a strategy for solving a class of system of bilinear equations. The main result is given in Subsection 3.3. The example in Section 4 shows how to use the main result in the case that there are more than one driving Wiener processes.

2. Statement of the problem

Consider the SDE

$$\begin{pmatrix} dX_t \\ dY_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} dt + \varepsilon \begin{pmatrix} f_{0,1}(X_t, Y_t) \\ f_{0,2}(X_t, Y_t) \end{pmatrix} dt + \sqrt{\varepsilon} \sum_{r=1}^{n} \begin{pmatrix} f_{r,1}(X_t, Y_t) \\ f_{r,2}(X_t, Y_t) \end{pmatrix} dW_t^{(r)} , \quad (7)$$

where $f_{r,k} : \mathbb{R}^2 \to \mathbb{R}$ are Lipschitz continuous functions (in the next section will be assumed polynomials), $r = 0, \ldots, n, \ , k = 1, 2$, $\{W_t^{(r)} \mid r = 1, \ldots, n\}$ is a family of independent Wiener processes and ε is a small parameter.

The dynamics of $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ is a small perturbation of the rigid rotation. We define new processes which describe the slow part of the motion:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} \tilde{X}_t \\ \tilde{Y}_t \end{pmatrix}$$

The new variables, after the time rescaling $t \to \varepsilon t$, satisfy the non-autonomous SDE

$$\begin{pmatrix} d\tilde{X}_t \\ d\tilde{Y}_t \end{pmatrix} = \begin{pmatrix} \tilde{f}_{0,1}(\tilde{X}_t, \tilde{Y}_t, \frac{t}{\varepsilon}) \\ \tilde{f}_{0,2}(\tilde{X}_t, \tilde{Y}_t, \frac{t}{\varepsilon}) \end{pmatrix} dt + \sum_{r=1}^{n} \begin{pmatrix} \tilde{f}_{r,1}(\tilde{X}_t, \tilde{Y}_t, \frac{t}{\varepsilon}) \\ \tilde{f}_{r,2}(\tilde{X}_t, \tilde{Y}_t, \frac{t}{\varepsilon}) \end{pmatrix} dW_t^{(r)} , \quad (8)$$

where, for $r = 0, 1, \ldots, n$,

$$\begin{pmatrix} \tilde{f}_{r,1}(x, y, t) \\ \tilde{f}_{r,2}(x, y, t) \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} f_{r,1}(\cos(t)x + \sin(t)y, -\sin(t)x + \cos(t)y) \\ f_{r,2}(\cos(t)x + \sin(t)y, -\sin(t)x + \cos(t)y) \end{pmatrix},$$

Note that (8) is of the form (4) and therefore stochastic averaging becomes applicable here. Since the coefficients of the SDE (8) are 2π-periodic functions in $\frac{t}{\varepsilon}$ the limits like (2) are obtained by averaging over $[0, 2\pi]$. The averaging method suggests that the SDE in any finite time interval can be approximated by an SDE like

$$\begin{pmatrix} d\hat{X}_t \\ d\hat{Y}_t \end{pmatrix} = \begin{pmatrix} \hat{f}_{0,1}(\hat{X}_t, \hat{Y}_t) \\ \hat{f}_{0,2}(\hat{X}_t, \hat{Y}_t) \end{pmatrix} dt + \sum_{r=1}^{\hat{n}} \begin{pmatrix} \hat{f}_{r,1}(\hat{X}_t, \hat{Y}_t) \\ \hat{f}_{r,2}(\hat{X}_t, \hat{Y}_t) \end{pmatrix} dW_t^{(r)} , \quad (9)$$

possibly with $\hat{n} = \infty$. In order to justify the averaging method one now would have to investigate the limiting behavior of the stochastic flow generated by (8) as $\varepsilon \to 0$. Here we do not consider this problem, mainly since it is very technical and does not lead to new insights. This problem has been dealt with, for instance, in [9] (Theorem 5.6.1) and in [4]. We concentrate on the determination of $\hat{n}$ and $\{\hat{f}_r\}_{r=0}^{\hat{n}}$. Otherwords, among the assumptions considered in [9] (Condition $(C.7)_k$ of Theorem 5.6.1) we concentrate on the first one expressed there. This condition is stated in our problem as for all $T \geq 0$ and

$x_1, y_1, x_2, y_2 \in \mathbb{R}$

$$\sum_{r=1}^{\hat{n}} \hat{f}_r(x_1, y_1)\hat{f}_r^\top(x_2, y_2) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_T^{T+\varepsilon} \sum_{r=1}^{n} \tilde{f}_r(x_1, y_1, \frac{t}{\varepsilon^2})\tilde{f}_r^\top(x_2, y_2, \frac{t}{\varepsilon^2})dt. \quad (10)$$

By the time periodicity of $\tilde{f}_r$s the limit in the right hand side can be simplified and one can consider the following relation, as in [4], instead of (10):

$$\sum_{r=1}^{\hat{n}} \hat{f}_r(x_1, y_1)\hat{f}_r^\top(x_2, y_2) = \frac{1}{2\pi} \int_0^{2\pi} \sum_{r=1}^{n} \tilde{f}_r(x_1, y_1, t)\tilde{f}_r^\top(x_2, y_2, t)dt \quad (11)$$

Now, there is no canonical way to construct $\hat{n}$ and $\{\hat{f}_r\}_{r=1}^{\hat{n}}$ satisfying (11). The problem is of the following form:

Problem I: For the matrix function $A : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$, find $\hat{n} \in \mathbb{N} \cup \{\infty\}$ and a family of vector fields $\{\hat{f}_r\}_{r=1}^{\hat{n}}$ such that

$$\sum_{r=1}^{\hat{n}} \hat{f}_r(x_1, y_1)\hat{f}_r^\top(x_2, y_2) = A(x_1, y_1, x_2, y_2). \quad (12)$$

A theoretical solution of Problem I (in a more general context) is discussed in [4], while [2] suggests a more constructive way. Here we mainly deal with solving Problem I but we want to treat this problem in a way that gives information about analytic relations between parameters of the vector fields $\{f_r \mid r = 1, \ldots, n\}$ from (7) and $\{\hat{f}_r\}_{r=1}^{\hat{n}}$ from (9). In the deterministic case, i.e. determination of $\hat{f}_0$, such analytic relations can be found by considering the equation in the complex plane (see Section 19.2A of [12]). Using this idea, we rewrite (7) in the complex plane using

$$X_t = \frac{Z_t + \bar{Z}_t}{2},\ Y_t = \frac{Z_t - \bar{Z}_t}{2i}\ .$$

Then (8) and (9), resp., are equivalent to

$$d\tilde{Z}_t = \tilde{g}_0(\tilde{Z}_t, \bar{\tilde{Z}}_t, \tfrac{t}{\varepsilon})dt + \sum_{r=1}^{n} \tilde{g}_r(\tilde{Z}_t, \bar{\tilde{Z}}_t, \tfrac{t}{\varepsilon})dW_t^{(r)}\ , \quad (13)$$

and

$$d\hat{Z}_t = \hat{g}_0(\hat{Z}_t, \bar{\hat{Z}}_t)dt + \sum_{r=1}^{n} \hat{g}_r(\hat{Z}_t, \bar{\hat{Z}}_t)dW_t^{(r)}\ , \quad (14)$$

where

$$\tilde{g}_r(z, \bar{z}, t) := \tilde{f}_{r,1}\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}, t\right) + i\tilde{f}_{r,2}\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}, t\right), \quad r = 0, \ldots, n$$

and

$$\hat{g}_r(z,\bar{z}) := \hat{f}_{r,1}\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}\right) + i\tilde{f}_{r,2}\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}\right), \quad r = 0,\ldots,\hat{n}.$$

Then a reformulation of (11) gives

$$\sum_{r=1}^{\hat{n}} \hat{g}_r(z_1,\bar{z}_1)\overline{\hat{g}_r(z_2,\bar{z}_2)} = \frac{1}{2\pi}\int_0^{2\pi}\sum_{r=1}^{n}\tilde{g}_r(z_1,\bar{z}_1,t)\overline{\tilde{g}_r(z_2,\bar{z}_2,t)}dt, \tag{15}$$

and

$$\sum_{r=1}^{\hat{n}} \hat{g}_r(z_1,\bar{z}_1)\hat{g}_r(z_2,\bar{z}_2) = \frac{1}{2\pi}\int_0^{2\pi}\sum_{r=1}^{n}\tilde{g}_r(z_1,\bar{z}_1,t)\tilde{g}_r(z_2,\bar{z}_2,t)dt. \tag{16}$$

Recall that for complex valued stochastic processes $\left\{\xi_t^{(1)} \mid t>0\right\}$ and $\left\{\xi_t^{(2)} \mid t>0\right\}$ the *infinitesimal covariance* is defined by

$$\mathcal{B}(t,z_1,z_2) = \lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}\left[\begin{pmatrix}\xi_{t+h}^{(1)}-z_1\\ \overline{\xi_{t+h}^{(1)}-z_1}\end{pmatrix}\left(\overline{\xi_{t+h}^{(2)}-z_2}\ \xi_{t+h}^{(2)}-z_2\right)\,\Big|\,(\xi_t^{(1)},\xi_t^{(2)})=(z_1,z_2)\right],$$

provided that the limit exists.

Remark 1 *The quantities appearing in the left hand sides or under the integrals in (15) and (16) are related to the infinitesimal covariances of the associated processes. For example consider* $\left\{\hat{Z}_t^{z_1} \mid t\geq 0\right\}$ *and* $\left\{\hat{Z}_t^{z_2} \mid t\geq 0\right\}$ *satisfying (14) with initial conditions* $\hat{Z}_0^{z_k} = z_k$, $k=1,2$. *Using the Itô formula, the infinitesimal covariance of these processes is*

$$\hat{\mathcal{B}}(t,z_1,z_2) = \hat{\mathcal{B}}(z_1,z_2) = \sum_{r=1}^{\hat{n}}\begin{pmatrix}\hat{g}_r(z_1,\bar{z}_1)\overline{\hat{g}_r(z_2,\bar{z}_2)} & \hat{g}_r(z_1,\bar{z}_1)\hat{g}_r(z_2,\bar{z}_2)\\ \overline{\hat{g}_r(z_1,\bar{z}_1)\hat{g}_r(z_2,\bar{z}_2)} & \overline{\hat{g}_r(z_1,\bar{z}_1)}\hat{g}_r(z_2,\bar{z}_2)\end{pmatrix}. \tag{17}$$

From (15) and (16) we infer that

$$\hat{\mathcal{B}}(z_1,z_2) = \frac{1}{2\pi}\sum_{r=1}^{n}\int_0^{2\pi}\begin{pmatrix}\tilde{g}_r(z_1,\bar{z}_1,t)\overline{\tilde{g}_r(z_2,\bar{z}_2,t)} & \tilde{g}_r(z_1,\bar{z}_1,t)\tilde{g}_r(z_2,\bar{z}_2,t)\\ \overline{\tilde{g}_r(z_1,\bar{z}_1,t)}\ \tilde{g}_r(z_2,\bar{z}_2,t) & \overline{\tilde{g}_r(z_1,\bar{z}_1,t)\tilde{g}_r(z_2,\bar{z}_2,t)}\end{pmatrix}dt. \tag{18}$$

Finally, comparing (17) and (18), it remains to solve the following problem for $\mathcal{B} = \hat{\mathcal{B}}$ computed by (18).

Problem II: For the matrix function $\mathcal{B}:\mathbb{C}\times\mathbb{C}\to\mathbb{C}^{2\times 2}$ find a family of complex valued functions $\{\hat{g}_r \mid r \in K\}$ with cardinality at most countable, such that

$$\begin{cases} \sum_{r\in K} \hat{g}_r(z_1,\bar{z}_1)\overline{\hat{g}_r(z_2,\bar{z}_2)} = \mathcal{B}_{11}(z_1,z_2), \\ \\ \sum_{r\in K} \hat{g}_r(z_1,\bar{z}_1)\hat{g}_r(z_2,\bar{z}_2) = \mathcal{B}_{12}(z_1,z_2). \end{cases} \tag{19}$$

3. The averaged SDE for polynomial diffusion coefficients

In this section, assuming the vector fields f_r, $r = 1,\dots,n$, in (7) to be polynomials, the vector fields $\hat{g}_r$, $r = 0,\dots,\hat{n}$, in (14) are explicitly constructed. In the first subsection the infinitesimal covariance $\hat{\mathcal{B}}$ is computed explicitly and it is found that this matrix function can be decomposed to simpler ones. In the second subsection a method for solving Problem II is developed and implemented for the matrix functions from the decomposition of the infinitesimal covariance. The results are given in Subsection 3.3.

3.1 *Averaging of the infinitesimal covariance*

The computation of $\hat{f}_0$, or equivalently $\hat{g}_0$, has been solved already. We consider Problem II. For simplicity let $n = 1$ and suppose that $f_{1,j}$ is a polynomial in x, y, $j = 1,2$ and put $m = \max\{\deg(f_{1,1}), \deg(f_{1,2})\}$. Therefore (7) takes the form

$$\begin{pmatrix} dX_t \\ dY_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} dt + \varepsilon \begin{pmatrix} f_{0,1}(X_t,Y_t) \\ f_{0,2}(X_t,Y_t) \end{pmatrix} dt + \sqrt{\varepsilon} \begin{pmatrix} \sum_{d=0}^{m}\sum_{k=0}^{d} a_{d,k} X_t^k Y_t^{d-k} \\ \sum_{d=0}^{m}\sum_{k=0}^{d} b_{d,k} X_t^k Y_t^{d-k} \end{pmatrix} dW_t \ , \tag{20}$$

Then $Z_t = X_t + iY_t$ satisfy

$$dZ_t = iZ_t dt + \varepsilon g_0(Z_t,\bar{Z}_t)dt + \sqrt{\varepsilon}\sum_{d=0}^{m}\sum_{k=0}^{d} \alpha_{d,k} Z_t^k \bar{Z}_t^{d-k} dW_t, \tag{21}$$

where $g_0(z,\bar{z}) := f_{0,1}(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}) + i f_{0,2}(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i})$ and

$$\alpha_{d,k} = \frac{1}{2^d}\sum_{j_1=0}^{k}\sum_{j_2=j_1}^{d-k+j_1} (a_{d,j_2} + i b_{d,j_2})\, i^{d-j_2}(-1)^{k-j_1}\binom{j_2}{j_1}\binom{d-j_2}{k-j_1}. \tag{22}$$

Since the SDE (21) is a perturbation of the rigid rotation around the origin, defining a new process by $Z_t = e^{it}\tilde{Z}_t$, we put the system in a form that we can apply stochastic averaging. This gives an SDE of the form (13) with

$$\tilde{g}_1(z,\bar{z},t) = \sum_{d=0}^{m}\sum_{k=0}^{d} \alpha_{d,k} z^k \overline{z}^{d-k} e^{i(2k-d-1)t} \tag{23}$$

For convenience, we rearrange the indices in (23). Let $I_m := \{(d,k) \mid d = 0,1,\ldots,m\ ,\ k = 0,\ldots,d\}$, $\delta_m := \#I_m$. Then we have the one-to-one correspondence

$$\begin{cases} I_m \to \{1,\ldots,\delta_m\} \\ (d,k) \to \frac{d(d+1)}{2} + k + 1. \end{cases}$$

The function $\tilde{g}_1$ can be written as

$$\tilde{g}_1(z,\bar{z},t) = \sum_{p=1}^{\delta_m} \tilde{\alpha}_p(t)\phi_p(z),$$

where $\tilde{\alpha}_p(t) := \alpha_{d,k}e^{i(2k-d-1)t}$ and $\phi_p(z) := z^k\overline{z}^{d-k}$, $(d,k) \in I_m$ is determined by $\frac{d(d+1)}{2} + k + 1 = p$, $p = 1,\ldots,\delta_m$.

Therefore the infinitesimal covariance of (13), the matrix function under the integral in (18), is

$$\tilde{\mathcal{B}}(t,z_1,z_2) = \varepsilon \sum_{p=1}^{\delta_m}\sum_{q=1}^{\delta_m} \begin{pmatrix} M_{p,q}(t)\phi_p(z_1)\overline{\phi_q(z_2)} & N_{p,q}(t)\phi_p(z_1)\phi_q(z_2) \\ \overline{N_{p,q}(t)\phi_p(z_1)\phi_q(z_2)} & \overline{M_{p,q}(t)\phi_p(z_1)}\phi_q(z_2) \end{pmatrix} \tag{24}$$

where

$$M_{p,q}(t) := \tilde{\alpha}_p(t)\overline{\tilde{\alpha}_q(t)} = e^{-((d_1-d_2)+2(k_1-k_2))it}\alpha_{d_1,k_1}\overline{\alpha_{d_2,k_2}}, \tag{25}$$

and

$$N_{p,q}(t) = \tilde{\alpha}_p(t)\tilde{\alpha}_q(t) = e^{-((d_1+d_2+2)+2(k_1+k_2))it}\alpha_{d_1,k_1}\alpha_{d_2,k_2}, \tag{26}$$

and (d_1,k_1) and (d_2,k_2) are determined by $\frac{d_1(d_1+1)}{2} + k_1 + 1 = p$ and $\frac{d_2(d_2+1)}{2} + k_2 + 1 = q$, $p,q = 1,\ldots,\delta_m$.

The computation of the averaged infinitesimal covariance is easy using formula (18), since $\mathcal{B}$ is 2π-periodic in t. We use the following notations for averages of functions defined in (25) and (26):

$$\hat{M}_{p,q} := \frac{1}{2\pi}\int_0^{2\pi} M_{p,q}(t)dt = \begin{cases} \alpha_{d_1,k_1}\bar{\alpha}_{d_2,k_2}\ d_1 - d_2 = 2(k_1-k_2) \\ \quad 0 \qquad\qquad \text{otherwise} \end{cases}, \tag{27}$$

$$\hat{N}_{p,q} := \frac{1}{2\pi}\int_0^{2\pi} N_{p,q}(t)dt = \begin{cases} \alpha_{d_1,k_1}\alpha_{d_2,k_2}\ d_1 + d_2 + 2 = 2(k_1+k_2) \\ \quad 0 \qquad\qquad \text{otherwise} \end{cases}. \tag{28}$$

Then

$$\hat{\mathcal{B}}(z_1,z_2) = \varepsilon \sum_{p=1}^{\delta_m}\sum_{q=1}^{\delta_m} \begin{pmatrix} \hat{M}_{p,q}\phi_p(z_1)\overline{\phi_q(z_2)} & \hat{N}_{p,q}\phi_p(z_1)\phi_q(z_2) \\ \overline{\hat{N}_{p,q}\phi_p(z_1)\phi_q(z_2)} & \overline{\hat{M}_{p,q}\phi_p(z_1)}\phi_q(z_2) \end{pmatrix} \tag{29}$$

Consistent with the relations (27) and (28) we make a partition on I_m by decomposing $I_m = \bigcup_{s=-m+1}^{m+1} I_{m,s}$ where

$$I_{m,s} = \{(d,k) \in I \mid d - 2k + 1 = s\}. \tag{30}$$

Since I_m and $\{1,\ldots,\delta_m\}$ are in one-to-one correspondence, by partioning the first set, we get a partition on the second one that is $\{1,\ldots,\delta_m\} = \bigcup_{s=-m+1}^{m+1} J_{m,s}$ where

$$J_{m,s} := \left\{ p \in \{1,\ldots,\delta_m\} \mid \exists (d,k) \in I_{m,s} : \frac{d(d+1)}{2} + k + 1 = p \right\}.$$

Note that any matrix under the summation sign in (29) is zero unless $|d_1 - 2k_1 + 1| = |d_2 - 2k_2 + 1|$ or equivalently $p,q \in J_{m,s}$ or $p \in J_{m,s}$ and $q \in J_{m,-s}$, for some $s \in \{-m+1,\ldots,m+1\}$. So we rewrite (29) as follows;

$$\hat{\mathcal{B}}(z_1,z_2) = \sum_{s=0}^{m+1} \hat{\mathcal{B}}_s(z_1,z_2), \tag{31}$$

where

$$\hat{\mathcal{B}}_0(z_1,z_2) = \sum_{p,q\in J_{m,0}} \begin{pmatrix} \hat{M}_{p,q}\phi_p(z_1)\overline{\phi_q(z_2)} & \hat{N}_{p,q}\phi_p(z_1)\phi_q(z_2) \\ \overline{\hat{N}_{p,q}}\overline{\phi_p(z_1)\phi_q(z_2)} & \overline{\hat{M}_{p,q}}\overline{\phi_p(z_1)}\phi_q(z_2) \end{pmatrix}, \tag{32}$$

for $s = 1,\ldots,m+1$

$$\begin{aligned}\hat{\mathcal{B}}_s(z_1,z_2) = &\sum_{p,q\in J_{m,s}} \begin{pmatrix} \hat{M}_{p,q}\phi_p(z_1)\overline{\phi_q(z_2)} & 0 \\ 0 & \overline{\hat{M}_{p,q}}\overline{\phi_p(z_1)}\phi_q(z_2) \end{pmatrix} \\ &+ \sum_{p,q\in J_{m,-s}} \begin{pmatrix} \hat{M}_{p,q}\phi_p(z_1)\overline{\phi_q(z_2)} & 0 \\ 0 & \overline{\hat{M}_{p,q}}\overline{\phi_p(z_1)}\phi_q(z_2) \end{pmatrix} \\ &+ \sum_{p\in J_{m,s}\ ,q\in J_{m,-s}} \begin{pmatrix} 0 & \hat{N}_{p,q}\phi_p(z_1)\phi_q(z_2) \\ \overline{\hat{N}_{p,q}}\overline{\phi_p(z_1)\phi_q(z_2)} & 0 \end{pmatrix} \\ &+ \sum_{p\in J_{m,-s},q\in J_{m,s}} \begin{pmatrix} 0 & \hat{N}_{p,q}\phi_p(z_1)\phi_q(z_2) \\ \overline{\hat{N}_{p,q}}\overline{\phi_p(z_1)\phi_q(z_2)} & 0 \end{pmatrix}.\end{aligned}$$

Note that for $s = m, m+1$ $J_{m,-s} = \emptyset$.

3.2 Solving problem II for $\mathcal{B} = \hat{\mathcal{B}}$

We now describe how Problem II can be solved for $\mathcal{B} = \hat{\mathcal{B}}_s,\ s = 1, \ldots, m-1$. Solving it for $\mathcal{B} = \hat{\mathcal{B}}_s,\ s = 0, m, m+1$, obeys a same strategy. Fix $s \in \{1, \ldots, m-1\}$. By our partitioning on I, the unknown functions must be of the form

$$\hat{g}_r(z, \bar{z}) = \sum_{p \in J_{m,s} \cup J_{m,-s}} \beta_{p,r} \phi_p(z) \tag{33}$$

Equating the coefficients of same terms in both sides of equations in (19) we obtain the following system of complex bilinear equations:

$$\begin{cases} \sum\limits_{r \in K} \beta_{p,r} \bar{\beta}_{q,r} = \hat{M}_{p,q} & p, q \in J_{m,s} \cup J_{m,-s} \\ \sum\limits_{r \in K} \beta_{p,r} \beta_{q,r} = \hat{N}_{p,q} & p, q \in J_{m,s} \cup J_{m,-s} \end{cases}, \tag{34}$$

where the set K remains to be determined.

Remark 2 *To give a method to solve (34) it is compared with the Cholesky decomposition which solves the system of equations*

$$\sum_{r=1}^{m} \beta_{p,r} \beta_{q,r} = a_{p,q} \qquad p, q = 1, \ldots, m$$

for unknowns $\beta_{p,r}, p, r = 1, \ldots, m$ in real numbers with some further assumptions on unknowns to ensure uniqueness of the solutions while the existence of the solution needs symmetry and non-negative difinity of the matrix $A := (a_{pq})_{p,q=1}^{n}$. We describe a similiar method, which can be called 'the complex Cholesky decomposition'. A presentation of the Cholesky decomposition method can be found in the appendix B of [2].

To determine K, we select it such that the number of unknowns and equations are the same though this does not imply existence and uniqueness of the solution. The number of equations is $2\delta_m^2$ and if we put $K = \{ r \in \mathbb{Z} \mid |r| \in J_{m,s} \cup J_{m,-s} \}$ then

$$\# \{ \beta_{p,r} \mid p \in J_{m,s} \cup J_{m,-s}, r \in K \} = 2\delta_m^2.$$

The non-uniqueness in solving (34) has three sources:

(1) If $p \neq q$ then the equations associated to (p, q) and (q, p) coincide.
(2) The equations are quadratic in the unknown variables.
(3) The matrices $M = (M_{p,q})$ and $N = (N_{p,q})$ may be degenerate.

To remove non-uniqueness we restrict the solution space by putting some conditions on $\{\beta_{pr}\}$;

(a) $\forall r \in K, p \in J_{m,s} \cup J_{m,-s} : \quad p < |r| \Rightarrow \beta_{pr} = 0$.
This removes the non-uniqueness source from (1).
(b) $\forall p \in J_{m,s} \cup J_{m,-s} : \quad |\beta_{p,p}| = |\beta_{p,-p}|$,
(b′) $\forall p \in J_{m,s} \cup J_{m,-s} : \quad \mathrm{Arg}(\beta_{p,p}) - \mathrm{Arg}(\alpha_p) \in [0, \pi)\ ,\ \mathrm{Arg}(\beta_{p,-p}) - \mathrm{Arg}(\beta_{p,p}) \in [0, \pi)$.
These remove the non-uniqueness source from (2).
(c) $\forall r \in K : \quad \beta_{rr} = 0 \Rightarrow \forall p \in J_{m,s} \cup J_{m,-s} :\ \beta_{p,r} = 0$.
This removes the (possible) non-uniqueness source from (3).

Condition (a) implies that (34) can be written as

$$\begin{cases} \sum\limits_{|r|\le \min\{p,q\}} \beta_{p,r}\bar{\beta}_{q,r} = \hat{M}_{p,q} & p,q \in J_{m,s}\cup J_{m,-s}, \\ \sum\limits_{|r|\le \min\{p,q\}} \beta_{p,r}\beta_{q,r} = \hat{N}_{p,q} & p,q \in J_{m,s}\cup J_{m,-s}. \end{cases} \tag{35}$$

Now let $p^* := \min\{p \in J_{m,s}\cup J_{m,-s} \mid \alpha_p \neq 0\}$. Since $\alpha_p = 0$, for $p \in J_{m,s}\cup J_{m,-s}$ with $|p| < p^*$, by (27) and (28), we have $\hat{M}_{p,q} = \hat{N}_{p,q} = 0$, for $p,q \in J_{m,s}\cup J_{m,-s}$ with $\min\{|p|,|q|\} < p^*$. Therefore letting

$$\beta_{p,r} = 0, \quad p < p^*, r \in K \tag{36}$$

results that the equations in (34) with $p,q < p^*$ hold.

The equations in (34) with $p = q = p^*$, using (36), is

$$\begin{cases} \beta_{p^*,p^*}\bar{\beta}_{p^*,p^*} + \beta_{p^*,-p^*}\bar{\beta}_{p^*,-p^*} = \hat{M}_{p^*,p^*} = \alpha_{p^*}\bar{\alpha_{p^*}} \\ \beta_{p^*,p^*}\beta_{p^*,p^*} + \beta_{p^*,-p^*}\beta_{p^*,-p^*} = \hat{N}_{p^*,p^*} = \quad 0 \end{cases}, \tag{37}$$

There are a continuous family of solutions for (37), while the unique solution satisfying (b) and (b$'$) is

$$\beta_{p^*,p^*} = \frac{\sqrt{2}}{2}\alpha_{p^*} \quad , \quad \beta_{p^*,-p^*} = \frac{i\sqrt{2}}{2}\alpha_{p^*}. \tag{38}$$

p^* and $q \in J_{m,s}\cup J_{m,-s}$ with $q > p^*$ are either in a same class, $J_{m,s}$ or $J_{m,-s}$, or $p^* \in J_{m,s}$ and $q \in J_{m,-s}$. In the first case, after using (36) and (38) in (35), we have to solve

$$\begin{cases} \beta_{q,p^*}\frac{\sqrt{2}}{2}\bar{\alpha}_{p^*} + \beta_{q,-p^*}\frac{-i\sqrt{2}}{2}\bar{\alpha}_{p^*} = \hat{M}_{q,p^*} = \alpha_q\bar{\alpha}_{p^*} \\ \beta_{q,p^*}\frac{\sqrt{2}}{2}\alpha_{p^*} + \beta_{q,-p^*}\frac{i\sqrt{2}}{2}\alpha_{p^*} = \hat{N}_{q,p^*} = \quad 0 \end{cases}, \tag{39}$$

which gives the unique solution

$$\beta_{q,p^*} = \frac{\sqrt{2}}{2}\alpha_q \quad , \quad \beta_{q,-p^*} = \frac{i\sqrt{2}}{2}\alpha_q. \tag{40}$$

In the second case, after using (36) and (38) in (35), we have to solve

$$\begin{cases} \beta_{q,p^*}\frac{\sqrt{2}}{2}\bar{\alpha}_{p^*} + \beta_{q,-p^*}\frac{-i\sqrt{2}}{2}\bar{\alpha}_{p^*} = \hat{M}_{q,p^*} = \quad 0 \\ \beta_{q,p^*}\frac{\sqrt{2}}{2}\alpha_{p^*} + \beta_{q,-p^*}\frac{i\sqrt{2}}{2}\alpha_{p^*} = \hat{N}_{q,p^*} = \alpha_q\alpha_{p^*} \end{cases}, \tag{41}$$

which gives the unique solution

$$\beta_{q,p^*} = \frac{\sqrt{2}}{2}\alpha_q \quad , \quad \beta_{q,-p^*} = \frac{-i\sqrt{2}}{2}\alpha_q. \tag{42}$$

After substituting (36), (38), (40) and (42) in the equations in (35) with $p,q > p^*$, it is found that

$$\beta_{p,r} = 0, \quad p \in J_{m,s} \cup J_{m,-s}, r \in K, |r| > p^*. \tag{43}$$

The result of the calculations performed above is formulated in Lemma 3.1.

3.3 *Results*

Lemma 3.1 Problem II for $\mathcal{B} = \hat{\mathcal{B}}_s$, $1 \leq s \leq m-1$ is solved by the pair of complex valued functions

$$\hat{g}_{2s}(z,\bar{z}) = \frac{\sqrt{2}}{2} \sum_{k=0}^{\left[\frac{m-s+1}{2}\right]} \alpha_{2k+s-1,k} z^k \bar{z}^{k+s-1} + \frac{\sqrt{2}}{2} \sum_{k=1+s}^{\left[\frac{m+s+1}{2}\right]} \alpha_{2k-s-1,k} z^k \bar{z}^{k-s-1},$$

and

$$\hat{g}_{2s+1}(z,\bar{z}) = \frac{i\sqrt{2}}{2} \sum_{k=0}^{\left[\frac{m-s+1}{2}\right]} \alpha_{2k+s-1,k} z^k \bar{z}^{k+s-1} - \frac{i\sqrt{2}}{2} \sum_{k=1+s}^{\left[\frac{m+s+1}{2}\right]} \alpha_{2k-s-1,k} z^k \bar{z}^{k-s-1}.$$

Proof. It can be checked directly, but we also have described the derivation of $\hat{g}_r$ in Subsection 3.2. Note that here we neglect the zero functions and unify the indices for all $s \in \{0, \dots, m+1\}$. Therefore the index of g_r is different from that used in Subsection 3.2. ■

Lemma 3.2 Problem II for $\mathcal{B} = \hat{\mathcal{B}}_0$ is solved by the complex valued function

$$\hat{g}_1(z,\bar{z}) = \sum_{k=1}^{\left[\frac{m+1}{2}\right]} \alpha_{2k-1,k} z^k \bar{z}^{k-1}.$$

Proof. This can be checked directly or found by a same procedure as performed in Subsection 3.2. ■

Lemma 3.3 Problem II for $\mathcal{B} = \hat{\mathcal{B}}_s$, $s = m, m+1$ is solved by the pair of complex valued functions

$$\hat{g}_{2s}(z,\bar{z}) = \frac{\sqrt{2}}{2} \sum_{k=0}^{\left[\frac{m-s+1}{2}\right]} \alpha_{2k+s-1,k} z^k \bar{z}^{k+s-1},$$

and

$$\hat{g}_{2s+1}(z,\bar{z}) = \frac{i\sqrt{2}}{2} \sum_{k=0}^{\left[\frac{m-s+1}{2}\right]} \alpha_{2k+s-1,k} z^k \bar{z}^{k+s-1}.$$

Proof. This is a special case of Lemma 3.1. ■

Theorem 3.4 The averaged SDE associated to the SDE (20) is

$$d\hat{Z}_t = \sum_{s=1}^{2m+3} \hat{g}_s\left(\hat{Z}_t,\ \bar{\hat{Z}}_t\right) dW_t^{(s)}$$

where the functions $\hat{g}_s$ are defined in Lemmas 3.2–3.3.

Proof. This is a consequence of Lemmas 3.2–3.3. ■

4. SDEs driven by several white noises

In the case where the SDE has several white noises, the derivation of the diffusion coefficients of the averaged SDE can be done seperately for each diffusion coefficient. At the end linearly dependent diffusion coefficients can be combined. As an example, consider the SDE

$$\begin{cases} dU_t = & V_t dt, \\ dV_t = \left(-U_t - \varepsilon^2 \lambda V_t - AU_t^3\right) dt + \varepsilon\sigma_1 U_t dW_t^{(1)} + \varepsilon\sigma_2 V_t dW_t^{(2)} + \sigma_3 U_t^2 dW_t^{(3)}. \end{cases}$$

We do not study the existence of the stochastic flow for this SDE here. But we determine an averaged SDE to describe the dynamics near zero. The new processes defined by $U_t = \varepsilon X_t$ and $V_t = \varepsilon Y_t$ satisfy

$$\begin{cases} dX_t = & Y_t dt \\ dY_t = -X_t dt - \varepsilon^2\left(\lambda Y_t + AX_t^3\right) dt + \varepsilon\left(\sigma_1 X_t dW_t^{(1)} + \sigma_2 Y_t dW_t^{(2)} + \sigma_3 X_t^2 dW_t^{(3)}\right), \end{cases}$$

which is of the form (7). Then using (22) we rewrite the equation in the complex plane

$$\begin{aligned} dZ_t = -iZ_t dt + \varepsilon^2 &\left(\frac{\lambda}{2}\bar{Z}_t - \frac{\lambda}{2}Z_t - \frac{Ai}{8}\bar{Z}_t^{\,3} - \frac{3Ai}{8}Z_t\bar{Z}_t^{\,2} - \frac{3Ai}{8}Z_t^2\bar{Z}_t - \frac{Ai}{8}Z_t^3\right) dt \\ &+\varepsilon\left(\frac{i\sigma_1}{2}\left(Z_t + \bar{Z}_t\right) dW_t^{(1)} + \frac{\sigma_2}{2}\left(Z_t - \bar{Z}_t\right) dW_t^{(2)} + \frac{i\sigma_3}{4}\left(Z_t^{\,2} + 2Z_t\bar{Z}_t + \bar{Z}_t^2\right) dW_t^{(3)}\right). \end{aligned}$$

Then we use Theorem 3.4 seperately for each vector field (the drift and the three difusion vector fields) to obtain the averaged SDE:

$$\begin{aligned} d\hat{Z}_t = &-\left(\frac{\lambda}{2}\hat{Z}_t + \frac{3Ai}{8}\hat{Z}_t^2\bar{\hat{Z}}_t\right) dt \\ &+\frac{i\sigma_1}{2}\left(\hat{Z}_t dW_t^{(1,1)} + \frac{\sqrt{2}}{2}\bar{\hat{Z}}_t dW_t^{(1,4)} + \frac{i\sqrt{2}}{2}\bar{\hat{Z}}_t dW_t^{(1,5)}\right) \\ &+\frac{i\sigma_2}{2}\left(\hat{Z}_t dW_t^{(2,1)} + \frac{-\sqrt{2}}{2}\bar{\hat{Z}}_t dW_t^{(2,4)} + \frac{-i\sqrt{2}}{2}\bar{\hat{Z}}_t dW_t^{(2,5)}\right) \\ &+\frac{i\sigma_3\sqrt{2}}{8}\left(\left(2\hat{Z}_t\bar{\hat{Z}}_t + \hat{Z}_t^2\right) dW_t^{(3,2)} + i\left(2\hat{Z}_t\bar{\hat{Z}}_t - \hat{Z}_t^2\right) dW_t^{(3,3)} + \bar{\hat{Z}}_t^2 dW_t^{(3,6)} + i\bar{\hat{Z}}_t^2 dW_t^{(3,7)}\right). \end{aligned}$$

By defining $\eta := \sqrt{\sigma_1^2 + \sigma_2^2}$, $\tilde{W}_t^{(1)} := \frac{\sigma_1}{\eta} W_t^{(1,1)} + \frac{\sigma_2}{\eta} W_t^{(2,1)}$ and $\tilde{W}_t^{(i)} := \frac{\sigma_1}{\eta} W_t^{(1,i)} + \frac{\sigma_2}{\eta} W_t^{(2,i)}$, $i = 4, 5$, the averaged SDE can be written as

$$
\begin{aligned}
d\hat{Z}_t = &-\left(\frac{\lambda}{2}\hat{Z}_t + \frac{3Ai}{8}\hat{Z}_t^2\bar{\hat{Z}}_t\right)dt \\
&+\frac{i\eta}{2}\hat{Z}_t d\tilde{W}_t^{(1)} + \frac{i\eta\sqrt{2}}{4}\bar{\hat{Z}}_t d\tilde{W}_t^{(4)} + \frac{-\eta\sqrt{2}}{4}\bar{\hat{Z}}_t d\tilde{W}_t^{(5)} \\
&+\frac{i\sigma_3\sqrt{2}}{8}\left(\left(2\hat{Z}_t\bar{\hat{Z}}_t + \hat{Z}_t^2\right)dW_t^{(3,2)} + i\left(2\hat{Z}_t\bar{\hat{Z}}_t - \hat{Z}_t^2\right)dW_t^{(3,3)} + \bar{\hat{Z}}_t^2 dW_t^{(3,6)} + i\bar{\hat{Z}}_t^2 dW_t^{(3,7)}\right).
\end{aligned}
$$

Acknowledgements

The author wishes to acknowledge helpful comments from his supervisors Hans Crauel and Hamid R. Z. Zangeneh.

References

[1] N. Abourashchi, A. Yu Veretennikov. On stochastic averaging and mixing, Theory Stoch. Process. 16, (1), (2010) 111-129.
[2] M. Alvand, Constructing an SDE from its two-point generator, Stoch. Dyn. DOI: 10.1142/S0219493715500252
[3] L. Arnold, Random Dynamical Systems, Springer-Verlag, 1998.
[4] P. H. Baxendale, Stochastic averaging and asymptotic behaviour of the stochastic Duffing - Van der Pol equation, Stochastic Process. Appl. 113, No. 2 (2004) 235-272.
[5] ———, Brownian motion in the diffeomorphisms group, Compositio Math. 53, No.1 (1984) 19-50.
[6] P. Bernard, Stochastic averaging, Nonlinear Stochastic Dynamics, (2002) 29-42.
[7] M. I. Freĭdlin, The factorization of nonnegative definite matrices, Teor. Verojatnost. i Primenen. 13 (1968) 375-378.
[8] Z. L. Huang and W. Q. Zhu, Stochastic averaging of quasi-generalized Hamiltonian systems, Int. J. Nonlinear Mech. No.44 (2009) 71-80.
[9] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, 1990.
[10] J. A. Sanders, F. Verhulst and J. Murdack, Averaging Methods in Nonlinear Dynamical Systems, 2nd edition, Springer, 2007.
[11] R. B. Sowers, Averaging of stochastic flows: Twist maps and escape from resonance, Stochastic Process. Appl. No. 119, (2009) 3549-3582.
[12] S. Wiggins, An Introduction to Applied Nonlonear Dynamical Systems and Chaos, second edition, Springer-Verlag, 2009.
[13] W.Q. Zhu, Z.L. Huang and Y. Suzuki, Stochastic averaging and Lyapunov exponent of quasi partially integrable Hamiltonian systems, Int. J. Nonlinear Mech. No.37 (2002) 419-437.