

APPROXIMATION OF DELAY DIFFERENTIAL EQUATIONS AT THE VERGE OF INSTABILITY BY EQUATIONS WITHOUT DELAY

NISHANTH LINGALA

University of Illinois, Urbana, IL, USA
 nlingala1@gmail.com

ABSTRACT. We consider linear delay differential equations at the verge of Hopf instability, i.e. a pair of roots of the characteristic equation are on the imaginary axis of the complex plane and all other roots have negative real parts. When nonlinear and noise perturbations are present, we show that the error in approximating the dynamics of the delay system by certain two dimensional stochastic differential equation *without delay* is small (in an appropriately defined sense). Two cases are considered: (i) linear perturbations and multiplicative noise (ii) cubic perturbations and additive noise. The approximation results are useful because, processes without delay are easier to simulate numerically as they do not require storage of the history of the process.

AMS (MSC2010) Subject Classification. 34K50, 34K27, 34K33.

1. INTRODUCTION

Consider the stochastic delay differential equation (SDDE)

$$(1.1) \quad dx(t) = (\mu x(t-1) + x^3(t)) dt + \varepsilon c_1 x(t-1) dV_1(t) + \varepsilon^2 c_2 dV_2(t),$$

where $0 < \varepsilon \ll 1$ and V_1, V_2 are Wiener processes. The above equation represents a noisy perturbation of the following deterministic system:

$$(1.2) \quad \dot{x}(t) = \mu x(t-1) + x^3(t).$$

The linear system corresponding to (1.2) is

$$(1.3) \quad \dot{x}(t) = \mu x(t-1).$$

Seeking a solution of the form $x(t) = e^{t\lambda}$ to the linear system, we find that λ must satisfy the characteristic equation $\lambda - \mu e^{-\lambda} = 0$. When $\mu \in (-\frac{\pi}{2}, 0)$, all roots of the characteristic equation have negative real parts (see corollary 3.3 on page 53 of [1]). When $\mu = -\frac{\pi}{2}$ a pair of roots $\pm i\frac{\pi}{2}$ are on the imaginary axis and all other roots have negative real parts. When $\mu < -\frac{\pi}{2}$ some of the roots have positive real part. Hence, the system (1.3) is on the verge of instability at $\mu = -\frac{\pi}{2}$. Close to the verge of instability, the behaviour of the solution is oscillatory with amplitude increasing

or decreasing depending on whether the root with the largest real part has positive real part or negative.

To study (1.1) close to the verge of instability, set $\mu = -\frac{\pi}{2} + \varepsilon^2 \tilde{\mu}$ and zoom-in near $x = 0$, i.e. write $y(t) = \varepsilon^{-1}x(t)$ for x governed by (1.1). We get

$$(1.4) \quad \begin{aligned} dy(t) = & -\frac{\pi}{2}y(t-1)dt + \varepsilon^2(y^3(t) + \tilde{\mu}y(t-1))dt \\ & + \varepsilon c_1 y(t-1)dV_1(t) + \varepsilon c_2 dV_2(t). \end{aligned}$$

The equations studied in this paper are of the above kind. Before stating the equations in more precise terms below, we describe briefly the motivation for studying such equations.

Delay equations at the verge of instability arise, for example, in machining processes [2], in the response of eye-pupil to incident light [3], in human balancing [4]. In machining processes, the motion of the cutting tool can be described by a DDE—the tool cuts a work-piece placed on a rotating shaft and the delay is the time-period of the rotating shaft. For each rotation period there is certain rate of cutting below which the tool is stable and above which the tool breaks. The inhomogenities in the properties of the material being cut can be modeled by noise (see [5]). The eye-pupil exhibits oscillations in response to incident light—however there is some delay in the response because neurons have finite processing speed. This phenomenon can be modeled using a DDE at the verge of instability [3]. So, a study of the effect of noise perturbations on ‘DDE at the verge of instability’ is indeed useful.

Now we describe the equations studied in this article in more precise terms.

Let $\{y_t : t \geq 0\}$ be an \mathbb{R} -valued process governed by an SDDE. Let $r > 0$ be the maximum of the delays involved in the drift and diffusion coefficients of the SDDE. To find the evolution at time t of the process, we need to keep track of y_s for $t-r \leq s \leq t$. For this purpose, let $\mathcal{C} := C([-r, 0]; \mathbb{R})$ be the space of \mathbb{R} -valued continuous functions on $[-r, 0]$, and equip \mathcal{C} with sup norm:

$$\|\eta\| := \sup_{\theta \in [-r, 0]} |\eta(\theta)|, \quad \text{for } \eta \in \mathcal{C}.$$

Define the *segment extractor* Π_t as follows: for $f \in C([-r, \infty); \mathbb{R})$,

$$(\Pi_t f)(\theta) = f(t + \theta), \quad \theta \in [-r, 0], \quad t \in [0, \infty).$$

Then, consider equation of the form:

$$(1.5) \quad dy(t) = L_0(\Pi_t y)dt + \varepsilon^2 G(\Pi_t y)dt + \varepsilon L_1(\Pi_t y)dV_1(t) + \varepsilon c_2 dV_2(t), \quad t \geq 0,$$

where $L_0, L_1 : \mathcal{C} \rightarrow \mathbb{R}$ are bounded linear operators, $G : \mathcal{C} \rightarrow \mathbb{R}$, and V_i are Wiener processes. Of course, as an initial condition we specify $\Pi_0 y = \xi \in \mathcal{C}$.

If we choose the maximum delay $r = 1$ and $L_0(\eta) = -\frac{\pi}{2}\eta(-1)$, $L_1(\eta) = c_1\eta(-1)$, $G(\eta) = \eta^3(0) + \tilde{\mu}\eta(-1)$, we see that (1.5) represents (1.4).

We make the following assumption on L_0 to reflect the Hopf-bifurcation scenario:

Assumption 1.1. We assume that the corresponding unperturbed DDE

$$(1.6) \quad \dot{x}(t) = L_0(\Pi_t x)$$

is critical, i.e. a pair of roots $(\pm i\omega_c)$ of the characteristic equation $\lambda - L_0(e^{\lambda \cdot}) = 0$ are on the imaginary axis (critical eigenvalues) and all other roots have negative real parts (stable eigenvalues).

Roughly speaking, under the above assumption, the solution of the unperturbed system (1.6) is oscillatory with constant amplitude. However for the perturbed system (1.5), it can be shown that for the amplitude of oscillation of y to change considerably, we need to wait for a time of order ε^{-2} . Hence we change the time scale, i.e. define $Y^\varepsilon(t) = y(t/\varepsilon^2)$.

To be able to put the rescaled process Y^ε in a form akin to (1.5) we need to define the *rescaled segment extractor* Π_t^ε as follows: for $f \in C([-\varepsilon^2 r, \infty); \mathbb{R})$,

$$(\Pi_t^\varepsilon f)(\theta) = f(t + \varepsilon^2 \theta), \quad \theta \in [-r, 0], \quad t \in [0, \infty).$$

Then, (1.5), with $Y^\varepsilon(t) = y(t/\varepsilon^2)$, can be written as

$$(1.7) \quad dY^\varepsilon(t) = \varepsilon^{-2}L_0(\Pi_t^\varepsilon Y^\varepsilon)dt + G(\Pi_t^\varepsilon Y^\varepsilon)dt + L_1(\Pi_t^\varepsilon Y^\varepsilon)dW_1(t) + c_2dW_2(t),$$

where $W_i(t) = \varepsilon V_i(t/\varepsilon^2)$ are again Wiener processes.

In this paper, we restrict to equations of the form:

$$(1.8) \quad \begin{cases} dX^\varepsilon(t) &= \varepsilon^{-2}L_0(\Pi_t^\varepsilon X^\varepsilon)dt + G(\Pi_t^\varepsilon X^\varepsilon)dt + \sigma dW(t), & t \in [0, T], \\ X^\varepsilon(t) &= \xi(\varepsilon^{-2}t), & t \in [-\varepsilon^2 r, 0], \quad \xi \in \mathcal{C}, \\ G(\eta) &= \int_{-r}^0 \eta(\theta) d\nu_1(\theta) + \int_{-r}^0 \eta^3(\theta) d\nu_3(\theta), \end{cases}$$

where for $i = 1, 3$, $\nu_i : [-r, 0] \rightarrow \mathbb{R}$, are bounded functions continuous from the left on $(-r, 0)$ and normalized with $\nu_i(0) = 0$; and also equations of the form:

$$(1.9) \quad \begin{cases} dX^\varepsilon(t) &= \varepsilon^{-2}L_0(\Pi_t^\varepsilon X^\varepsilon)dt + G(\Pi_t^\varepsilon X^\varepsilon)dt + L_1(\Pi_t^\varepsilon X^\varepsilon)dW(t), & t \in [0, T], \\ X^\varepsilon(t) &= \xi(\varepsilon^{-2}t), & t \in [-\varepsilon^2 r, 0], \quad \xi \in \mathcal{C}, \\ & |G(\eta_1) - G(\eta_2)| \leq K_G \|\eta_1 - \eta_2\|, & G(0) = 0. \end{cases}$$

We refer to (1.8) as the *additive noise case* and (1.9) as the *multiplicative noise case*. In both cases we assume that the initial condition ξ is deterministic (not a random variable).

Equations of the form $dX^\varepsilon(t) = \varepsilon^{-2}L_0(\Pi_t^\varepsilon X^\varepsilon)dt + G(\Pi_t^\varepsilon X^\varepsilon)dt + \sigma dW(t)$ were studied in [6] *but the coefficient G was assumed to be Lipschitz*. A quantity \mathcal{H}^ε was

identified which, roughly speaking, gives the amplitude of oscillations of X^ε . It was shown that the distribution of \mathcal{H}^ε converges weakly to the distribution of a process \mathcal{H}^0 governed by a stochastic differential equation (SDE) *without delay*. For small ε , this \mathcal{H}^0 gives good approximation for the dynamics of X^ε . The advantage is three fold: (i) equations without delay are easier to analyze, (ii) for numerical simulations, X^ε requires storage of $\Pi_t^\varepsilon X^\varepsilon$ (the entire segment) whereas \mathcal{H}^0 requires just the storage of current value \mathcal{H}_t^0 , (iii) numerical simulation of X^ε requires very small time-step for integration because the drift coefficient is of the order ε^{-2} , whereas \mathcal{H}^0 does not require such a small time-step.

In this article we relax the Lipschitz assumption on the coefficient G for the additive noise case. Note that the presence of ν_3 in (1.8) makes G non-Lipschitz. The process \mathcal{H}^ε mentioned above encodes information only about the critical eigenspace (space spanned by the eigenvectors corresponding to the imaginary roots of the characteristic equation), and to obtain the convergence to \mathcal{H}^0 one needs to show that the projection of X^ε onto stable eigenspace is small (details would be provided later). In [6] this was easy to show because of the Lipschitz condition on G . In this article we need to follow a different approach.

The case of multiplicative noise (1.9) is also considered here. However, for the multiplicative noise case the Lipschitz condition could not be relaxed. The presence of cubic nonlinearities causes the following problem: in trying to estimate a moment of certain order we encounter terms with higher order moments.

[7] discusses the approaches in the literature towards SDDE at the verge of instability and shows the mistakes and shortcomings of those approaches (see section 1 and appendix A of [7]). Hence, here we refrain from mentioning these works again.

Though here we discuss rigorously only \mathbb{R} -valued processes, the multi-dimensional processes are dealt with in [7] without proofs. An applications-oriented reader would benefit from [7] rather than this article.

Before stating the goals of this paper, we give a brief overview of the unperturbed system (1.6), and the variation-of-constants formula relating the solutions of (1.8) and (1.9) with (1.6). The material in section 1.1 can be found in chapter 7 of [8] (see also [9]).

1.1. The unperturbed system (1.6). The solution of (1.6) gives rise to the strongly continuous semigroup $T(t) : \mathcal{C} \rightarrow \mathcal{C}$, $t \geq 0$, defined by $T(t)\Pi_0 x = \Pi_t x$.

The state space \mathcal{C} can be split in the form $\mathcal{C} = P \oplus Q$ where $P = \text{span}_{\mathbb{R}}\{\Phi_1, \Phi_2\}$ where

$$\Phi_1(\theta) = \cos(\omega_c \theta), \quad \Phi_2(\theta) = \sin(\omega_c \theta), \quad \theta \in [-r, 0].$$

Write $\Phi = [\Phi_1, \Phi_2]$. Any $\eta \in P$ can be written as $\Phi z = z_1\Phi_1 + z_2\Phi_2$ for $z \in \mathbb{R}^2$, i.e. Φ is a basis for the two-dimensional space P and the z are coordinates of $\eta \in P$ with respect to the basis Φ .

Let π denote the projection of \mathcal{C} onto P along Q , i.e. $\pi : \mathcal{C} \rightarrow P$ with $\pi^2 = \pi$ and $\pi(\eta) = 0$ for $\eta \in Q$. The operator π can be written down explicitly, but we would not need the explicit form.

1.1.1. *Behaviour of the solution on P and Q .* It is easy to see that $\Pi_t x = \cos(\omega_c(t+\cdot))$ is a solution to (1.6) with the initial condition $\Pi_0 x = \cos(\omega_c \cdot)$, and $\Pi_t x = \sin(\omega_c(t+\cdot))$ is a solution to (1.6) with the initial condition $\Pi_0 x = \sin(\omega_c \cdot)$. Using the identity $\cos(\omega_c(t+\cdot)) = \cos(\omega_c t) \cos(\omega_c \cdot) - \sin(\omega_c t) \sin(\omega_c \cdot)$ and the linearity of L_0 , it can be shown that

$$(1.10) \quad T(t)\Phi(\cdot) = \Phi(\cdot)e^{Bt}, \quad B = \begin{bmatrix} 0 & \omega_c \\ -\omega_c & 0 \end{bmatrix}.$$

There exists positive constants κ and K such that

$$(1.11) \quad \|T(t)\phi\| \leq Ke^{-\kappa t}\|\phi\|, \quad \forall \phi \in Q.$$

The above is a consequence of the fact that, except for the roots $\pm i\omega_c$ all other roots of the characteristic equation have negative real parts.

Write the solution to (1.6) as

$$\Pi_t x = \pi \Pi_t x + (1 - \pi)\Pi_t x =: \Phi z(t) + y_t$$

where z is \mathbb{R}^2 -valued and y is \mathcal{C} valued. Then we find that¹ z oscillate harmonically according to $\dot{z}(t) = Bz(t)$ and $\|y_t\|$ decays exponentially fast, i.e.

$$(1.12) \quad \|y_t\| \leq Ke^{-\kappa t}\|y_0\|.$$

1.2. **The variation-of-constants formula.** The solution of the perturbed systems (1.8) or (1.9) can be expressed in terms of the solution of (1.6) with the initial condition $\Pi_0 x = \mathbf{1}_{\{0\}}$ where

$$\mathbf{1}_{\{0\}}(\theta) = \begin{cases} 1, & \theta = 0, \\ 0, & \theta \in [-r, 0). \end{cases}$$

However, note that $\mathbf{1}_{\{0\}}$ does not belong to \mathcal{C} and so we need to extend the space \mathcal{C} to accommodate the discontinuity.

See p. 192–193, 206–207 of [10] for the results pertaining to the extension. Let $\hat{\mathcal{C}} := \hat{C}([-r, 0]; \mathbb{R})$ be the Banach space of all bounded measurable maps $[-r, 0] \rightarrow \mathbb{R}$,

¹Multiply (1.10) by $z(0)$ and realize (using the fact T commutes with π) that $T(t)\Phi(\cdot)z(0) = T(t)\pi\Pi_0 x = \pi T(t)\Pi_0 x = \pi\Pi_t x = \Phi z(t)$ to get that $\Phi z(t) = \Phi e^{Bt}z(0)$ from which $\dot{z} = Bz$ follows.

given the sup norm. Solving the unperturbed system (1.6) for initial conditions in $\hat{\mathcal{C}}$, we can extend the semigroup T to one on $\hat{\mathcal{C}}$. Denote the extension by \hat{T} . Again $\hat{\mathcal{C}}$ splits in the form $\hat{\mathcal{C}} = P \oplus \hat{Q}$. The projection π can be extended to $\hat{\mathcal{C}}$. The extension is denoted by $\hat{\pi}$. There exists a two component column vector $\tilde{\Psi} \in \mathbb{R}^2$ such that

$$(1.13) \quad \hat{\pi} \mathbf{1}_{\{0\}} = \Phi \tilde{\Psi} = \tilde{\Psi}_1 \Phi_1 + \tilde{\Psi}_2 \Phi_2.$$

Also, there exists positive constants κ and K such that

$$(1.14) \quad \|\hat{T}(t)\phi\| \leq K e^{-\kappa t} \|\phi\|, \quad \forall \phi \in \hat{Q}.$$

1.2.1. *Additive noise case.* The solution of (1.8) can be represented as (see theorem IV.4.1 on page 200 in [10])

$$(1.15) \quad \begin{aligned} \Pi_t^\varepsilon X^\varepsilon &= \hat{T}(t/\varepsilon^2) \Pi_0^\varepsilon X^\varepsilon + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \\ &\quad + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) \mathbf{1}_{\{0\}} \sigma dW_u. \end{aligned}$$

The third term in the RHS of (1.15) is an element in \mathcal{C} and its value at $\theta \in [-r, 0]$ is given by $\int_0^t \left(\hat{T}\left(\frac{t-u}{\varepsilon^2}\right) \mathbf{1}_{\{0\}} \right) (\theta) \sigma dW_u$. Write

$$\Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon + y_t^\varepsilon.$$

Here $(y_t^\varepsilon)_{t \geq 0}$ is the \mathcal{C} -valued process $y_t^\varepsilon = (1 - \pi) \Pi_t^\varepsilon X^\varepsilon$ and $\Phi z_t = \pi \Pi_t^\varepsilon X^\varepsilon$. Note that z is \mathbb{R}^2 -valued process. Taking projection of (1.15) onto the space P , and using the facts (i) $\pi \Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon$, (ii) \hat{T} commutes with $\hat{\pi}$, (iii) $\hat{\pi} \mathbf{1}_{\{0\}} = \Phi \tilde{\Psi}$, (iv) $\hat{T}(t) \Phi z = \Phi e^{tB} z$, we get for z^ε (see corollary IV.4.1.1 on page 207 in [10])

$$(1.16) \quad dz_t^\varepsilon = \varepsilon^{-2} B z_t^\varepsilon dt + \tilde{\Psi} G(\Phi z_t^\varepsilon + y_t^\varepsilon) dt + \tilde{\Psi} \sigma dW_t, \quad \Phi z_0^\varepsilon = \pi \Pi_0^\varepsilon X^\varepsilon.$$

Using the fact that \hat{T} commutes with $\hat{\pi}$, y_t^ε satisfies

$$(1.17) \quad \begin{aligned} y_t^\varepsilon &= \hat{T}(t/\varepsilon^2) y_0^\varepsilon + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} G(\Phi z_u^\varepsilon + y_u^\varepsilon) du \\ &\quad + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} \sigma dW_u. \end{aligned}$$

1.2.2. *Multiplicative noise case.* The solution of (1.9) can be represented in a form analogous to (1.15) with σ replaced by $L_1(\Pi_u^\varepsilon X^\varepsilon)$ (see [11]):

$$(1.18) \quad \begin{aligned} \Pi_t^\varepsilon X^\varepsilon &= \hat{T}(t/\varepsilon^2) \Pi_0^\varepsilon X^\varepsilon + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \\ &\quad + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) \mathbf{1}_{\{0\}} L_1(\Pi_u^\varepsilon X^\varepsilon) dW_u. \end{aligned}$$

For the projections onto P and Q we have:

$$(1.19) \quad dz_t^\varepsilon = \varepsilon^{-2} B z_t^\varepsilon dt + \tilde{\Psi} G(\Phi z_t^\varepsilon + y_t^\varepsilon) dt + \tilde{\Psi} L_1(\Phi z_t^\varepsilon + y_t^\varepsilon) dW, \quad \Phi z_0^\varepsilon = \pi \Pi_0^\varepsilon X^\varepsilon,$$

$$(1.20) \quad y_t^\varepsilon = \hat{T}(t/\varepsilon^2)y_0^\varepsilon + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) (1 - \hat{\pi})\mathbf{1}_{\{0\}}G(\Phi z_u^\varepsilon + y_u^\varepsilon)du + \int_0^t \hat{T}\left(\frac{t-u}{\varepsilon^2}\right) (1 - \hat{\pi})\mathbf{1}_{\{0\}}L_1(\Phi z_u^\varepsilon + y_u^\varepsilon)dW_u.$$

Crucial role would be played in proofs by

$$(1.21) \quad \gamma(t) := (\hat{T}(t)(1 - \hat{\pi})\mathbf{1}_{\{0\}})(0).$$

From (1.14) we already know that

$$(1.22) \quad |\gamma(t)| \leq Ke^{-\kappa t}\|(1 - \hat{\pi})\mathbf{1}_{\{0\}}\|, \quad t \geq 0.$$

Further, for $t > 0$

$$(1.23) \quad \left| \frac{d}{dt}\gamma(t) \right| = \left| L_0\left(\hat{T}(t)(1 - \hat{\pi})\mathbf{1}_{\{0\}}\right) \right| \leq \|L_0\|_{op} \|\hat{T}(t)(1 - \hat{\pi})\mathbf{1}_{\{0\}}\| \leq \|L_0\|_{op} \|(1 - \hat{\pi})\mathbf{1}_{\{0\}}\| Ke^{-\kappa t}.$$

Thus, both γ and γ' have exponential decay.

1.3. Goal of this paper. Let X^ε evolve according to either (1.8) or (1.9). Write $\Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon + y_t^\varepsilon$, and define

$$(1.24) \quad \mathfrak{Y}_t^\varepsilon := \hat{T}(t/\varepsilon^2)y_0^\varepsilon, \quad \mathfrak{Z}_t^\varepsilon := y_t^\varepsilon - \mathfrak{Y}_t^\varepsilon.$$

Note that $\mathfrak{Y}_t^\varepsilon$ depends only on the unperturbed system (1.6). Given the initial condition $\Pi_0^\varepsilon X^\varepsilon$, $\mathfrak{Y}_t^\varepsilon$ is a deterministic quantity. Note that $\|\mathfrak{Y}_t^\varepsilon\|$ decays exponentially fast:

$$(1.25) \quad \|\mathfrak{Y}_t^\varepsilon\| \leq Ke^{-\kappa t/\varepsilon^2}\|(1 - \pi)\Pi_0^\varepsilon X^\varepsilon\|.$$

1.3.1. *Goal for the multiplicative noise case* (1.9). Roughly speaking, the goals are

- (i) Show that, until time $T > 0$, $\mathbb{E} \sup_{t \in [0, T]} \|\mathfrak{Z}_t^\varepsilon\|^n \xrightarrow{\varepsilon \rightarrow 0} 0$, so that we can approximate y_t^ε with the deterministic quantity $\mathfrak{Y}_t^\varepsilon$
- (ii) Consider the process

$$(1.26) \quad dz_t^\varepsilon = \varepsilon^{-2}Bz_t^\varepsilon dt + \tilde{\Psi}G(\Phi z_t^\varepsilon)dt + \tilde{\Psi}L_1(\Phi z_t^\varepsilon)dW, \quad \Phi z_0^\varepsilon = \pi\Pi_0^\varepsilon X^\varepsilon.$$

Note that \mathbf{z}^ε is two-dimensional process without delay totally ignoring y^ε . Show that

$$(1.27) \quad \mathbb{E} \sup_{t \in [0, T]} \|\mathbf{z}_t^\varepsilon - z_t^\varepsilon\|_2^2 \xrightarrow{\varepsilon \rightarrow 0} 0$$

where $\|\cdot\|_2$ is ℓ_2 norm of vectors in \mathbb{R}^2 .

The above tasks justify the approximation of $\Pi_t^\varepsilon X^\varepsilon$ by $\Phi z_t^\varepsilon + \mathfrak{Y}_t^\varepsilon$ for small ε . Note that \mathbf{z}^ε is a two-dimensional process without delay and $\mathfrak{Y}_t^\varepsilon$ is an exponentially

decaying deterministic process. For small ε one could study this non-delay system instead of the original stochastic DDE (1.9). The advantage is that the 2-dimensional system without delay would be easier to analyze or simulate numerically.

The approximation result is stated in theorem 2.8. The process $\tilde{\mathfrak{z}}_t^\varepsilon$ in (2.8) is related to the process \mathbf{z}_t^ε of (1.26) by $e^{tB/\varepsilon^2}\tilde{\mathfrak{z}}_t^\varepsilon = \mathbf{z}_t^\varepsilon$.

Further simplification can be obtained by studying the process

$$(1.28) \quad \mathcal{H}_t^\varepsilon := \frac{1}{2} \|z_t^\varepsilon\|_2^2.$$

Roughly speaking², $\sqrt{2\mathcal{H}^\varepsilon}$ is the amplitude of oscillations of X^ε . We will show that there is a constant C such that

$$(1.29) \quad \mathbb{E} \sup_{t \in [0, T]} \|\mathbf{z}_t^\varepsilon\|_2^2 < C$$

for all ε smaller than some ε_* . Using (1.29) and (1.27) it follows that

$$(1.30) \quad \mathbb{E} \sup_{t \in [0, T]} |\mathcal{H}_t^\varepsilon - \frac{1}{2} \|z_t^\varepsilon\|_2^2| \xrightarrow{\varepsilon \rightarrow 0} 0.$$

One can use standard averaging techniques for stochastic differential equations (without delay) to show that the distribution of $\frac{1}{2} \|z^\varepsilon\|_2^2$ converges to the distribution of some one-dimensional process \mathcal{H}^0 without delay. By theorem 3.1 in [12], the distribution of \mathcal{H}^ε converges to the distribution of \mathcal{H}^0 . For small ε , the distribution of \mathcal{H}^0 gives a good approximation to the distribution of \mathcal{H}^ε . The advantages of having \mathcal{H}^0 were mentioned in section 1.

The above result concerning \mathcal{H}^ε is stated in theorem 2.10. The process $\tilde{\mathfrak{z}}_t^\varepsilon$ in (2.10) is related to \mathbf{z}_t^ε of (1.26) by $\mathbf{z}_t^\varepsilon = e^{tB/\varepsilon^2}\tilde{\mathfrak{z}}_t^\varepsilon$.

1.3.2. *Goal for the additive noise case* (1.8). The presence of cubic nonlinearities causes the following problem: in trying to estimate a moment of certain order we face the task of estimating terms with higher order moments. So the approach taken for (1.9) does not work here. We take the following approach.

Recall the projection operator $\pi : \mathcal{C} \rightarrow P$. Fix a constant $C_\varepsilon > 0$ and define the stopping time $\mathbf{e}^\varepsilon = \inf\{t \geq 0 : \|\pi \Pi_t^\varepsilon X^\varepsilon\| \geq C_\varepsilon\}$. (Note that the stopping time depends on ε).

- Show that for $t \in [0, T \wedge \mathbf{e}^\varepsilon]$, $\|\mathcal{Y}_t^\varepsilon\|$ is small with high probability
- Define a 2-dimensional process \mathbf{z}^ε as

$$(1.31) \quad d\mathbf{z}_t^\varepsilon = \varepsilon^{-2} B \mathbf{z}_t^\varepsilon dt + \tilde{\Psi} G(\Phi \mathbf{z}_t^\varepsilon) dt + \tilde{\Psi} \sigma dW, \quad \Phi \mathbf{z}_0^\varepsilon = \pi \Pi_0^\varepsilon X^\varepsilon.$$

²Writing $\Pi_t^\varepsilon X^\varepsilon = \Phi z(t) + y_t$ and showing y is small, we can write $X^\varepsilon(t) = \Pi_t^\varepsilon X^\varepsilon(0) \approx \Phi(0)z_t = (z_t)_1$. Since the dynamics of z is small perturbation of a predominant oscillation according $\dot{z} = Bz$, the approximate amplitude of $(z)_1$ is $\sqrt{(z)_1^2 + (z)_2^2} = \|z\|_2 = \sqrt{2\mathcal{H}}$.

Note that \mathbf{z}^ε is a 2 dimensional process without delay. Show that for $t \in [0, T \wedge \varepsilon^\varepsilon]$, error in approximating z^ε by \mathbf{z}^ε is small with high probability

- Using estimates on \mathbf{z}^ε process, get rid of the stopping time and obtain approximation results until time T , by leveraging some arbitrarily small probability.

The stopping time helps in arriving at a bound on the norm of stable-mode $(1 - \pi)\Pi_t^\varepsilon X^\varepsilon$ without worrying about what happens to the critical-mode $\pi\Pi_t^\varepsilon X^\varepsilon$.

The approximation result is stated in theorem 3.10. The process $\tilde{\mathfrak{Z}}_t^\varepsilon$ in (3.9) is related to the process \mathbf{z}_t^ε of (1.31) by $e^{tB/\varepsilon^2} \tilde{\mathfrak{Z}}_t^\varepsilon = \mathbf{z}_t^\varepsilon$.

Examples illustrating the usefulness of the above approximation results are shown in sections 2.1 and 3.1.

For related work on stochastic partial differential equations see [13]. However note that in [13] the bifurcation scenario is different—analogous situation in the DDE framework would be if one root of characteristic equation is zero and all other roots have negative real parts.

2. MULTIPLICATIVE NOISE

In this section we consider (1.9) with $T > 0$ fixed. The constants here can depend on T .

The first goal is to show that $\mathbb{E} \sup_{t \in [0, T]} \|\mathcal{Y}_t^\varepsilon\|^n \rightarrow 0$, which is the content of proposition 2.4. For this purpose, we use the variation of constants formulas (1.18)–(1.20). Recalling the definition of \mathcal{Y}^ε from (1.24), to estimate $\mathbb{E} \sup_{t \in [0, T]} \|\mathcal{Y}_t^\varepsilon\|^n$, we need to estimate the last two terms on the RHS of (1.20).

Roughly speaking, the integral in the last term of RHS of (1.20) can be split as $\int_0^s = \int_0^{s-\varepsilon^\delta r} + \int_{s-\varepsilon^\delta r}^s$ with $0 < \delta < 2$. For $\int_0^{s-\varepsilon^\delta r}$ we can use exponential decay of \hat{T} on \hat{Q} . For $\int_{s-\varepsilon^\delta r}^s$, making note that the length of the interval of integration is small ($r\varepsilon^\delta$), we need to be concerned with increments of Wiener process over small intervals, i.e. the modulus of continuity of the Wiener process.

Lemma 2.1 is needed to be able to use the results of [14] on moments of modulus-of-continuity of Ito processes. Using results from [14], proposition 2.3 shows that the stochastic term in (1.20) is small. Then, straight forward estimation yields proposition 2.4 which is the result that we need.

Lemma 2.1. *Fix $n \geq 0$. There exists constants $\mathfrak{C} > 0$ and $\varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$,*

$$(2.1) \quad \mathbb{E} \sup_{t \in [0, T]} \|\Pi_t^\varepsilon X^\varepsilon\|^n < \mathfrak{C}.$$

Proof is given in appendix A.1. Note that, though one of the drift coefficients in (1.9) is of the order ε^{-2} , the constant \mathfrak{C} above does not depend on ε . Proof uses: (i) the variation-of-constants formula to exploit the exponential decay (1.22) and

(1.23) on \hat{Q} , and oscillatory behaviour on P ; (ii) Burkholder-Davis-Gundy inequality to estimate supremum of martingales by their quadratic-variation; and then (iii) Gronwall inequality.

Definition 2.2. Define the modulus of continuity for $f : [0, \infty) \rightarrow \mathbb{R}$ by

$$\mathfrak{w}(a, b; f) = \sup_{\substack{|u-v| \leq a \\ u, v \in [0, b]}} |f(u) - f(v)|.$$

Define

$$(2.2) \quad \Upsilon_s^\varepsilon := \sup_{\theta \in [-r, 0]} \left| \int_0^s \left(\hat{T}\left(\frac{s-u}{\varepsilon^2}\right)(1 - \hat{\pi})\mathbf{1}_{\{0\}} \right) (\theta) dZ_u \right|, \quad Z_t := \int_0^t L_1(\Pi_s^\varepsilon X^\varepsilon) dW_s.$$

Note that Z depends also on ε .

Proposition 2.3. Fix $n \geq 1$. There exists constant $\hat{C} > 0$ and a family of constants $\hat{\varepsilon}_\delta > 0$ (indexed by $0 < \delta < 2$) such that, given $\delta \in (0, 2)$ we have for $\varepsilon < \hat{\varepsilon}_\delta$

$$(2.3) \quad \mathbb{E} \sup_{s \in [0, T]} (\Upsilon_s^\varepsilon)^n \leq \hat{C} \left(r \varepsilon^\delta \ln \left(\frac{2T}{r \varepsilon^\delta} \right) \right)^{n/2} \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Proof is given in appendix A.2. The essential idea of writing $\int_0^s = \int_0^{s-\varepsilon^\delta r} + \int_{s-\varepsilon^\delta r}^s$ and using [14] is mentioned earlier.

Let $\mathfrak{Y}_t^\varepsilon$ and $\mathfrak{Z}_t^\varepsilon$ be as defined in (1.24).

Proposition 2.4. Fix $n \geq 1$. $\exists \varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$,

$$(2.4) \quad \mathbb{E} \sup_{s \in [0, T]} \|\mathfrak{Y}_s^\varepsilon\|^n \leq \varepsilon^{2n} 2^{n-1} \left(\frac{K_G K}{\kappa} \right)^n \mathfrak{C} + 2^{n-1} \mathbb{E} \sup_{s \in [0, T]} (\Upsilon_s^\varepsilon)^n \xrightarrow{\varepsilon \rightarrow 0} 0,$$

where \mathfrak{C} is from lemma 2.1.

Proof given in appendix A.3.

Recall that when we write $\Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon + y_t^\varepsilon$, the \mathbb{R}^2 -valued process z^ε satisfies equation (1.19). Removing the fast rotation induced by B , i.e. writing $\mathfrak{z}_t^\varepsilon = e^{-tB/\varepsilon^2} z_t^\varepsilon$ we have

$$d\mathfrak{z}^\varepsilon = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \tilde{\Psi} L_1(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) dW_t, \quad \mathfrak{z}_0^\varepsilon = z_0^\varepsilon.$$

Let $\widehat{\mathfrak{z}}^\varepsilon$ be governed by

$$d\widehat{\mathfrak{z}}^\varepsilon = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \tilde{\Psi} L_1(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon) dW_t, \quad \widehat{\mathfrak{z}}_0^\varepsilon = \mathfrak{z}_0^\varepsilon.$$

i.e. we are totally ignoring y part except for the effect of initial condition (note that $\mathfrak{Y}_t^\varepsilon = \hat{T}(t/\varepsilon^2)y_0^\varepsilon$).

As an intermediate step towards the end goal, we want to show that, until time T the error in approximating \mathfrak{z}^ε by $\widehat{\mathfrak{z}}^\varepsilon$ is small. For this purpose, define

$$\alpha_t^\varepsilon = \frac{1}{2} \|\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon\|_2^2 = \frac{1}{2} ((\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1^2 + (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2^2).$$

Here $(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_i$ denotes the i^{th} component of the \mathbb{R}^2 -valued vector $\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon$. Let

$$(2.5) \quad \Gamma_t = \sum_{i=1}^2 (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_i (e^{-tB/\varepsilon^2} \tilde{\Psi})_i.$$

Then α_t^ε is governed by

$$d\alpha_t^\varepsilon = \mathcal{B}_t dt + \Sigma_t dW_t, \quad \alpha_0^\varepsilon = 0,$$

where

$$\begin{aligned} \mathcal{B}_t &= \Gamma_t \left(G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) - G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon) \right) \\ &\quad + \frac{1}{2} \|\tilde{\Psi}\|_2^2 \left(L_1(\Phi e^{tB/\varepsilon^2} (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)) + L_1(y_t^\varepsilon - \mathfrak{Y}_t^\varepsilon) \right)^2, \end{aligned}$$

and

$$\Sigma_t = \Gamma_t \left(L_1(\Phi e^{tB/\varepsilon^2} (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)) + L_1(y_t^\varepsilon - \mathfrak{Y}_t^\varepsilon) \right).$$

The following lemma gives processes dominating \mathcal{B}_t and Σ_t . These help in applying Gronwall inequality to arrive at proposition 2.6.

Lemma 2.5. *Define*

$$\begin{aligned} \mathfrak{B}(\alpha, p) &:= C_{\mathcal{B}}(\alpha + p^2), & C_{\mathcal{B}} &= 2\|\tilde{\Psi}\|_2^2 \|L_1\|^2 + 3\|\tilde{\Psi}\|_2 K_G, \\ \mathfrak{S}^2(\alpha, p) &:= C_{\Sigma}(\alpha^2 + p^4), & C_{\Sigma} &= 16\|\tilde{\Psi}\|_2^2 \|L_1\|^2. \end{aligned}$$

Then $|\mathcal{B}_t| \leq \mathfrak{B}(\alpha_t^\varepsilon, \|\mathfrak{Y}_t^\varepsilon\|)$ and $\Sigma_t^2 \leq \mathfrak{S}^2(\alpha_t^\varepsilon, \|\mathfrak{Y}_t^\varepsilon\|)$ for $t \geq 0$.

Proof given in appendix A.4

Proposition 2.6. *Fix $\delta \in (0, 2)$. There exists constants $C, \hat{\varepsilon}_\delta > 0$ such that $\forall \varepsilon < \hat{\varepsilon}_\delta$*

$$\mathbb{E} \sup_{s \in [0, T]} (\alpha_s^\varepsilon)^2 \leq C \left(r\varepsilon^\delta \ln\left(\frac{2T}{r\varepsilon^\delta}\right) \right)^2 \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Proof is given in appendix A.5 and is by using lemma 2.5, result (2.4), applying Gronwall pathwise (see [15]) and Doob's L^p inequalities.

As final step, consider the system

$$(2.6) \quad d\widetilde{\mathfrak{z}}_t^\varepsilon = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \widetilde{\mathfrak{z}}_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \tilde{\Psi} L_1(\Phi e^{tB/\varepsilon^2} \widetilde{\mathfrak{z}}_t^\varepsilon) dW, \quad \widetilde{\mathfrak{z}}_0^\varepsilon = z_0^\varepsilon,$$

i.e. we are totally ignoring the Q part—even the effect \mathfrak{Y} of the initial condition.

Define $\beta_t^\varepsilon = \frac{1}{2} \|\widetilde{\mathfrak{z}}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon\|_2^2$. Using exactly the same technique as the one employed for

α^ε and using the exponential decay of \mathfrak{Y}^ε , it is trivial to get the following result analogous to proposition 2.6.

Proposition 2.7. *There exists $C > 0$ and $\varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$*

$$\mathbb{E} \sup_{s \in [0, T]} (\beta_s^\varepsilon)^2 \leq C\varepsilon^2.$$

Proof given in appendix A.6.

Combining propositions 2.4, 2.6 and 2.7 we get

Theorem 2.8. *Fix $\delta \in (0, 2)$. There exists constants $C, \hat{\varepsilon}_\delta > 0$ such that $\forall \varepsilon < \hat{\varepsilon}_\delta$*

$$(2.7) \quad \mathbb{E} \sup_{t \in [0, T]} |X^\varepsilon(t) - (\Phi(0)e^{tB/\varepsilon^2} \widehat{\mathfrak{Z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon(0))|^4 \leq C \left(r\varepsilon^\delta \ln\left(\frac{2T}{r\varepsilon^\delta}\right) \right)^2 \xrightarrow{\varepsilon \rightarrow 0} 0.$$

There exists constants $C > 0$ and $\varepsilon_ > 0$ such that $\forall \varepsilon < \varepsilon_*$*

$$(2.8) \quad \mathbb{E} \sup_{t \in [0, T]} |X^\varepsilon(t) - (\Phi(0)e^{tB/\varepsilon^2} \widetilde{\mathfrak{Z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon(0))|^4 \leq C\varepsilon^2.$$

Proof given in appendix A.7

Note that both the *approximating processes* $\widehat{\mathfrak{Z}}^\varepsilon$ and $\widetilde{\mathfrak{Z}}^\varepsilon$ are *processes without delay*. However, $\widehat{\mathfrak{Z}}^\varepsilon$ considers the effect of the initial condition y_0^ε , but $\widetilde{\mathfrak{Z}}^\varepsilon$ ignores it. Hence the approximation (2.7) using $\widehat{\mathfrak{Z}}^\varepsilon$ is better than the approximation (2.8). For example, choosing δ close to two in (2.7) we can get the bound $O(\varepsilon^{4-})$ whereas the bound in (2.8) is $O(\varepsilon^2)$.

Now we revisit the goals stated in section 1.3.1.

Note that for \mathbf{z}_t^ε defined in (1.26) we have $\mathbf{z}_t^\varepsilon = e^{tB/\varepsilon^2} \widetilde{\mathfrak{Z}}_t^\varepsilon$. Hence, $\mathbf{z}_t^\varepsilon - z_t^\varepsilon = e^{tB/\varepsilon^2} (\widetilde{\mathfrak{Z}}_t^\varepsilon - \mathfrak{Z}_t^\varepsilon)$. Using the results of this section and the fact that for any \mathbb{R}^2 -vector v , $\|e^{tB/\varepsilon^2} v\|_2 = \|v\|_2$, we can easily see that (1.27) is satisfied. The condition (1.29) is equivalent to the following condition (2.9). Lemma 2.9 is proved in appendix A.8.

Lemma 2.9. *There exists constants C and $\varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$*

$$(2.9) \quad \mathbb{E} \sup_{t \in [0, T]} \|\widetilde{\mathfrak{Z}}_t^\varepsilon\|_2^2 < C.$$

Hence, (1.30) follows. We summarize the discussion in section 1.3.1 in the following theorem.

Theorem 2.10. *Define $\mathcal{H}_t^\varepsilon := \frac{1}{2} \|z_t^\varepsilon\|_2^2$ where z^ε are given by $\pi \Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon$. Let $\widetilde{\mathfrak{Z}}^\varepsilon$ be the two-dimensional process (without delay) defined in (2.6). Then*

$$(2.10) \quad \mathbb{E} \sup_{t \in [0, T]} \left| \mathcal{H}_t^\varepsilon - \frac{1}{2} \|\widetilde{\mathfrak{Z}}_t^\varepsilon\|_2^2 \right| \xrightarrow{\varepsilon \rightarrow 0} 0.$$

If the process $\frac{1}{2} \|\widetilde{\mathfrak{Z}}^\varepsilon\|_2^2$ converges weakly to a process \mathcal{H}^0 , then \mathcal{H}^ε converges weakly to \mathcal{H}^0 .

Remark 2.1. Because $\tilde{\mathfrak{z}}^\varepsilon$ is a process without delay, weak convergence of $\frac{1}{2}\|\tilde{\mathfrak{z}}^\varepsilon\|_2^2$ can be dealt using standard averaging techniques for stochastic differential equations.

2.1. **Example.** Consider (1.9) with $G \equiv 0$. The corresponding $\tilde{\mathfrak{z}}^\varepsilon$ satisfies

$$d\tilde{\mathfrak{z}}^\varepsilon = e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon dW, \quad M = \tilde{\Psi} L_1 \Phi,$$

with $\tilde{\mathfrak{z}}_0^\varepsilon$ such that $\Phi \tilde{\mathfrak{z}}_0^\varepsilon = \pi \xi$. Let $\mathcal{H}_t^\varepsilon := \frac{1}{2} \|z_t^\varepsilon\|_2^2$ and $H_t^\varepsilon := \frac{1}{2} \|\tilde{\mathfrak{z}}_t^\varepsilon\|_2^2$. Then applying Ito formula we have

$$(2.11) \quad dH_t^\varepsilon = (\tilde{\mathfrak{z}}_t^\varepsilon)^* e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon dW_t + \frac{1}{2} \left(\left(e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon \right)_1^2 + \left(e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon \right)_2^2 \right) dt.$$

Averaging out the fast oscillations of $\tilde{\mathfrak{z}}^\varepsilon$, it can be shown that³ as $\varepsilon \rightarrow 0$, the distribution of H^ε converges weakly to the distribution of

$$(2.12) \quad d\mathcal{H}_t^0 = C_1 \mathcal{H}_t^0 dW_t + C_2 \mathcal{H}_t^0 dt,$$

where $2C_1^2 = 3(M_{11}^2 + M_{22}^2) + (M_{12} + M_{21})^2 + 2M_{11}M_{22}$ and $C_2 = \frac{1}{2}(\sum_{i,j=1}^2 M_{ij}^2)$. Using $M = \tilde{\Psi} L_1 \Phi$ we get

$$(2.13) \quad d\mathcal{H}_t^0 = \sqrt{\left(\frac{1}{2} \|\tilde{\Psi}\|_2^2 \|L_1 \Phi\|_2^2 + (L_1 \Phi \tilde{\Psi})^2 \right)} \mathcal{H}_t^0 dW_t + \frac{1}{2} \|\tilde{\Psi}\|_2^2 \|L_1 \Phi\|_2^2 \mathcal{H}_t^0 dt.$$

For (2.13) solution can be written explicitly. For small ε , the distribution of \mathcal{H}^0 gives good approximation to the distribution of \mathcal{H}^ε . Note that, roughly speaking, $\sqrt{2\mathcal{H}^\varepsilon}$ is the amplitude of oscillations of X^ε . Hence \mathcal{H}^0 can be used to understand the dynamics of X^ε . The advantage is that \mathcal{H}^0 does not involve any delay and is one-dimensional, and hence easier to analyze and simulate numerically (see [7] for examples involving numerical simulations).

3. ADDITIVE NOISE

In this section we consider (1.8) with $T > 0$ fixed. The constants here can depend on T .

The strategy employed for (1.9) in the previous section does not work for (1.8) due to the problem of moment-closure, i.e. in trying to estimate a lower moment we end up with the task of estimating a higher moment (because of the cubic nonlinearity). For (1.8) we employ the strategy stated in section 1.3.2.

Define

$$(3.1) \quad \Upsilon_s^\varepsilon := \sup_{\theta \in [-r, 0]} \left| \int_0^s \left(\hat{T} \left(\frac{s-u}{\varepsilon^2} \right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} \right) (\theta) dZ_u \right|, \quad Z_t := \sigma W_t.$$

³If $\frac{1}{2} \|\tilde{\mathfrak{z}}^\varepsilon\|_2^2 = \mathcal{H}^0$ then $\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^\varepsilon \frac{1}{2} \|e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}^\varepsilon\|_2^2 dt = C_2 \mathcal{H}^0$ and $\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^\varepsilon \left((\tilde{\mathfrak{z}}^\varepsilon)^* e^{-tB/\varepsilon^2} M e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}^\varepsilon \right)^2 dt = (C_1 \mathcal{H}^0)^2$.

Proposition 3.1. *Fix $n \geq 1$. There exists constant $\hat{C} > 0$ and a family of constants $\hat{\varepsilon}_\delta > 0$ (indexed by $0 < \delta < 2$) such that, given $\delta \in (0, 2)$ we have for $\varepsilon < \hat{\varepsilon}_\delta$*

$$(3.2) \quad \mathbb{E} \sup_{s \in [0, T]} (\Upsilon_s^\varepsilon)^n \leq \hat{C} \left(r\varepsilon^\delta \ln \left(\frac{2T}{r\varepsilon^\delta} \right) \right)^{n/2} \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Proof is same as that of proposition 2.3 with appropriate changes to account for $Z_t = \sigma W_t$; and we dont need anything analogous to lemma 2.1.

Fix a constant $C_\varepsilon > 0$ and define the stopping time

$$(3.3) \quad \mathbf{e}^\varepsilon = \inf\{t \geq 0 : \|\pi \Pi_t^\varepsilon X^\varepsilon\| \geq C_\varepsilon\}.$$

The stopping time helps in arriving at a bound on the norm of stable-mode $(1 - \pi)\Pi_t^\varepsilon X^\varepsilon$ (until time $T \wedge \mathbf{e}^\varepsilon$) without worrying about what happens to the critical-mode $\pi \Pi_t^\varepsilon X^\varepsilon$. Hence, as an intermediate step we establish results that hold until time $T \wedge \mathbf{e}^\varepsilon$ and later get rid of the stopping time \mathbf{e}^ε .

Write $\Pi_t^\varepsilon X^\varepsilon = \Phi z_t^\varepsilon + y_t^\varepsilon$. Then z_t^ε and y_t^ε satisfy the variation-of-constants formula (1.16) and (1.17). Define \mathfrak{Y}^ε and \mathfrak{Z}^ε as in (1.24).

Proposition 3.2. *Let \hat{C} and $\hat{\varepsilon}_\delta$ be the same as in proposition 3.1. There exists a family of constants $\varepsilon_{a, C_\varepsilon} > 0$ such that, given $a \in [0, 1)$ and $\delta \in (2a, 2)$, we have for $\varepsilon < \min\{\hat{\varepsilon}_\delta, \varepsilon_{a, C_\varepsilon}\}$*

$$(3.4) \quad \mathbb{P} \left[\sup_{s \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\mathfrak{Z}_s^\varepsilon\| \leq 8\varepsilon^a \right] \geq 1 - \hat{C}\varepsilon^{-a} \sqrt{r\varepsilon^\delta \ln \left(\frac{T}{r\varepsilon^\delta} \right)}.$$

Here $\varepsilon_{a, C_\varepsilon}$ is of the order $\mathcal{O}(\min\{C_\varepsilon^{-3/(2-a)}, C_\varepsilon^{-3/2a}\})$ for large C_ε .

In (3.4) we obtain a bound on $\|\mathfrak{Z}^\varepsilon\|$ which does not depend on C_ε in spite of the cubic nonlinearity—hence the ε should be made really small. Larger the C_ε , smaller the ε we need to consider. Proof is by straight forward application of exponential decay on \hat{Q} , Markov and Gronwall inequalities. Proof is given in appendix B.1

Removing the fast rotation induced by B , i.e. writing $\mathfrak{z}_t^\varepsilon = e^{-tB/\varepsilon^2} z_t^\varepsilon$ we have

$$d\mathfrak{z}_t^\varepsilon = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \tilde{\Psi} \sigma dW_t, \quad \mathfrak{z}_0^\varepsilon = z_0^\varepsilon.$$

Let $\widehat{\mathfrak{z}}^\varepsilon$ be governed by

$$d\widehat{\mathfrak{z}}_t^\varepsilon = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \tilde{\Psi} \sigma dW_t, \quad \widehat{\mathfrak{z}}_0^\varepsilon = z_0^\varepsilon,$$

i.e., in $\widehat{\mathfrak{z}}^\varepsilon$ we are totally ignoring y part except for the effect of the initial condition ($\mathfrak{Y}_t^\varepsilon = \hat{T}(t/\varepsilon^2)y_0^\varepsilon$). Note that $\widehat{\mathfrak{z}}^\varepsilon$ is a process without delay.

We want to show that until time $T \wedge \epsilon^\varepsilon$, error in approximating \mathfrak{z}^ε by $\widehat{\mathfrak{z}}^\varepsilon$ is small. For this purpose, define

$$\alpha_t^\varepsilon = \frac{1}{2} \|\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon\|_2^2 = \frac{1}{2} ((\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1^2 + (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2^2).$$

and let $\Gamma_t = \left(\sum_{i=1}^2 (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_i (e^{-tB/\varepsilon^2} \tilde{\Psi})_i \right)$. Then α_t^ε is governed by

$$d\alpha_t^\varepsilon = \mathcal{B}_t dt, \quad \alpha_0^\varepsilon = 0,$$

where

$$\mathcal{B}_t = \Gamma_t \left(G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t) - G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{y}_t^\varepsilon) \right).$$

The following lemma gives a process dominating \mathcal{B}_t . This helps in applying Gronwall inequality to arrive at proposition 3.4.

Lemma 3.3. $\exists C > 0$ (is of the order $\mathcal{O}(C_\epsilon^2)$ for large C_ϵ) such that if \mathfrak{B} is defined by

$$(3.5) \quad \mathfrak{B}(\alpha, p) := C\sqrt{2\alpha} \sum_{j=1}^3 \left(p^j + (\sqrt{2\alpha})^j \right),$$

then $|\mathcal{B}_t| \leq \mathfrak{B}(\alpha_t, \|\mathfrak{y}_t^\varepsilon\|)$ for $t \in [0, T \wedge \epsilon^\varepsilon]$.

Proof given in appendix B.2.

Proposition 3.4. Let \hat{C} and $\hat{\varepsilon}_\delta$ be the same as in proposition 3.1. There exists two families of constants $\varepsilon_{a,C_\epsilon,1} > 0$, $\varepsilon_{a,C_\epsilon,2} > 0$ such that, given $a \in [0, 1)$ and $\delta \in (2a, 2)$, we have for $\varepsilon < \min\{\hat{\varepsilon}_\delta, \varepsilon_{a,C_\epsilon,1}, \varepsilon_{a,C_\epsilon,2}\}$

$$(3.6) \quad \mathbb{P} \left[\sup_{t \in [0, T \wedge \epsilon^\varepsilon]} \alpha_t^\varepsilon \leq \varepsilon^{a/2} \right] \geq 1 - \hat{C} \varepsilon^{-a} \sqrt{r \varepsilon^\delta \ln \left(\frac{T}{r \varepsilon^\delta} \right)} =: p_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} 1.$$

Here $\varepsilon_{a,C_\epsilon,1}$ is of the order $\mathcal{O}(\min\{C_\epsilon^{-3/(2-a)}, C_\epsilon^{-3/2a}\})$ (these are from proposition 3.2) and $\varepsilon_{a,C_\epsilon,2}$ is of the order $\mathcal{O}(\exp(-30C_\epsilon^2 T/a))$ for large C_ϵ .

Proof is given in appendix B.3.

Finally, the stopping time ϵ^ε can be got rid as follows.

Let Ω be the set of all realizations of the Brownian motion W and $\omega \in \Omega$ denote one particular realization.

Definition 3.5. Given $T > 0$ and $q > 0$, we say that “ $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$ ” if $\exists C_\epsilon, \varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$, we have $\mathbb{P}[E^\varepsilon] \geq 1 - q$ where

$$(3.7) \quad E^\varepsilon := \left\{ \omega : \sup_{t \in [0, T]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\| < 0.99C_\epsilon \right\}.$$

Theorem 3.6. *Fix $T > 0$. Define*

$$H^\varepsilon := \left\{ \omega : \sup_{t \in [0, T]} \alpha_t^\varepsilon \geq \varepsilon^{a/2} \right\}, \quad S^\varepsilon := \left\{ \omega : \sup_{t \in [0, T]} \|\mathcal{Y}_t^\varepsilon\| \geq 8\varepsilon^a \right\},$$

for $a \in [0, 1)$. Fix $q > 0$ and assume $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$. Then $\exists \varepsilon_q > 0$ such that $\forall \varepsilon < \varepsilon_q$,

$$\mathbb{P}[H^\varepsilon] < q + 2(1 - p_\varepsilon), \quad \mathbb{P}[S^\varepsilon] < q + 2(1 - p_\varepsilon),$$

where $p_\varepsilon \rightarrow 1$ as $\varepsilon \rightarrow 0$ and is given explicitly in (3.6).

Proof is given in appendix B.4.

Note that we have extended our results on $[0, T \wedge \varepsilon^\varepsilon]$ to $[0, T]$ by leveraging a small probability q , provided that $\widehat{\mathfrak{z}}^\varepsilon$ system possess property $\mathcal{P}(T, q)$. Now we discuss under what conditions does $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$ for arbitrary $q > 0$.

Fix $T > 0$. In general one cannot expect $\mathcal{P}(T, q)$ to hold for arbitrary $q > 0$ —for example, if the cubic nonlinearities have a destabilizing effect then there is a non-zero probability that trajectories blow-up in finite time. Similar situation arises in stochastic partial differential equations—see remark 5.2 in [13]. When cubic nonlinearities have stabilizing effect, it is reasonable to expect $\mathcal{P}(T, q)$ to hold for arbitrary $q > 0$ (see proposition 3.8 below).

The following two propositions help in checking if the property $\mathcal{P}(T, q)$ is satisfied. Proofs of them are similar in nature to the proof of Theorem 5.1 in [13]. [13] deals with stochastic partial differential equations and the instability scenario there is different—analogue situation in delay equations case would be that “one root of the characteristic equation is zero, and all other roots have negative real parts”. For the scenario that we are considering in this paper, one pair of roots lie on the imaginary axis, and so there are oscillations in the system and the proofs requires a bit more work than that in [13].

Proposition 3.7 does not assume anything about the nature of the nonlinearity G —consequently its result is weak. Proposition 3.8 assumes that the nonlinearity is stabilizing and concludes that $\widehat{\mathfrak{z}}^\varepsilon$ possesses the property $\mathcal{P}(T, q)$ for any $q > 0$.

Proposition 3.7. *Fix $q > 0$. Then $\exists T_q > 0$ such that the $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$ for $T \in [0, T_q]$.*

Proof is given in appendix B.5.

Proposition 3.8. *Fix $T > 0$. Assume the cubic nonlinearity of G is stabilizing, i.e., $\exists C_G > 0$ such that*

$$(3.8) \quad \frac{\omega_c}{2\pi} \int_0^{2\pi/\omega_c} \left((e^{tB}z)^* \tilde{\Psi} \int_{-r}^0 (\Phi(\theta)e^{tB}z)^3 d\nu_3(\theta) \right) dt < -C_G \|z\|_2^4, \quad \forall z \in \mathbb{R}^2.$$

Then the $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$ for arbitrary $q > 0$.

Proof is given in appendix B.6.

Now consider the system

$$d\widetilde{\mathfrak{z}}_t^\varepsilon = e^{-tB/\varepsilon^2} \widetilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \widetilde{\mathfrak{z}}_t^\varepsilon) dt + e^{-tB/\varepsilon^2} \widetilde{\Psi} L_1(\Phi e^{tB/\varepsilon^2} \widetilde{\mathfrak{z}}_t^\varepsilon) dW, \quad \widetilde{\mathfrak{z}}_0^\varepsilon = z_0^\varepsilon,$$

i.e. we are totally ignoring the Q part—even the effect \mathfrak{Y} of the initial condition.

Define

$$\beta_t^\varepsilon = \frac{1}{2} \|\widetilde{\mathfrak{z}}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon\|_2^2.$$

Proposition 3.9. *Assume the cubic nonlinearity is such that (3.8) is satisfied, i.e. nonlinearity is stabilizing. Fix $T > 0$. Given any $q > 0$, $\exists C > 0$ and $\varepsilon_\circ > 0$ such that $\forall \varepsilon < \varepsilon_\circ$*

$$\mathbb{P} \left[\sup_{t \in [0, T]} \beta_t^\varepsilon \geq C\varepsilon^4 \right] \leq q.$$

Proof is in appendix B.7.

Combining theorem 3.6 and proposition 3.9 we get the following result.

Theorem 3.10. *Assume the cubic nonlinearity is such that (3.8) is satisfied, i.e. nonlinearity is stabilizing. Fix any $a \in [0, 1)$. For any $q > 0$, $\exists \varepsilon_q > 0$ such that $\forall \varepsilon < \varepsilon_q$*

$$(3.9) \quad \mathbb{P} \left[\sup_{s \in [0, T]} |X_s^\varepsilon - (\Phi(0)e^{sB/\varepsilon^2} \widetilde{\mathfrak{z}}_s^\varepsilon + \mathfrak{Y}_s^\varepsilon)| > 6\varepsilon^{a/4} \right] < 3q + 4(1 - p_\varepsilon)$$

where $p_\varepsilon \rightarrow 1$ as $\varepsilon \rightarrow 0$ and is given explicitly in (3.6).

Proof. Using $X_s^\varepsilon = \Phi(0)e^{sB/\varepsilon^2} \mathfrak{z}_s^\varepsilon + \mathcal{Y}_s^\varepsilon + \mathfrak{Y}_s^\varepsilon$ the above probability is bounded by

$$\begin{aligned} & \mathbb{P} \left[\sup_{s \in [0, T]} |\Phi(0)e^{sB/\varepsilon^2} (\mathfrak{z}_s^\varepsilon - \widetilde{\mathfrak{z}}_s^\varepsilon)| + |\Phi(0)e^{sB/\varepsilon^2} (\widetilde{\mathfrak{z}}_s^\varepsilon - \widetilde{\mathfrak{z}}_s^\varepsilon)| + |\mathcal{Y}_s^\varepsilon(0)| > 6\varepsilon^{a/4} \right] \\ & \leq \mathbb{P} \left[\sup_{s \in [0, T]} \sqrt{2\alpha_t^\varepsilon} > 2\varepsilon^{a/4} \right] + \mathbb{P} \left[\sup_{s \in [0, T]} \sqrt{2\beta_t^\varepsilon} > 2\varepsilon^{a/4} \right] + \mathbb{P} \left[\sup_{s \in [0, T]} \|\mathcal{Y}_s^\varepsilon\| > 2\varepsilon^{a/4} \right] \\ & < q + 2(1 - p_\varepsilon) + q + q + 2(1 - p_\varepsilon). \quad (\text{for } \varepsilon \text{ sufficiently small.}) \end{aligned}$$

□

Remark 3.1. Note that $\widetilde{\mathfrak{z}}^\varepsilon$ is a 2-dimensional system without delay and \mathfrak{Y}^ε is a deterministic process that has exponential decay. The above theorem shows that, for small enough ε , the delay system X^ε can be approximated by the $\widetilde{\mathfrak{z}}^\varepsilon$ system without delay, with probability close to 1.

When the nonlinearity is stabilizing, using standard averaging techniques for equations without delay (see for example [16]), it can be shown that the distribution

of $\tilde{\mathfrak{z}}^\varepsilon$ converges as $\varepsilon \rightarrow 0$ to the distribution of a 2-dimensional process $\tilde{\mathfrak{z}}^0$. Theorem 3.6 and propositions 3.9 show that $\sup_{t \in [0, T]} \beta_t^\varepsilon$ and $\sup_{t \in [0, T]} \alpha_t^\varepsilon$ converge to zero in probability. Hence, by theorem 3.1 in [12], the distribution of \mathfrak{z}^ε converges as $\varepsilon \rightarrow 0$ to the distribution of $\tilde{\mathfrak{z}}^0$. Also, the distribution of \mathcal{H}^ε process, where $\mathcal{H}_t^\varepsilon := \frac{1}{2} \|z_t^\varepsilon\|_2^2 = \frac{1}{2} \|\mathfrak{z}_t^\varepsilon\|_2^2$, converges as $\varepsilon \rightarrow 0$ to the distribution of \mathcal{H}^0 , where \mathcal{H}^0 is the weak-limit as $\varepsilon \rightarrow 0$ of the process $H_t^\varepsilon = \frac{1}{2} \|\tilde{\mathfrak{z}}_t^\varepsilon\|_2^2$.

Great simplification can be obtained when $\frac{1}{2} \|\mathfrak{z}^\varepsilon\|_2^2$ can be used to approximate the required quantities. For example, consider the exit time

$$(3.10) \quad \tau^\varepsilon := \inf\{t \geq 0 : |X_t^\varepsilon| \geq \sqrt{2H^*}\}$$

where H^* is fixed and is such that $\sqrt{2H^*} \gg \|(I - \pi)\Pi_0^\varepsilon X^\varepsilon\|$. Noting that $\Phi(0)e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon = (\mathfrak{z}_t^\varepsilon)_1 \cos(\omega t/\varepsilon^2) + (\mathfrak{z}_t^\varepsilon)_2 \sin(\omega t/\varepsilon^2)$; because of the fast oscillations of \mathfrak{z}^ε and fast decay of \mathfrak{Y}^ε and smallness of \mathfrak{Z}^ε , the exit time τ^ε would be very close to the exit time τ'^ε where

$$(3.11) \quad \tau'^\varepsilon := \inf\{t \geq 0 : \sqrt{(\mathfrak{z}_t^\varepsilon)_1^2 + (\mathfrak{z}_t^\varepsilon)_2^2} \geq \sqrt{2H^*}\}.$$

To approximate the distribution of τ^ε , one can study $H_t^\varepsilon := \frac{1}{2} \|\tilde{\mathfrak{z}}_t^\varepsilon\|_2^2$ and consider the distribution of

$$\tau^{\varepsilon, \hbar} := \inf\{t \geq 0 : H_t^\varepsilon \geq H^*\}.$$

The distribution of $\tau^{\varepsilon, \hbar}$ would be close to that of τ^ε . Since $\tilde{\mathfrak{z}}_t^\varepsilon$ does not involve any delay, standard averaging techniques can be used to show that the distribution of H^ε converges as $\varepsilon \rightarrow 0$ to the distribution of a specific 1-dimensional process \mathcal{H}^0 (without delay). Then the exit times

$$\tau^{\hbar} := \inf\{t \geq 0 : \mathcal{H}_t^0 \geq H^*\}$$

would closely approximate τ^ε . The advantages in doing so are: (i) \mathcal{H}^0 is a process without delay and hence easier to simulate (ii) numerical simulation of \mathcal{H}^0 can be done with a much coarser numerical mesh than that required for X^ε .

3.1. Example. Consider the following equation:

$$(3.12) \quad dX^\varepsilon(t) = \varepsilon^{-2} L_0(\Pi_t^\varepsilon X^\varepsilon) dt + G(\Pi_t^\varepsilon X^\varepsilon) dt + \sigma dW, \quad \Pi_0^\varepsilon X^\varepsilon = \xi,$$

where $L_0\eta = -\frac{\pi}{2}\eta(-1)$ and $G(\eta) = \gamma_c\eta^3(-1)$. The characteristic equation $\lambda + \frac{\pi}{2}e^{-\lambda} = 0$ has countably infinite roots on the complex plane. The roots with the largest real part are $\pm i\frac{\pi}{2}$. Hence L_0 satisfies the assumption 1.1. The basis Φ for P and the vector $\tilde{\Psi}$ can be evaluated as

$$\Phi(\theta) = \left[\cos\left(\frac{\pi}{2}\theta\right) \sin\left(\frac{\pi}{2}\theta\right) \right], \quad \tilde{\Psi} = \frac{2}{(1 + (\pi/2)^2)} \begin{bmatrix} 1 \\ \pi/2 \end{bmatrix}.$$

The corresponding $\tilde{\mathfrak{z}}^\varepsilon$ satisfies

$$d\tilde{\mathfrak{z}}^\varepsilon = \gamma_c e^{-tB/\varepsilon^2} \tilde{\Psi} \left(\Phi(-1) e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}^\varepsilon \right)^3 dt + e^{-tB/\varepsilon^2} \tilde{\Psi} \sigma dW, \quad \Phi \tilde{\mathfrak{z}}_0^\varepsilon = \pi \xi.$$

Let $H_t^\varepsilon = \frac{1}{2}((\tilde{\mathfrak{z}}_1^\varepsilon)^2 + (\tilde{\mathfrak{z}}_2^\varepsilon)^2)_t$. Applying Ito formula we have

$$dH_t^\varepsilon = \gamma_c (e^{-tB/\varepsilon^2} \tilde{\Psi})^* \tilde{\mathfrak{z}}^\varepsilon (\Phi(-1) e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}^\varepsilon)^3 dt + (e^{-tB/\varepsilon^2} \tilde{\Psi})^* \tilde{\mathfrak{z}}^\varepsilon \sigma dW + \frac{1}{2} \sigma^2 \left((e^{-tB/\varepsilon^2} \tilde{\Psi})_1^2 + (e^{-tB/\varepsilon^2} \tilde{\Psi})_2^2 \right) dt.$$

Averaging the fast oscillations we get that the probability distribution of H^ε converges as $\varepsilon \rightarrow 0$ to the probability distribution of

$$(3.13) \quad d\mathcal{H}_t^0 = \left(-\frac{3\gamma_c \pi/2}{1 + (\pi/2)^2} (\mathcal{H}_t^0)^2 + \frac{2\sigma^2}{1 + (\pi/2)^2} \right) dt + \frac{2}{\sqrt{1 + (\pi/2)^2}} \sqrt{\mathcal{H}_t^0} \sigma dW.$$

Now we illustrate our results employing numerical simulations.

Draw a random sample of N_{samp} particles with initial H values $\{h_i\}_{i=1}^{N_{samp}}$. Simulate them according to (3.13) for $0 \leq t \leq T_{end}$.

Simulate (3.12) for $0 \leq t \leq T_{end}$ using initial trajectories $\{\sqrt{2h_i} \cos(\omega_c \cdot)\}_{i=1}^{N_{samp}}$.

Let $\tau^\varepsilon := \inf\{t \geq 0 : |X_t^\varepsilon| \geq \sqrt{2H^*}\}$ and $\tau^h := \inf\{t \geq 0 : H^0(t) \geq H^*\}$

We can check whether the following pairs are close.

1. the distribution of $\frac{1}{2}((z_{T_{end}}^\varepsilon)_1^2 + (z_{T_{end}}^\varepsilon)_2^2)$ from (3.12) and the distribution of $\mathcal{H}_{T_{end}}^0$ from (3.13),
2. distribution of τ^ε and the distribution of τ^h .

We took $\varepsilon = 0.025$, $H^* = 1.5$, $T_{end} = 2$, $N_{samp} = 4000$, and $\sqrt{2\{h_i\}_{i=1}^{N_{samp}}} = 1.2$. Figures 1 and 2 answer the above questions. Two cases are considered with $\sigma = 1$ fixed: $\gamma_c = 1$ and $\gamma_c = 0$.

More examples (oscillators with cubic nonlinearity) are discussed in⁴ [7].

⁴[7] employs complex coordinates and so the form of answers would differ from this paper. However the numerical values would be same. For example, in this paper we write an element in P as $z_1 \cos(\omega_c \cdot) + z_2 \sin(\omega_c \cdot)$ with $z_i \in \mathbb{R}$. But in [7] we write $z_1 e^{i\omega_c \cdot} + z_2 e^{-i\omega_c \cdot}$ with z_1 and z_2 being complex conjugates. For multidimensional systems as treated in [7] this complex coordinates is more convenient.

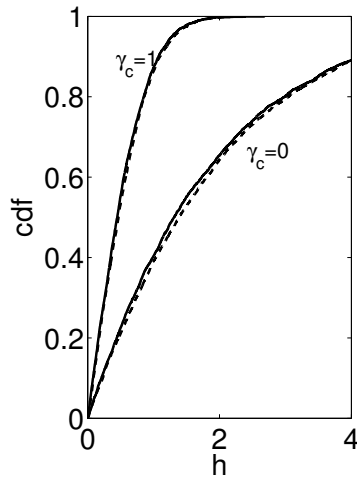


FIGURE 1. cdf of $\frac{1}{2}((z_T^\varepsilon)_1^2 + (z_T^\varepsilon)_2^2)$ (dashed line) and \mathcal{H}_T^0 (solid line).

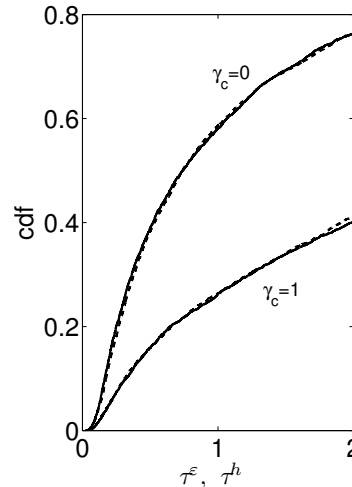


FIGURE 2. cdf of τ^ε (dashed line) and cdf of τ^h (solid line). The cdf value at $\tau^\varepsilon = 2$ indicates the fraction of particles whose modulus exceeded $\sqrt{2H^*}$ before the time $T = 2$.

REFERENCES

- [1] G. Stepan, *Retarded dynamical systems: stability and characteristic functions*, Longman Scientific & Technical, 1989.
- [2] G. Stepan, *Modelling nonlinear regenerative effects in metal cutting*, Phil. Trans. R. Soc. Lond. A, **359**:739–757, 2001.
- [3] A. Longtin, J.G. Milton, J.E. Bos and M.C. Mackey, *Noise and critical behavior of the pupil light reflex at oscillation onset*, Phys. Rev. A, **41**:6992–7005, 1990.
- [4] W. Yao, P. Yu and C. Essex, *Delayed stochastic differential model for quiet standing*, Phys. Rev. E, **63**:021902, 2001.
- [5] E. Buckwar, R. Kuske, B. L’Esperance and T. Soo, *Noise-sensitivity in machine tool vibrations*, International Journal of Bifurcation and Chaos, **16**(08):2407–2416, 2006.
- [6] N. Lingala and N. Sri Namachchivaya, *Nonlinear and additive white noise perturbations of linear delay differential equations at the verge of instability: an averaging approach*, Stochastics and Dynamics, **16**(4), 2016.
- [7] N. Lingala and N. Sri Namachchivaya, *Perturbations of linear delay differential equations at the verge of instability*, Physical Review E, **93**:062104, 2016.
- [8] J. K. Hale and S.M. Verduyn Lunel, *Introduction to functional differential equations*, Springer Verlag, 1993.
- [9] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel and H.O. Walther, *Delay equations*, Springer Verlag, 1995.
- [10] S-E. A. Mohammed, *Stochastic functional differential equations*, Pitman, 1984.
- [11] M. Rei, M. Riedle and O. van Gaans, *On Émery’s Inequality and a Variation-of-Constants Formula*, Stochastic Analysis and Applications, **25**(2):353–379, 2007.
- [12] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons, 2nd edition, 1999.

[13] D. Blomker, *Amplitude equations for locally cubic nonautonomous nonlinearities*, SIAM Journal on Applied Dynamical Systems, **2**(3):464–486, 2003.
 [14] M. Fischer and G. Nappo. *On the Moments of the Modulus of Continuity of Ito Processes*, Stochastic Analysis and Applications, **28**(1):103–122, 2009.
 [15] M. Scheutzw, *A stochastic Gronwall lemma*, arxiv:1304.5424
 [16] N. Sri Namachchivaya and R. Sowers, *Rigorous stochastic averaging at a center*, Meccanica, **37**:85–114, 2002.
 [17] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Springer, 1999.
 [18] D. Bainov and P. Simeonov, *Integral inequalities and applications*, Kluwer Academic, 1992.
 [19] J. K. Hale, *Ordinary differential equations*, Krieger Publishing Company, 2nd edition, 1980.

Appendix A. PROOFS OF RESULTS IN SECTION 2

A.1. **Proof of Lemma 2.1.** For $\eta \in \mathcal{C}$, by $\eta(\theta)$ we mean η evaluated at $\theta \in [-r, 0]$.

Let $X_{q,t}^\varepsilon := ((1 - \pi)\Pi_t^\varepsilon X^\varepsilon)(0)$, and $X_{p,t}^\varepsilon := (\pi\Pi_t^\varepsilon X^\varepsilon)(0)$. And for the unperturbed system (1.6), let $x_{q,t} := ((1 - \pi)\Pi_t x)(0)$, and $x_{p,t} := (\pi\Pi_t x)(0)$ with the initial condition $\Pi_0 x = \Pi_0^\varepsilon X^\varepsilon = \xi$. Let

$$Z_t := \int_0^t L_1(\Pi_s^\varepsilon X^\varepsilon) dW_s.$$

In (1.21), γ was defined. Let

$$\chi(t) = (\hat{T}(t)\hat{\pi}\mathbf{1}_{\{0\}})(0).$$

Using $\hat{\pi}\mathbf{1}_{\{0\}} = \Phi\tilde{\Psi}$ from (1.13) and $\hat{T}(t)\Phi = \Phi e^{tB}$, we get $\chi(t) = \Phi(0)e^{tB}\tilde{\Psi}$.

Using the variation-of-constants formula (1.18)-(1.20) we have for $t \geq 0$

$$X_t^\varepsilon = X_{p,t}^\varepsilon + X_{q,t}^\varepsilon = x_{p,t/\varepsilon^2} + x_{q,t/\varepsilon^2} + D_{p,t} + D_{q,t} + A_{p,t} + A_{q,t}$$

with

$$\begin{aligned} D_{q,t} &:= \int_0^t \gamma\left(\frac{t-s}{\varepsilon^2}\right) G(\Pi_s^\varepsilon X^\varepsilon) ds, & A_{q,t} &:= \int_0^t \gamma\left(\frac{t-s}{\varepsilon^2}\right) dZ_s, \\ D_{p,t} &:= \int_0^t \chi\left(\frac{t-s}{\varepsilon^2}\right) G(\Pi_s^\varepsilon X^\varepsilon) ds, & A_{p,t} &:= \int_0^t \chi\left(\frac{t-s}{\varepsilon^2}\right) dZ_s. \end{aligned}$$

For any process M , we define $M_t^* := \sup_{0 \leq s \leq t} |M_s|$. Now, what we mean by $D_{q,t}^*$, $A_{q,t}^*$ and $x_{q,t}^*$ etc is clear. Also define

$$\mathfrak{X}_t^\varepsilon := \sup_{s \in [0,t]} |X_s^\varepsilon|^n.$$

We then have,

$$\begin{aligned} (A.1) \quad 2^{-5(n-1)} \mathbb{E} \mathfrak{X}_t^\varepsilon &\leq \mathbb{E} |x_{p,t/\varepsilon^2}^*|^n + \mathbb{E} |x_{q,t/\varepsilon^2}^*|^n \\ &\quad + \mathbb{E} |D_{p,t}^*|^n + \mathbb{E} |D_{q,t}^*|^n + \mathbb{E} |A_{p,t}^*|^n + \mathbb{E} |A_{q,t}^*|^n. \end{aligned}$$

First we focus on the terms involving the process A . Using integration by parts we have

$$A_{q,s} = \gamma(0)Z_s + \int_0^s \varepsilon^{-2} \gamma' \left(\frac{s-u}{\varepsilon^2} \right) Z_u du.$$

Using Minkowski inequality,

$$(A.2) \quad \mathbb{E} |A_{q,t}^*|^n \leq 2^{n-1} |\gamma(0)|^n \mathbb{E} \sup_{s \in [0,t]} |Z_s|^n \\ + 2^{n-1} \mathbb{E} \sup_{s \in [0,t]} \left| \int_0^s \varepsilon^{-2} \gamma' \left(\frac{s-u}{\varepsilon^2} \right) Z_u du \right|^n.$$

The second term on the RHS of (A.2) is bounded above (using the exponential decay (1.23)) by

$$2^{n-1} \mathbb{E} \sup_{s \in [0,t]} \left| \int_0^s \varepsilon^{-2} \left| \gamma' \left(\frac{s-u}{\varepsilon^2} \right) \right| |Z_u| du \right|^n \\ \leq 2^{n-1} \mathbb{E} \sup_{s \in [0,t]} \left| \int_0^s \varepsilon^{-2} \tilde{K} \|L_0\| e^{-\kappa(s-u)/\varepsilon^2} |Z_u| du \right|^n \\ \leq 2^{n-1} (\tilde{K} \|L_0\|/\kappa)^n \mathbb{E} \sup_{s \in [0,t]} |Z_s|^n,$$

where $\tilde{K} = K \|(1 - \hat{\pi})\mathbf{1}_{\{0\}}\|$. Hence, using Burkholder-Davis-Gundy inequality and Holder inequality

$$\mathbb{E} |A_{q,t}^*|^n \leq 2^{n-1} \left(|\gamma(0)|^n + (\tilde{K} \|L_0\|/\kappa)^n \right) \mathbb{E} \sup_{s \in [0,t]} |Z_s|^n \\ \leq 2^{n-1} \left(|\gamma(0)|^n + (\tilde{K} \|L_0\|/\kappa)^n \right) C_m \mathbb{E} \left(\int_0^t L_1^2(\Pi_u^\varepsilon X^\varepsilon) du \right)^{n/2} \\ \leq 2^{n-1} \left(|\gamma(0)|^n + (\tilde{K} \|L_0\|/\kappa)^n \right) C_m t^{\frac{n-2}{2}} \|L_1\|^n \left(\int_0^t \mathbb{E} \mathfrak{X}_u^\varepsilon du + \mathbb{E} \|\xi\|^n (t \wedge \varepsilon^2 r) \right) \\ = C_{m,L} t^{\frac{n-2}{2}} \left(\int_0^t \mathbb{E} \mathfrak{X}_u^\varepsilon du + \mathbb{E} \|\xi\|^n (t \wedge \varepsilon^2 r) \right)$$

where $C_{m,L} = 2^{n-1} \left(|\gamma(0)|^n + (\tilde{K} \|L_0\|/\kappa)^n \right) C_m \|L_1\|^n$ and $t \wedge \varepsilon^2 r$ means $\min\{t, \varepsilon^2 r\}$. Now we focus on $A_{p,t}$.

$$A_{p,t} = \int_0^t \chi \left(\frac{t-s}{\varepsilon^2} \right) dZ_s \\ = \int_0^t \left(\Phi(0) e^{(t-s)B/\varepsilon^2} \tilde{\Psi} \right) dZ_s = \Phi(0) e^{tB/\varepsilon^2} \left(\int_0^t e^{-sB/\varepsilon^2} dZ_s \right) \tilde{\Psi}.$$

Let $M_t^c = \int_0^t \cos(\omega_c s/\varepsilon^2) dZ_s$ and $M_t^s = \int_0^t \sin(\omega_c s/\varepsilon^2) dZ_s$. Then

$$\mathbb{E} |A_{p,t}^*|^n \leq (\|\Phi(0)\|_2 \|\tilde{\Psi}\|_2)^n \mathbb{E} (M_t^{c,*} + M_t^{s,*})^n.$$

Using BDG and Holder inequalities,

$$\mathbb{E}|A_{p,t}^*|^n \leq 2^{n-1}(\|\Phi(0)\|_2\|\tilde{\Psi}\|_2)^n C t^{\frac{n-2}{2}}\|L_1\|^n \left(\int_0^t \mathbb{E}\mathfrak{X}_u^\varepsilon du + \mathbb{E}\|\xi\|^n(t \wedge \varepsilon^2 r) \right).$$

Now we focus on the process D . Using exponential decay of γ we have

$$\begin{aligned} |D_{q,t}| &\leq \int_0^t \left| \gamma \left(\frac{t-s}{\varepsilon^2} \right) \right| |G(\Pi_s^\varepsilon X^\varepsilon)| ds \\ &\leq \sup_{s \in [0,t]} |G(\Pi_s^\varepsilon X^\varepsilon)| \int_0^t \tilde{K} e^{-\kappa(t-s)/\varepsilon^2} ds \leq \varepsilon^2 (\tilde{K}/\kappa) \sup_{s \in [0,t]} |G(\Pi_s^\varepsilon X^\varepsilon)|. \end{aligned}$$

Hence, using the Lipschitz condition $|G(\eta)| \leq K_G \|\eta\|$ we have,

$$\mathbb{E}|D_{q,t}^*|^n \leq \varepsilon^{2n} (\tilde{K} K_G / \kappa)^n (\mathbb{E}\mathfrak{X}_t^\varepsilon + \mathbb{E}\|\xi\|^n).$$

Now,

$$\begin{aligned} |D_{p,t}| &\leq \int_0^t \left| \chi \left(\frac{t-s}{\varepsilon^2} \right) \right| |G(\Pi_s^\varepsilon X^\varepsilon)| ds = \int_0^t \left| \Phi(0) e^{(t-s)B/\varepsilon^2} \tilde{\Psi} \right| |G(\Pi_s^\varepsilon X^\varepsilon)| ds \\ &\leq \|\Phi(0)\|_2 \|\tilde{\Psi}\|_2 \int_0^t |G(\Pi_s^\varepsilon X^\varepsilon)| ds \leq \|\Phi(0)\|_2 \|\tilde{\Psi}\|_2 K_G \int_0^t \|\Pi_s^\varepsilon X^\varepsilon\| ds. \end{aligned}$$

Hence, using BDG and Holder inequalities,

$$\mathbb{E}|D_{p,t}^*|^n \leq (\|\Phi(0)\|_2 \|\tilde{\Psi}\|_2 K_G)^n t^{n-1} \left(\int_0^t \mathbb{E}\mathfrak{X}_s^\varepsilon ds + \mathbb{E}\|\xi\|^n(t \wedge \varepsilon^2 r) \right).$$

Now, we focus on the deterministic terms. Because of our assumption on L_0 , there exists $C_{L_0} > 0$ such that $x_{q,t/\varepsilon^2}^* \leq \sqrt{C_{L_0}} \|(1-\pi)\xi\| e^{-\kappa t/\varepsilon^2}$ and $x_{p,t/\varepsilon^2}^* \leq \sqrt{C_{L_0}} \|\pi\xi\|$.

Collecting all the above results in (A.1), we have for $n > 2$,

$$(2^{-5(n-1)} - \varepsilon^{2n} (\tilde{K} K_G / \kappa)^n) \mathbb{E}\mathfrak{X}_t^\varepsilon \leq C_1 + C_2 \int_0^t \mathbb{E}\mathfrak{X}_s^\varepsilon ds,$$

where

$$\begin{aligned} C_1 &= C_{L_0}^{n/2} (\mathbb{E}\|\pi\xi\|^n + \mathbb{E}\|(1-\pi)\xi\|^n) + \varepsilon^{2n} (\tilde{K} K_G / \kappa)^n \mathbb{E}\|\xi\|^n \\ &\quad + (\|\Phi(0)\|_2 \|\tilde{\Psi}\|_2)^n (T^{n-1} K_G^n + 2^{n-1} \|L_1\|^n C T^{(n-2)/2}) \mathbb{E}\|\xi\|^n \varepsilon^2 r, \end{aligned}$$

$$C_2 = (\|\Phi(0)\|_2 \|\tilde{\Psi}\|_2 K_G)^n T^{n-1} + C_{m,L} T^{(n-2)/2} + 2^{n-1} \|\Phi(0)\|_2^n \|\tilde{\Psi}\|_2^n C T^{\frac{n-2}{2}} \|L_1\|^n.$$

The initial condition ξ is assumed to be deterministic and hence C_1 can be written as $C_1 = C_{L_0}^{n/2} (\|\pi\xi\|^n + \|(1-\pi)\xi\|^n) + \varepsilon^{2n} (K K_G / \kappa)^n \|\xi\|^n$.

Applying Gronwall inequality we have $\mathbb{E}\mathfrak{X}_T^\varepsilon \leq \frac{2C_1}{2^{-5(n-1)}} \exp\left(\frac{2C_2}{2^{-5(n-1)}} T\right)$ for small enough ε .

For $0 \leq n \leq 2$ we can use $\mathbb{E} \sup_{t \in [0,T]} |X_t^\varepsilon|^n \leq 1 + \mathbb{E} \sup_{t \in [0,T]} |X_t^\varepsilon|^3$.

A.2. Proof of Proposition 2.3. Recall the γ defined in (1.21). We have

$$\left(\hat{T}\left(\frac{s-u}{\varepsilon^2}\right)(1-\hat{\pi})\mathbf{1}_{\{0\}} \right) (\theta) = \begin{cases} \gamma\left(\frac{s+\varepsilon^2\theta-u}{\varepsilon^2}\right), & s+\varepsilon^2\theta-u \geq 0 \\ \hat{\pi}\mathbf{1}_{\{0\}}\left(\frac{s+\varepsilon^2\theta-u}{\varepsilon^2}\right), & s+\varepsilon^2\theta-u < 0. \end{cases}$$

Note that

$$(A.3) \quad \sup_{s \in [0, T]} \Upsilon_s^\varepsilon \leq \sup_{s \in [0, T]} \sup_{\theta \in [-r, 0]} \left| \int_0^{(s+\varepsilon^2\theta) \vee 0} \gamma\left(\frac{s+\varepsilon^2\theta-u}{\varepsilon^2}\right) dZ_u \right| \\ + \sup_{s \in [0, T]} \sup_{\theta \in [-r, 0]} \left| \int_{(s+\varepsilon^2\theta) \vee 0}^s \hat{\pi}\mathbf{1}_{\{0\}}\left(\frac{s+\varepsilon^2\theta-u}{\varepsilon^2}\right) dZ_u \right| =: \mathcal{J}_1 + \mathcal{J}_2.$$

In the above $t \vee s$ means $\max\{t, s\}$.

For \mathcal{J}_1 we have (with $\delta \in (0, 2)$)

$$\mathcal{J}_1 = \sup_{t \in [0, T]} \left| \int_0^t \gamma\left(\frac{t-u}{\varepsilon^2}\right) dZ_u \right| \\ \leq \sup_{t \in [r\varepsilon^\delta, T]} \left| \int_0^{t-r\varepsilon^\delta} \gamma\left(\frac{t-u}{\varepsilon^2}\right) dZ_u \right| + \sup_{t \in [0, T]} \left| \int_{(t-r\varepsilon^\delta) \vee 0}^t \gamma\left(\frac{t-u}{\varepsilon^2}\right) dZ_u \right| \\ =: \mathcal{J}_{1a} + \mathcal{J}_{1b}.$$

Using integration by parts and exponential decay of γ and γ' (see (1.22)–(1.23)) in \mathcal{J}_{1a} we have

$$\mathcal{J}_{1a} \leq \sup_{t \in [r\varepsilon^\delta, T]} |\gamma(r\varepsilon^{\delta-2})Z_{t-r\varepsilon^\delta}| + \sup_{t \in [r\varepsilon^\delta, T]} \frac{1}{\varepsilon^2} \int_0^{t-r\varepsilon^\delta} \left| \gamma'\left(\frac{t-u}{\varepsilon^2}\right) \right| |Z_u| du \\ \leq \tilde{K} e^{-\kappa r\varepsilon^{\delta-2}} \sup_{t \in [0, T]} |Z_t| + \sup_{t \in [r\varepsilon^\delta, T]} \frac{1}{\varepsilon^2} \|L_0\| \tilde{K} \int_0^{t-r\varepsilon^\delta} e^{-\kappa(t-u)/\varepsilon^2} |Z_u| du \\ \leq \tilde{K} \left(1 + \frac{\|L_0\|}{\kappa} \right) e^{-\kappa r\varepsilon^{\delta-2}} \sup_{t \in [0, T]} |Z_t|,$$

where $\tilde{K} = K\|(1-\pi)\mathbf{1}_{\{0\}}\|$. For \mathcal{J}_{1b} we use

$$\gamma\left(\frac{t-u}{\varepsilon^2}\right) = \gamma\left(\frac{t - ((t-r\varepsilon^\delta) \vee 0)}{\varepsilon^2}\right) - \frac{1}{\varepsilon^2} \int_{(t-r\varepsilon^\delta) \vee 0}^u \gamma'\left(\frac{t-\tau}{\varepsilon^2}\right) d\tau, \\ \text{for } u \in [(t-r\varepsilon^\delta) \vee 0, t].$$

Now, using the definition 2.2 of modulus of continuity, we have

$$\mathcal{J}_{1b} \leq \sup_{t \in [0, r\varepsilon^\delta]} |\gamma(t\varepsilon^{-2})| |Z_t| + \sup_{t \in [r\varepsilon^\delta, T]} |\gamma(r\varepsilon^{\delta-2})| |Z_t - Z_{t-r\varepsilon^\delta}| \\ + \sup_{t \in [0, T]} \frac{1}{\varepsilon^2} \left| \int_{(t-r\varepsilon^\delta) \vee 0}^t \left(\int_{(t-r\varepsilon^\delta) \vee 0}^u \gamma'\left(\frac{t-\tau}{\varepsilon^2}\right) d\tau \right) dZ_u \right|$$

$$\begin{aligned}
&\leq 2\tilde{K}\mathfrak{w}(r\varepsilon^\delta, T; Z) + \sup_{t \in [0, T]} \frac{1}{\varepsilon^2} \left| \int_{(t-r\varepsilon^\delta) \vee 0}^t \left(\int_\tau^t dZ_u \right) \gamma' \left(\frac{t-\tau}{\varepsilon^2} \right) d\tau \right| \\
&\leq 2\tilde{K}\mathfrak{w}(r\varepsilon^\delta, T; Z) + \sup_{t \in [0, T]} \frac{1}{\varepsilon^2} \int_{(t-r\varepsilon^\delta) \vee 0}^t |Z_t - Z_\tau| \left| \gamma' \left(\frac{t-\tau}{\varepsilon^2} \right) \right| d\tau \\
&\leq 2\tilde{K}\mathfrak{w}(r\varepsilon^\delta, T; Z) + \mathfrak{w}(r\varepsilon^\delta, T; Z) \sup_{t \in [0, T]} \frac{1}{\varepsilon^2} \int_{(t-r\varepsilon^\delta) \vee 0}^t K \|L_0\| e^{-\kappa(t-\tau)/\varepsilon^2} d\tau \\
&\leq 2\tilde{K} \left(1 + \frac{\|L_0\|}{2\kappa} \right) \mathfrak{w}(r\varepsilon^\delta, T; Z).
\end{aligned}$$

For \mathcal{J}_2 we make use of the following facts:

$$\begin{aligned}
\hat{\pi} \mathbf{1}_{\{0\}}(v-u) &= \tilde{\Psi}_1 \cos(\omega_c(v-u)) + \tilde{\Psi}_2 \sin(\omega_c(v-u)) \\
&= (\tilde{\Psi}_1 \cos \omega_c v + \tilde{\Psi}_2 \sin \omega_c v) \cos \omega_c u + (\tilde{\Psi}_1 \sin \omega_c v - \tilde{\Psi}_2 \cos \omega_c v) \sin \omega_c u,
\end{aligned}$$

and $|\tilde{\Psi}_1 \cos \omega_c v + \tilde{\Psi}_2 \sin \omega_c v| \leq \sqrt{\tilde{\Psi}_1^2 + \tilde{\Psi}_2^2} = \|\tilde{\Psi}\|_2$. Using these it is easy to see that the

$$(A.4) \quad \mathcal{J}_2 \leq \|\tilde{\Psi}\|_2 \sup_{s \in [0, T]} \sup_{\theta \in [-r, 0]} \left(\left| M_s^{c, \varepsilon} - M_{(s+\varepsilon^2\theta) \vee 0}^{c, \varepsilon} \right| + \left| M_s^{s, \varepsilon} - M_{(s+\varepsilon^2\theta) \vee 0}^{s, \varepsilon} \right| \right)$$

where

$$M_t^{c, \varepsilon} = \int_0^t \cos(\omega_c u / \varepsilon^2) dZ_u, \quad M_t^{s, \varepsilon} = \int_0^t \sin(\omega_c u / \varepsilon^2) dZ_u.$$

Using the definition 2.2 of modulus of continuity, we have

$$\mathcal{J}_2 \leq \|\tilde{\Psi}\|_2 \left(\mathfrak{w}(\varepsilon^2 r, T; M^{c, \varepsilon}) + \mathfrak{w}(\varepsilon^2 r, T; M^{s, \varepsilon}) \right).$$

Collecting all the above estimates in (A.3) we have

$$\begin{aligned}
\sup_{s \in [0, T]} \Upsilon_s^\varepsilon &\leq \tilde{K} \left(1 + \frac{\|L_0\|}{\kappa} \right) e^{-\kappa r \varepsilon^{\delta-2}} \sup_{t \in [0, T]} |Z_t| \\
&\quad + 2\tilde{K} \left(1 + \frac{\|L_0\|}{2\kappa} \right) \mathfrak{w}(r\varepsilon^\delta, T; Z) \\
&\quad + \|\tilde{\Psi}\|_2 \left(\mathfrak{w}(\varepsilon^2 r, T; M^{c, \varepsilon}) + \mathfrak{w}(\varepsilon^2 r, T; M^{s, \varepsilon}) \right).
\end{aligned}$$

Now we take expectations. Using Burkholder-Davis-Gundy inequality and lemma 2.1, we have for $n \geq 1$,

$$\mathbb{E} \sup_{t \in [0, T]} |Z_t|^n \leq C \mathbb{E} \langle Z \rangle_T^{n/2} \leq C \mathbb{E} \left(T \sup_{t \in [0, T]} |L_1(\Pi_t^\varepsilon X^\varepsilon)|^2 \right)^{n/2} \leq CT^{n/2} \mathfrak{C}.$$

Using the Theorem 1 in section 3 of [14] and lemma 2.1, we get that there exists constants $C_{\mathfrak{w}}$, $C_{\mathfrak{w}}^c$ and $C_{\mathfrak{w}}^s$ such that, $\mathbb{E} \mathfrak{w}^n(r\varepsilon^\delta, T; Z) \leq C_{\mathfrak{w}} (r\varepsilon^\delta \ln(\frac{2T}{r\varepsilon^\delta}))^{n/2}$, $\mathbb{E} \mathfrak{w}^n(\varepsilon^2 r, T; M^{c, \varepsilon}) \leq C_{\mathfrak{w}}^c (\varepsilon^2 r \ln(\frac{2T}{\varepsilon^2 r}))^{n/2}$ and $\mathbb{E} \mathfrak{w}^n(\varepsilon^2 r, T; M^{s, \varepsilon}) \leq C_{\mathfrak{w}}^s (\varepsilon^2 r \ln(\frac{2T}{\varepsilon^2 r}))^{n/2}$.

Collecting all, we have

$$\begin{aligned} 2^{-5(n-1)}\mathbb{E} \sup_{s \in [0, T]} (\Upsilon_s^\varepsilon)^n &\leq \tilde{K}^n \left(1 + \frac{\|L_0\|}{\kappa}\right)^n C T^{n/2} \mathfrak{C} e^{-n\kappa r \varepsilon^{\delta-2}} \\ &\quad + \left(2\tilde{K} \left(1 + \frac{\|L_0\|}{2\kappa}\right)\right)^n C_{\mathfrak{w}} \left(\varepsilon^\delta r \ln \left(\frac{2T}{\varepsilon^\delta r}\right)\right)^{n/2} \\ &\quad + \|\tilde{\Psi}\|_2^n (C_{\mathfrak{w}}^c + C_{\mathfrak{w}}^s) \left(\varepsilon^2 r \ln \left(\frac{2T}{\varepsilon^2 r}\right)\right)^{n/2}. \end{aligned}$$

As $\varepsilon \rightarrow 0$, the 2nd term on the RHS dominates and hence we have (2.3).

A.3. Proof of Proposition 2.4. Using the variation of constants formula (1.20), we have

$$(A.5) \quad \|\mathcal{Y}_s^\varepsilon\| \leq \left\| \int_0^s \hat{T} \left(\frac{s-u}{\varepsilon^2}\right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \right\| + \Upsilon_s^\varepsilon.$$

Using the exponential decay (1.14) we have

$$\begin{aligned} \left\| \int_0^s \hat{T} \left(\frac{s-u}{\varepsilon^2}\right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \right\| &\leq \int_0^s \|\hat{T} \left(\frac{s-u}{\varepsilon^2}\right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}}\| \|G(\Pi_u^\varepsilon X^\varepsilon)\| du \\ &\leq K_G \sup_{u \in [0, s]} \|\Pi_u^\varepsilon X^\varepsilon\| \int_0^s \tilde{K} e^{-\kappa(s-u)/\varepsilon^2} du \leq (\varepsilon^2 K_G \tilde{K} / \kappa) \sup_{u \in [0, s]} \|\Pi_u^\varepsilon X^\varepsilon\|, \end{aligned}$$

where $\tilde{K} = K \|(1 - \hat{\pi}) \mathbf{1}_{\{0\}}\|$. Hence for $s \in [0, T]$

$$\|\mathcal{Y}_s^\varepsilon\| \leq (\varepsilon^2 K_G \tilde{K} / \kappa) \sup_{u \in [0, s]} \|\Pi_u^\varepsilon X^\varepsilon\| + \Upsilon_s^\varepsilon.$$

Hence

$$\sup_{s \in [0, t]} \|\mathcal{Y}_s^\varepsilon\| \leq (\varepsilon^2 K_G \tilde{K} / \kappa) \sup_{s \in [0, t]} \|\Pi_s^\varepsilon X^\varepsilon\| + \sup_{s \in [0, t]} \Upsilon_s^\varepsilon.$$

Raise to power n , take expectation and apply lemma 2.1 for the first term on the RHS and proposition 2.3 for the second term to get (2.4).

A.4. Proof of lemma 2.5. For any \mathbb{R}^2 -vector v , and $\theta \in [-r, 0]$, we have $\Phi(\theta) e^{tB/\varepsilon^2} v = v_1 \cos((\omega_c t / \varepsilon^2) + \theta) + v_2 \sin((\omega_c t / \varepsilon^2) + \theta)$. Hence

$$(A.6) \quad \|\Phi e^{tB/\varepsilon^2} v\| = \sup_{\theta \in [-r, 0]} |\Phi(\theta) e^{tB/\varepsilon^2} v| \leq \sqrt{v_1^2 + v_2^2}.$$

Using Lipschitz condition on G , and then using $y_t^\varepsilon - \mathfrak{Y}_t^\varepsilon = \mathcal{Y}_t^\varepsilon$ and (A.6), we get

$$(A.7) \quad \left| G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) - G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon) \right| \leq K_G (\|\mathcal{Y}_t^\varepsilon\| + \sqrt{2\alpha_t^\varepsilon}).$$

Using the definition (2.5) of Γ_t we have

$$\begin{aligned} |\Gamma_t| &= \left| (\tilde{\Psi}_1(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1 + \tilde{\Psi}_2(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2) \cos(\omega_c t / \varepsilon^2) \right. \\ &\quad \left. + (\tilde{\Psi}_1(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2 - \tilde{\Psi}_2(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1) \sin(\omega_c t / \varepsilon^2) \right| \end{aligned}$$

$$\begin{aligned} &\leq \sqrt{(\tilde{\Psi}_1(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1 + \tilde{\Psi}_2(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2)^2 + (\tilde{\Psi}_1(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_2 - \tilde{\Psi}_2(\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)_1)^2} \\ &= \|\tilde{\Psi}\|_2 \sqrt{2\alpha_t^\varepsilon}. \end{aligned}$$

Using the above inequality and (A.7) in the definition of \mathcal{B}_t we get

$$\begin{aligned} |\mathcal{B}_t| &\leq \|\tilde{\Psi}\|_2 \sqrt{2\alpha_t^\varepsilon} K_G (\|\mathcal{Y}_t^\varepsilon\| + \sqrt{2\alpha_t^\varepsilon}) + \frac{1}{2} \|\tilde{\Psi}\|_2^2 \|L_1\|^2 (\sqrt{2\alpha_t^\varepsilon} + \|\mathcal{Y}_t^\varepsilon\|)^2 \\ &\leq \|\tilde{\Psi}\|_2 K_G \left(\frac{\|\mathcal{Y}_t^\varepsilon\|^2 + 2\alpha_t^\varepsilon}{2} + 2\alpha_t^\varepsilon \right) + \|\tilde{\Psi}\|_2^2 \|L_1\|^2 (2\alpha_t^\varepsilon + \|\mathcal{Y}_t^\varepsilon\|^2) \\ &\leq C_{\mathcal{B}} (\alpha_t^\varepsilon + \|\mathcal{Y}_t^\varepsilon\|^2). \end{aligned}$$

Using $|\Gamma_t| \leq \|\tilde{\Psi}\|_2 \sqrt{2\alpha_t^\varepsilon}$ in the definition of Σ_t we get

$$\begin{aligned} \Sigma_t^2 &\leq \Gamma_t^2 \|L_1\|^2 (\sqrt{2\alpha_t^\varepsilon} + \|\mathcal{Y}_t^\varepsilon\|)^2 \\ &\leq \|\tilde{\Psi}\|_2^2 \|L_1\|^2 2\alpha_t^\varepsilon (\sqrt{2\alpha_t^\varepsilon} + \|\mathcal{Y}_t^\varepsilon\|)^2 \leq 16 \|\tilde{\Psi}\|_2^2 \|L_1\|^2 ((\alpha_t^\varepsilon)^2 + \|\mathcal{Y}_t^\varepsilon\|^4). \end{aligned}$$

A.5. Proof of Proposition 2.6. Using lemma 2.5 we have that

$$d\alpha_t^\varepsilon \leq C_{\mathcal{B}} (\alpha_t^\varepsilon + \|\mathcal{Y}_t^\varepsilon\|^2) dt + \Sigma_t dW_t.$$

Let

$$H_t := C_{\mathcal{B}} \int_0^t \|\mathcal{Y}_s^\varepsilon\|^2 ds, \quad M_t := \int_0^t \Sigma_s dW_s, \quad L_t := \int_0^t e^{-C_{\mathcal{B}}s} dM_s.$$

Then,

$$(A.8) \quad \alpha_t^\varepsilon \leq \int_0^t C_{\mathcal{B}} \alpha_s^\varepsilon ds + H_t + M_t.$$

Applying Gronwall inequality pathwise, we get,

$$(A.9) \quad \alpha_t^\varepsilon e^{-C_{\mathcal{B}}t} \leq (H_t + M_t) e^{-C_{\mathcal{B}}t} + \int_0^t (H_s + M_s) C_{\mathcal{B}} e^{-C_{\mathcal{B}}s} ds.$$

Using integration by parts we get

$$\begin{aligned} \int_0^t H_s C_{\mathcal{B}} e^{-C_{\mathcal{B}}s} ds &= -H_t e^{-C_{\mathcal{B}}t} + \int_0^t e^{-C_{\mathcal{B}}s} dH_s \\ &\leq -H_t e^{-C_{\mathcal{B}}t} + \int_0^t dH_s = -H_t e^{-C_{\mathcal{B}}t} + H_t. \end{aligned}$$

Using integration by parts we get $\int_0^t M_s C_{\mathcal{B}} e^{-C_{\mathcal{B}}s} ds = -M_t e^{-C_{\mathcal{B}}t} + L_t$. Using these results in (A.9) we get

$$0 \leq \alpha_t^\varepsilon e^{-C_{\mathcal{B}}t} \leq L_t + H_t.$$

Note that L is a martingale. We have

$$\begin{aligned} \mathbb{E} \sup_{s \in [0, t]} (\alpha_s^\varepsilon e^{-C_{\mathcal{B}}s})^2 &\leq \mathbb{E} \sup_{s \in [0, t]} (L_s + H_s)^2 \leq 2\mathbb{E} \sup_{s \in [0, t]} L_s^2 + 2\mathbb{E} \sup_{s \in [0, t]} H_s^2 \\ &\leq 8\mathbb{E} L_t^2 + 2\mathbb{E} H_t^2 \end{aligned}$$

where in the last step we have used Doob's L^p inequality (Theorem 2.1.7 in [17]) and the fact that H is non-decreasing. Now, using BDG inequality

$$\begin{aligned}\mathbb{E}L_t^2 &= \mathbb{E} \int_0^t e^{-2C_{\mathcal{B}}s} \Sigma_s^2 ds \leq C_{\Sigma} \mathbb{E} \int_0^t e^{-2C_{\mathcal{B}}s} ((\alpha_s^\varepsilon)^2 + \|\mathcal{Y}_s^\varepsilon\|^4) ds, \\ &\leq C_{\Sigma} \int_0^t \mathbb{E} \sup_{u \in [0, s]} (\alpha_u^\varepsilon e^{-C_{\mathcal{B}}u})^2 ds + C_{\Sigma} \int_0^t \mathbb{E} \|\mathcal{Y}_s^\varepsilon\|^4 ds.\end{aligned}$$

Using Holder inequality we have

$$2\mathbb{E}H_t^2 = 2\mathbb{E} \left(C_{\mathcal{B}} \int_0^t \|\mathcal{Y}_s^\varepsilon\|^2 ds \right)^2 \leq 2C_{\mathcal{B}}^2 t \int_0^t \mathbb{E} \|\mathcal{Y}_s^\varepsilon\|^4 ds.$$

Hence,

$$\mathbb{E} \sup_{s \in [0, t]} (\alpha_s^\varepsilon e^{-C_{\mathcal{B}}s})^2 \leq 8C_{\Sigma} \int_0^t \mathbb{E} \sup_{u \in [0, s]} (\alpha_u^\varepsilon e^{-C_{\mathcal{B}}u})^2 ds + (8C_{\Sigma} + 2C_{\mathcal{B}}^2 t) \int_0^t \mathbb{E} \|\mathcal{Y}_s^\varepsilon\|^4 ds.$$

Using Gronwall and then (2.4) we have

$$\begin{aligned}\mathbb{E} \sup_{s \in [0, T]} (\alpha_s^\varepsilon e^{-C_{\mathcal{B}}s})^2 &\leq (8C_{\Sigma} + 2C_{\mathcal{B}}^2 T) T^2 \left(\varepsilon^8 2^3 \left(\frac{K_G K}{\kappa} \right)^4 \mathfrak{C}^4 + 2^3 \mathbb{E} \sup_{s \in [0, T]} (\Upsilon_s^\varepsilon)^4 \right) e^{8C_{\Sigma} T} \\ &\leq C \left(r\varepsilon^\delta \ln \left(\frac{2T}{r\varepsilon^\delta} \right) \right)^2, \quad \text{for small enough } \varepsilon.\end{aligned}$$

Hence

$$\mathbb{E} \sup_{s \in [0, T]} (\alpha_s^\varepsilon)^2 \leq C e^{2C_{\mathcal{B}}T} \left(r\varepsilon^\delta \ln \left(\frac{2T}{r\varepsilon^\delta} \right) \right)^2, \quad \text{for small enough } \varepsilon.$$

A.6. Proof of Proposition 2.7. Following exactly the same technique as for α^ε , we arrive at

$$\mathbb{E} \sup_{s \in [0, t]} (\beta_s e^{-C_{\mathcal{B}}s})^2 \leq 8C_{\Sigma} \int_0^t \mathbb{E} \sup_{u \in [0, s]} (\beta_u e^{-C_{\mathcal{B}}u})^2 ds + (8C_{\Sigma} + 2C_{\mathcal{B}}^2 t) \int_0^t \mathbb{E} \|\mathfrak{Y}_s^\varepsilon\|^4 ds.$$

Using the exponential decay (1.22) we have that $\int_0^t \mathbb{E} \|\mathfrak{Y}_s^\varepsilon\|^4 ds \leq \tilde{K}^4 \int_0^t e^{-4\kappa s/\varepsilon^2} ds \leq \varepsilon^2 (\tilde{K}^4/4\kappa)$ where $\tilde{K} = K \|(I - \hat{\pi}) \mathbf{1}_{\{0\}}\|$. Using Gronwall inequality we have

$$\mathbb{E} \sup_{s \in [0, T]} (\beta_s e^{-C_{\mathcal{B}}s})^2 \leq \varepsilon^2 (8C_{\Sigma} + 2C_{\mathcal{B}}^2 T) T (\tilde{K}^4/4\kappa) e^{8C_{\Sigma} T}.$$

A.7. Proof of theorem 2.8. Using $X^\varepsilon(t) = \Phi(0) e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon(0)$ and Minkowski inequality in (2.8) and then using (A.6) we have

$$\begin{aligned}& \left| X^\varepsilon(t) - (\Phi(0) e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon(0)) \right|^4 \\ & \leq 8 \left(\|\Phi e^{tB/\varepsilon^2} (\mathfrak{z}_t^\varepsilon - \tilde{\mathfrak{z}}_t^\varepsilon)\|^4 + \|\Phi e^{tB/\varepsilon^2} (\tilde{\mathfrak{z}}_t^\varepsilon - \mathfrak{z}_t^\varepsilon)\|^4 + \|y_t^\varepsilon - \mathfrak{Y}_t^\varepsilon\|^4 \right) \\ & \leq 8 \left(\|\mathfrak{z}_t^\varepsilon - \tilde{\mathfrak{z}}_t^\varepsilon\|_2^4 + \|\tilde{\mathfrak{z}}_t^\varepsilon - \mathfrak{z}_t^\varepsilon\|_2^4 + \|\mathfrak{Y}_t^\varepsilon\|^4 \right) \\ & \leq 8 \left(4(\alpha_t^\varepsilon)^2 + 4(\beta_t^\varepsilon)^2 + \|\mathfrak{Y}_t^\varepsilon\|^4 \right).\end{aligned}$$

Combining propositions 2.4, 2.6 and 2.7 and realizing that $(r\varepsilon^\delta \ln(\frac{2T}{r\varepsilon^\delta}))^2 \ll \varepsilon^2$ for small enough ε when $\delta \in (1, 2)$, we get (2.8). Similar is the proof for (2.7).

A.8. Proof of lemma 2.9. Define $\zeta_t^\varepsilon = \frac{1}{2}\|\tilde{\mathfrak{F}}_t^\varepsilon\|_2^2$. Using Ito formula we have $d\zeta_t^\varepsilon = \tilde{\mathcal{B}}_t dt + \tilde{\Sigma}_t dW_t$ where

$$\tilde{\mathcal{B}}_t = \tilde{\Gamma}_t G(\Phi e^{tB/\varepsilon^2} \tilde{\mathfrak{F}}_t^\varepsilon) + \frac{1}{2} \|e^{-tB/\varepsilon^2} \tilde{\Psi}\|_2^2 (L_1(\Phi e^{tB/\varepsilon^2} \tilde{\mathfrak{F}}_t^\varepsilon))^2, \quad \tilde{\Sigma}_t = \tilde{\Gamma}_t L_1(\Phi e^{tB/\varepsilon^2} \tilde{\mathfrak{F}}_t^\varepsilon),$$

and $\tilde{\Gamma}_t = \sum_{i=1}^2 (\tilde{\mathfrak{F}}_t^\varepsilon)_i (e^{-tB/\varepsilon^2} \tilde{\Psi})_i$. Using similar technique as in proof of lemma 2.5 it can be shown that $|\tilde{\mathcal{B}}_t| \leq C_{\tilde{\mathcal{B}}} \zeta_t^\varepsilon$ and $\tilde{\Sigma}_t^2 \leq C_{\tilde{\Sigma}} (\zeta_t^\varepsilon)^2$ where $C_{\tilde{\mathcal{B}}} = 2\|\tilde{\Psi}\|_2 K_G + \|\tilde{\Psi}\|_2^2 \|L_1\|^2$ and $C_{\tilde{\Sigma}} = 4\|\tilde{\Psi}\|_2^2 \|L_1\|^2$. Hence we have

$$\zeta_t^\varepsilon \leq \int_0^t C_{\tilde{\mathcal{B}}} \zeta_s^\varepsilon ds + \tilde{H}_t + \tilde{M}_t, \quad \tilde{H}_t := \zeta_0^\varepsilon, \quad \tilde{M}_t := \int_0^t \tilde{\Sigma}_s dW_s,$$

which is analogous to (A.8). Following the same technique as in section A.5 we get

$$\mathbb{E} \sup_{s \in [0, T]} (\zeta_s^\varepsilon)^2 \leq \mathbb{E} (\zeta_0^\varepsilon)^2 e^{(2C_{\tilde{\mathcal{B}}} + 8C_{\tilde{\Sigma}})T}.$$

Appendix B. PROOFS OF RESULTS IN SECTION 3

B.1. Proof of proposition 3.2. Using the variation of constants formula (1.17) and definition (3.1), we have

$$(B.1) \quad \|\mathcal{Y}_s^\varepsilon\| \leq \left\| \int_0^s \hat{T} \left(\frac{s-u}{\varepsilon^2} \right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \right\| + \Upsilon_s^\varepsilon.$$

For G defined in (1.8) we have

$$(B.2) \quad |G(\eta)| \leq \int |\pi\eta| |d\nu_1| + \int |(1-\pi)\eta| |d\nu_1| + \sum_{j=0}^3 \binom{3}{j} \int_{-r}^0 |\pi\eta|^{3-j} |(1-\pi)\eta|^j |d\nu_3|.$$

For $s \in [0, T \wedge \mathbf{e}^\varepsilon]$ we have that $\|\pi \Pi_s^\varepsilon X^\varepsilon\| \leq C_\varepsilon$. Using this fact and $\|(1-\pi)\Pi_s^\varepsilon X^\varepsilon\| \leq \|\mathcal{Y}_s^\varepsilon\| + \|\mathfrak{Y}_s^\varepsilon\|$ in (B.2), and using inequalities $q \leq 1 + q^3$, $q^2 \leq 1 + q^3$ for $q > 0$; we have for $s \in [0, T \wedge \mathbf{e}^\varepsilon]$

$$|G(\Pi_s^\varepsilon X^\varepsilon)| \leq C(1 + \|\mathcal{Y}_s^\varepsilon\|^3 + \|\mathfrak{Y}_s^\varepsilon\|^3).$$

This C is of the order of C_ε^3 for large C_ε . Now, using the above inequality and the exponential decays (1.14) and (1.25) we have

$$\begin{aligned} \left\| \int_0^s \hat{T} \left(\frac{s-u}{\varepsilon^2} \right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}} G(\Pi_u^\varepsilon X^\varepsilon) du \right\| &\leq \int_0^s \|\hat{T} \left(\frac{s-u}{\varepsilon^2} \right) (1 - \hat{\pi}) \mathbf{1}_{\{0\}}\| |G(\Pi_u^\varepsilon X^\varepsilon)| du \\ &\leq C \int_0^s e^{-\kappa(s-u)/\varepsilon^2} (1 + \|\mathcal{Y}_u^\varepsilon\|^3 + \|\mathfrak{Y}_u^\varepsilon\|^3) du \\ &\leq (C\varepsilon^2/\kappa)(1 + K^3 \|(1-\pi)\Pi_0^\varepsilon X^\varepsilon\|^3/2) + C \int_0^s e^{-\kappa(s-u)/\varepsilon^2} \|\mathcal{Y}_u^\varepsilon\|^3 du. \end{aligned}$$

Plugging the above inequality in (B.1) we have for $s \in [0, T \wedge \mathbf{e}^\varepsilon]$

$$\|\mathcal{Y}_s^\varepsilon\| - \left(C\varepsilon^2 + C \int_0^s e^{-\kappa(s-u)/\varepsilon^2} \|\mathcal{Y}_u^\varepsilon\|^3 du \right) \leq \Upsilon_s^\varepsilon,$$

where C above is of the order of C_ε^3 for large C_ε . For the RHS of the above inequality we use Markov inequality, i.e.

$$\mathbb{P} \left[\sup_{s \in [0, T \wedge \mathbf{e}^\varepsilon]} \Upsilon_s^\varepsilon \geq \varepsilon^a \right] \leq \varepsilon^{-a} \mathbb{E} \left[\sup_{s \in [0, T]} \Upsilon_s^\varepsilon \right]$$

and then proposition 3.1. Then we have the following statement:

Fix $a \in [0, 1)$. For $\delta \in (2a, 2)$, there exists constants $\hat{C} > 0$ (independent of δ and a) and $\varepsilon_\delta > 0$, such that for $\varepsilon < \varepsilon_\delta$

$$\begin{aligned} \mathbb{P} \left[\forall s \in [0, T \wedge \mathbf{e}^\varepsilon], \quad \|\mathcal{Y}_s^\varepsilon\| \leq C\varepsilon^2 + C \int_0^s e^{-\kappa(s-u)/\varepsilon^2} \|\mathcal{Y}_u^\varepsilon\|^3 du + 2\varepsilon^a \right] \\ \geq 1 - \hat{C}\varepsilon^{-a} \sqrt{r\varepsilon^\delta \ln \left(\frac{T}{r\varepsilon^\delta} \right)}. \end{aligned}$$

Using Gronwall kind of inequality (Theorem 2.4.8 in [18]) we have that LHS of above inequality is bounded above by

$$\mathbb{P} \left[\forall s \in [0, T \wedge \mathbf{e}^\varepsilon], \quad \|\mathcal{Y}_s^\varepsilon\| \leq \frac{C\varepsilon^2 + 2\varepsilon^a}{\sqrt{1 - 2 \int_0^s (C\varepsilon^2 + 2\varepsilon^a)^2 C du}} \right]$$

which is bounded above by (for small enough ε , i.e. $\varepsilon \ll (2/C)^{1/(2-a)}$)

$$\mathbb{P} \left[\forall s \in [0, T \wedge \mathbf{e}^\varepsilon], \quad \|\mathcal{Y}_s^\varepsilon\| \leq \frac{4\varepsilon^a}{\sqrt{1 - 2CT(4\varepsilon^a)^2}} \right].$$

which is bounded above by (for small enough ε , i.e. $\varepsilon \ll (1/C)^{1/2a}$)

$$\mathbb{P} \left[\forall s \in [0, T \wedge \mathbf{e}^\varepsilon], \quad \|\mathcal{Y}_s^\varepsilon\| \leq 8\varepsilon^a \right].$$

Hence (3.4) follows.

B.2. Proof of lemma 3.3. Recall that $G(\eta) = \int_{-r}^0 \eta(\theta) d\nu_1(\theta) + \int_{-r}^0 \eta^3(\theta) d\nu_3(\theta)$. For brevity, let \mathbf{e} denote e^{tB/ε^2} . Now,

$$\begin{aligned} & \left| \int_{-r}^0 (\Phi \mathbf{e} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon)^3 d\nu_3 - \int_{-r}^0 (\Phi \widehat{\mathfrak{z}}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon)^3 d\nu_3 \right| \\ &= \left| \int_{-r}^0 (\Phi \mathbf{e} \mathfrak{z}_t^\varepsilon + \mathfrak{Y}_t^\varepsilon + \mathcal{Y}_t^\varepsilon)^3 d\nu_3 - \int_{-r}^0 (\Phi \mathbf{e} \mathfrak{z}_t^\varepsilon + \Phi(\widehat{\mathfrak{z}}_t^\varepsilon - \mathfrak{z}_t^\varepsilon) + \mathfrak{Y}_t^\varepsilon)^3 d\nu_3 \right| \\ &\leq \sum_{j=1}^3 \binom{3}{j} \left| \int_{-r}^0 (\Phi \mathbf{e} \mathfrak{z}_t^\varepsilon)^{3-j} ((\mathfrak{Y}_t^\varepsilon + \mathcal{Y}_t^\varepsilon)^j - (\Phi(\widehat{\mathfrak{z}}_t^\varepsilon - \mathfrak{z}_t^\varepsilon) + \mathfrak{Y}_t^\varepsilon)^j) d\nu_3 \right| \end{aligned}$$

$$\leq \left(3 \sum_{j=0}^2 \|\Phi \mathbf{e}_{\mathfrak{z}_t^\varepsilon}\|^j \right) \left(3 \sum_{j=0}^2 \|\mathfrak{Y}_t^\varepsilon\|^j \right) \left(\int |d\nu_1| + |d\nu_3| \right) \sum_{j=1}^3 (\|\mathfrak{Y}_t^\varepsilon\|^j + \|\Phi \mathbf{e}(\widehat{\mathfrak{z}}_t^\varepsilon - \mathfrak{z}_t^\varepsilon)\|^j).$$

Note that for $t \in [0, T \wedge \mathbf{e}^\varepsilon]$, \mathfrak{z}^ε is bounded. Also, due to the exponential decay (1.25) we have that $\|\mathfrak{Y}_t^\varepsilon\| < K\|(I - \pi)\Pi_0^\varepsilon X^\varepsilon\|$. Hence we have,

$$\left| \int_{-r}^0 (\Phi \mathbf{e}_{\mathfrak{z}_t^\varepsilon} + y_t^\varepsilon)^3 d\nu_3 - \int_{-r}^0 (\Phi \mathbf{e}_{\widehat{\mathfrak{z}}_t^\varepsilon} + \mathfrak{Y}_t^\varepsilon)^3 d\nu_3 \right| \leq C \sum_{j=1}^3 (\|\mathfrak{Y}_t^\varepsilon\|^j + (\sqrt{2\alpha_t^\varepsilon})^j),$$

where C in the above inequality is of the order of C_ε^2 for large C_ε .

Similarly,

$$\left| \int_{-r}^0 (\Phi \mathbf{e}_{\mathfrak{z}_t^\varepsilon} + y_t^\varepsilon) d\nu_1 - \int_{-r}^0 (\Phi \mathbf{e}_{\widehat{\mathfrak{z}}_t^\varepsilon} + \mathfrak{Y}_t^\varepsilon) d\nu_1 \right| \leq C(\|\mathfrak{Y}_t^\varepsilon\| + (\sqrt{2\alpha_t^\varepsilon})).$$

Combining, we get that for $t \in [0, T \wedge \mathbf{e}^\varepsilon]$,

$$|\Gamma_t| \left| G(\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon + y_t^\varepsilon) - G(\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon) \right| \leq |\Gamma_t| C \sum_{j=1}^3 (\|\mathfrak{Y}_t^\varepsilon\|^j + (\sqrt{2\alpha_t^\varepsilon})^j).$$

We have shown $|\Gamma_t| \leq \|\tilde{\Psi}\|_2 \sqrt{2\alpha_t^\varepsilon}$ in section A.4. Hence, if we define \mathfrak{B} by (3.5), then we have, $|\mathcal{B}_t| \leq \mathfrak{B}(\alpha_t, \|\mathfrak{Y}_t^\varepsilon\|)$ for $t \in [0, T \wedge \mathbf{e}^\varepsilon]$.

B.3. Proof of proposition 3.4. Using lemma 3.3 and $\sqrt{2\alpha}(\sqrt{2\alpha})^2 \leq 4\alpha(1 + \alpha)$ we have

$$d\alpha_t^\varepsilon \leq (6C\alpha_t^\varepsilon + 8C(\alpha_t^\varepsilon)^2)dt + C\sqrt{2\alpha_t^\varepsilon} \left(\sum_{j=1}^3 \|\mathfrak{Y}_t^\varepsilon\|^j \right) dt, \quad t \in [0, T \wedge \mathbf{e}^\varepsilon], \quad \alpha_0^\varepsilon = 0,$$

where C is from lemma 3.3. This C is of the order $\mathcal{O}(C_\varepsilon^2)$ for large C_ε . Let $\varepsilon_{a,C_\varepsilon}$ be as in proposition 3.2 and define $\varepsilon_{a,C_\varepsilon,1} = \min\{1, \varepsilon_{a,C_\varepsilon}\}$. Then we have, for $\varepsilon < \min\{\hat{\varepsilon}_\delta, \varepsilon_{a,C_\varepsilon,1}\}$, with probability at least $p_\varepsilon := 1 - \hat{C}\varepsilon^{-a} \sqrt{r\varepsilon^\delta \ln\left(\frac{T}{r\varepsilon^\delta}\right)}$,

$$\sum_{j=1}^3 \|\mathfrak{Y}_t^\varepsilon\|^j \leq 24\varepsilon^a, \quad \forall t \in [0, T \wedge \mathbf{e}^\varepsilon].$$

(We have used that for $\varepsilon \leq 1$, $\varepsilon^{3a} \leq \varepsilon^a$.)

Let $\mathfrak{s} := \inf\{t \geq 0 : \alpha_t^\varepsilon \geq 1\}$. Using $\sqrt{2\alpha} \leq 2(1 + \alpha)$, and $\alpha^2 < \alpha$ when $\alpha < 1$; we have for $t \in [0, T \wedge \mathbf{e}^\varepsilon \wedge \mathfrak{s}]$

$$\frac{1}{C} d\alpha_t^\varepsilon \leq \alpha_t^\varepsilon (6 + 8 + 48\varepsilon^a) dt + 48\varepsilon^a dt$$

Using Gronwall we get for $t \in [0, T \wedge \mathbf{e}^\varepsilon \wedge \mathfrak{s}]$

$$\alpha_t^\varepsilon \leq \frac{48\varepsilon^a}{14 + 48\varepsilon^a} (e^{(14+48\varepsilon^a)Ct} - 1).$$

Define $\varepsilon_{a,C_\epsilon,2} = \min\{48^{-1/a}, (14e^{-15CT}/48)^{2/a}\}$. Then for $\varepsilon < \min\{\hat{\varepsilon}_\delta, \varepsilon_{a,C_\epsilon,1}, \varepsilon_{a,C_\epsilon,2}\}$ we have $\alpha_t^\varepsilon \leq \varepsilon^{a/2}$ for $t \in [0, T \wedge \mathbf{e}^\varepsilon \wedge \mathfrak{s}]$. But, since $\varepsilon^{a/2} < 1$ we have $\mathfrak{s} > T \wedge \mathbf{e}^\varepsilon$ and hence $\alpha_t^\varepsilon \leq \varepsilon^{a/2}$ for $t \in [0, T \wedge \mathbf{e}^\varepsilon]$.

Note that C is of the order $\mathcal{O}(C_\epsilon^2)$ and hence $\varepsilon_{a,C_\epsilon,2}$ is of the order $\mathcal{O}(e^{-30C_\epsilon^2 T/a})$.

B.4. Proof of theorem 3.6. Since the $\widehat{\mathfrak{z}}^\varepsilon$ system possesses the property $\mathcal{P}(T, q)$, $\exists C_\epsilon, \varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$, we have $\mathbb{P}[E^\varepsilon] \geq 1 - q$ where E^ε is given in (3.7).

The stopping time \mathbf{e}^ε was defined at (3.3).

Using

$$\begin{aligned} \Omega &= E^\varepsilon \cup (\Omega \setminus E^\varepsilon) \\ &= (E^\varepsilon \cap \{\mathbf{e}^\varepsilon \leq T\}) \cup (E^\varepsilon \cap \{\mathbf{e}^\varepsilon > T\}) \cup (\Omega \setminus E^\varepsilon), \end{aligned}$$

we have

$$(B.3) \quad H^\varepsilon = (E^\varepsilon \cap \{\mathbf{e}^\varepsilon \leq T\} \cap H^\varepsilon) \cup (E^\varepsilon \cap \{\mathbf{e}^\varepsilon > T\} \cap H^\varepsilon) \cup (H^\varepsilon \cap (\Omega \setminus E^\varepsilon)).$$

Now we deal with the first term on the RHS of (B.3). Note that

$$(B.4) \quad \begin{aligned} \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \sqrt{2\alpha_t^\varepsilon} &\geq \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\Phi e^{tB/\varepsilon^2} (\mathfrak{z}_t^\varepsilon - \widehat{\mathfrak{z}}_t^\varepsilon)\| \\ &\geq \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon\| - \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\|. \end{aligned}$$

In E^ε we have $\sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\| < 0.99C_\epsilon$, and in $\{\mathbf{e}^\varepsilon \leq T\}$ we have

$$\sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\Phi e^{tB/\varepsilon^2} \mathfrak{z}_t^\varepsilon\| \geq C_\epsilon.$$

Hence, in $E^\varepsilon \cap \{\mathbf{e}^\varepsilon \leq T\}$ we have that $\sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \sqrt{2\alpha_t^\varepsilon} > 0.01C_\epsilon$. Hence $E^\varepsilon \cap \{\mathbf{e}^\varepsilon \leq T\} \subset J^\varepsilon$ where $J^\varepsilon := \{\omega : \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \alpha_t^\varepsilon \geq \frac{1}{2}(0.01C_\epsilon)^2\}$. By proposition 3.4, $\exists \varepsilon_1$ such that $\forall \varepsilon < \varepsilon_1$, $\mathbb{P}[J^\varepsilon] < 1 - p_\varepsilon$.

Now we deal with the second term on the RHS of (B.3). Note that $\{\mathbf{e}^\varepsilon > T\} \cap H^\varepsilon \subset \widetilde{J}^\varepsilon$ where

$$\widetilde{J}^\varepsilon := \left\{ \omega : \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \alpha_t^\varepsilon \geq \varepsilon^{a/2} \right\}.$$

By proposition 3.4, $\exists \varepsilon_2$ such that $\forall \varepsilon < \varepsilon_2$, $\mathbb{P}[\widetilde{J}^\varepsilon] < 1 - p_\varepsilon$.

And $\forall \varepsilon < \varepsilon_*$ $\mathbb{P}[\Omega \setminus E^\varepsilon] < q$.

Combining we have, when $\varepsilon < \min\{\varepsilon_1, \varepsilon_2, \varepsilon_*\} =: \varepsilon_q$,

$$\mathbb{P}[H^\varepsilon] < q + 2(1 - p_\varepsilon).$$

Note that (B.3) is true with H^ε replaced by S^ε . Using that

$$\{\mathbf{e}^\varepsilon > T\} \cap S^\varepsilon \subset \{\omega : \sup_{t \in [0, T \wedge \mathbf{e}^\varepsilon]} \|\mathcal{Y}_t^\varepsilon\| \geq 8\varepsilon^a\}$$

and also that the probability of the latter set is bounded above by $1 - p_\varepsilon$ we get the desired result.

B.5. Proof of proposition 3.7. We have

$$(B.5) \quad \widehat{\mathfrak{z}}_t^\varepsilon = \widehat{\mathfrak{z}}_0^\varepsilon + \int_0^t e^{-sB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{sB/\varepsilon^2} \widehat{\mathfrak{z}}_s^\varepsilon + \mathfrak{Y}_s^\varepsilon) ds + \mathfrak{w}_t, \quad \mathfrak{w}_t := \int_0^t e^{-sB/\varepsilon^2} \tilde{\Psi} \sigma dW_s.$$

To keep things simple, we prove assuming $\|\mathfrak{Y}_0^\varepsilon\| = 0$ (which ensures that $\|\mathfrak{Y}_t^\varepsilon\| = 0$ for all $t \geq 0$). Using $\int_0^t \|\mathfrak{Y}_s^\varepsilon\|^n ds \leq \varepsilon^2 (K/nk) \|\mathfrak{Y}_0^\varepsilon\|^n$ (because of exponential decay (1.25)), it is easy to see that the following ideas work even if we assume that $\|\mathfrak{Y}_0^\varepsilon\| \neq 0$ (we assume the initial condition is deterministic).

We will make use of the inequality⁵ that for \mathbb{R}^2 vector v ,

$$(B.6) \quad \|\Phi e^{tB/\varepsilon^2} v\| \leq \|v\|_1$$

where $\|\cdot\|_1$ indicates the 1-norm. Using the structure of G specified at (1.8) in (B.5) we have (with some $K_G > 0$)

$$(B.7) \quad \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 \leq \|\widehat{\mathfrak{z}}_0^\varepsilon\|_1 + \|\tilde{\Psi}\|_1 \int_0^t K_G (\|\widehat{\mathfrak{z}}_s^\varepsilon\|_1 + \|\widehat{\mathfrak{z}}_s^\varepsilon\|_1^3) ds + \|\mathfrak{w}_t\|_1.$$

Because the initial condition is deterministic, we have a $C_0 > 0$ such that $\|\widehat{\mathfrak{z}}_0^\varepsilon\|_1 < C_0$. For any $C_a > 4C_0$, define $T_{C_a} := \left(2(1 + C_a^2)K_G\|\tilde{\Psi}\|_1\right)^{-1}$.

Suppose that $\sup_{t \in [0, T]} \|\mathfrak{w}_t\|_1 < C_a/4$. If $T \leq T_{C_a}$, as long as $\|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 < C_a$, we have (using (B.7)) for $t \in [0, T]$

$$\begin{aligned} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 &< C_0 + K_G \|\tilde{\Psi}\|_1 (C_a + C_a^3) T + \frac{1}{4} C_a \\ &< \frac{1}{4} C_a + K_G \|\tilde{\Psi}\|_1 (C_a + C_a^3) T_{C_a} + \frac{1}{4} C_a = C_a. \end{aligned}$$

This means that, if $C_a > 4C_0$ and $T \leq T_{C_a}$, then we have $\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 < C_a$ provided $\sup_{t \in [0, T]} \|\mathfrak{w}_t\|_1 < C_a/4$.

Hence, for $C_a > 4C_0$ and $T \leq T_{C_a}$,

$$(B.8) \quad \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 \geq C_a \right] \leq \mathbb{P} \left[\sup_{t \in [0, T]} \|\mathfrak{w}_t\|_1 \geq C_a/4 \right].$$

Using Markov inequality and Burkholder-Davis-Gundy inequality we have

$$\begin{aligned} \mathbb{P} \left[\sup_{t \in [0, T]} \|\mathfrak{w}_t\|_1 \geq C_a/4 \right] &\leq \frac{\mathbb{E} \sup_{t \in [0, T]} \|\mathfrak{w}_t\|_1}{C_a/4} \\ &\leq \frac{\sum_{j=1}^2 C_{bdg} \mathbb{E} \sqrt{\int_0^T (e^{-sB/\varepsilon^2} \tilde{\Psi} \sigma)^2 ds}}{C_a/4} \leq \frac{8|\sigma| \|\tilde{\Psi}\|_2 \sqrt{T} C_{bdg}}{C_a}. \end{aligned}$$

⁵Note that $\|\Phi e^{tB/\varepsilon^2} v\| = \sup_{\theta \in [-r, 0]} |v_1 \cos(\theta + \omega_c t/\varepsilon^2) + v_2 \sin(\theta + \omega_c t/\varepsilon^2)| \leq |v_1| + |v_2|$.

Using the above inequality in (B.8) we have for $C_a > 4C_0$ and $T \leq T_{C_a}$

$$(B.9) \quad \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 \geq C_a \right] \leq \frac{8|\sigma| \|\tilde{\Psi}\|_2 C_{bdg}}{C_a \sqrt{2K_G \|\tilde{\Psi}\|_1 (1 + C_a^2)}} =: f(C_a).$$

Given $q > 0$, let $C_{a,q} > 4C_0$ be such that $f(C_a) < q$, $\forall C_a \geq C_{a,q}$. Such a $C_{a,q}$ exists because f is monotonically decreasing in C_a . Set $T_q = T_{C_{a,q}}$. Choose $C_\varepsilon > C_{a,q}/0.99$. Let

$$(B.10) \quad \widetilde{E}^\varepsilon := \left\{ \omega : \sup_{t \in [0, T_q]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\| < 0.99C_\varepsilon \right\}.$$

Now, using (B.6) and (B.9)

$$\begin{aligned} \mathbb{P}[\Omega \setminus \widetilde{E}^\varepsilon] &= \mathbb{P} \left[\sup_{t \in [0, T_q]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\| \geq 0.99C_\varepsilon \right] \\ &\leq \mathbb{P} \left[\sup_{t \in [0, T_q]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 \geq 0.99C_\varepsilon \right] \leq \mathbb{P} \left[\sup_{t \in [0, T_q]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_1 \geq C_{a,q} \right] \leq f(C_{a,q}) < q. \end{aligned}$$

Hence $\mathbb{P}[\widetilde{E}^\varepsilon] \geq 1 - q$. But, for $T \leq T_q$ the set E^ε defined in (3.7) contains $\widetilde{E}^\varepsilon$ and hence we have that for $T \in [0, T_q]$, $\mathbb{P}[E^\varepsilon] \geq 1 - q$. Hence (1.8) possesses the property $\mathcal{P}(T, q)$ for $T \in [0, T_q]$.

B.6. Proof of proposition 3.8. To keep things simple, we prove assuming $\|\mathfrak{Y}_0^\varepsilon\| = 0$ (which ensures that $\|\mathfrak{Y}_t^\varepsilon\| = 0$ for all $t \geq 0$). Using $\int_0^t \|\mathfrak{Y}_s^\varepsilon\|^n ds \leq \varepsilon^2 (K/nk) \|\mathfrak{Y}_0^\varepsilon\|^n$ (because of exponential decay (1.25)), it is easy to see that the following ideas work even if we assume that $\|\mathfrak{Y}_0^\varepsilon\| \neq 0$.

For simplicity of notation we write $G = G_1 + G_3$ where G_1 is the linear part and G_3 is the cubic part.

We have

$$\widehat{\mathfrak{z}}_t^\varepsilon = \widehat{\mathfrak{z}}_0^\varepsilon + \int_0^t e^{-sB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{sB/\varepsilon^2} \widehat{\mathfrak{z}}_s^\varepsilon) ds + \mathfrak{w}_t, \quad \mathfrak{w}_t := \int_0^t e^{-sB/\varepsilon^2} \tilde{\Psi} \sigma dW_s.$$

Writing $\eta_t = \widehat{\mathfrak{z}}_t^\varepsilon - \mathfrak{w}_t$, we have

$$(B.11) \quad \dot{\eta}_t = e^{-tB/\varepsilon^2} \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} (\eta_t + \mathfrak{w}_t))$$

from which we can write (using that the transpose of e^{tB/ε^2} is e^{-tB/ε^2})

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\eta_t\|_2^2 &= (e^{tB/\varepsilon^2} \eta_t)^* \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} (\eta_t + \mathfrak{w}_t)) \\ &= (e^{tB/\varepsilon^2} \eta_t)^* \tilde{\Psi} G(\Phi e^{tB/\varepsilon^2} \eta_t) \\ &\quad + (e^{tB/\varepsilon^2} \eta_t)^* \tilde{\Psi} \left(G(\Phi e^{tB/\varepsilon^2} (\eta_t + \mathfrak{w}_t)) - G(\Phi e^{tB/\varepsilon^2} \eta_t) \right). \end{aligned}$$

Using $G = G_1 + G_3$, and the Lipschitz condition on the linear part $|G_1(\eta_1) - G_1(\eta_2)| \leq K_G \|\eta_1 - \eta_2\|$, and that $\|\Phi e^{tB/\varepsilon^2} \mathbf{w}_t\| \leq \|\mathbf{w}_t\|_1$, and $|(e^{tB/\varepsilon^2} \boldsymbol{\eta}_t)^* \tilde{\Psi}| \leq \|\tilde{\Psi}\|_2 \|\boldsymbol{\eta}_t\|_2$, we have

$$(B.12) \quad \frac{1}{2} \frac{d}{dt} \|\boldsymbol{\eta}_t\|_2^2 \leq K_G \|\tilde{\Psi}\|_2 \|\boldsymbol{\eta}_t\|_2^2 + (e^{tB/\varepsilon^2} \boldsymbol{\eta}_t)^* \tilde{\Psi} G_3(\Phi e^{tB/\varepsilon^2} \boldsymbol{\eta}_t) \\ + \|\tilde{\Psi}\|_2 \|\boldsymbol{\eta}_t\|_2 K_G \|\mathbf{w}_t\|_1 + \sum_{j=1}^3 c_j \|\boldsymbol{\eta}_t\|_2^{4-j} \|\mathbf{w}_t\|_1^j,$$

for some constants $c_j > 0$.

Define the time averaging operator \mathbb{T} as follows: For a periodic function $f : \mathbb{R} \rightarrow \mathbb{R}$ with period $2\pi/\omega_c$, the action of \mathbb{T} is given by $\mathbb{T}(f) = \frac{1}{2\pi/\omega_c} \int_0^{2\pi/\omega_c} f(s) ds$. Note that the condition (3.8) means that

$$(B.13) \quad \mathbb{T} \left((e^{\cdot B} z)^* \tilde{\Psi} G_3(\Phi e^{\cdot B} z) \right) < -C_G \|z\|_2^4.$$

Define

$$(B.14) \quad \tilde{G}_3(z, t) := (e^{tB} z)^* \tilde{\Psi} G_3(\Phi e^{tB} z) - \mathbb{T}((e^{\cdot B} z)^* \tilde{\Psi} G_3(\Phi e^{\cdot B} z)).$$

Then, using (B.14) and (B.13) in (B.12) we have

$$\frac{1}{2} \frac{d}{dt} \|\boldsymbol{\eta}_t\|_2^2 \leq K_G \|\tilde{\Psi}\|_2 \|\boldsymbol{\eta}_t\|_2^2 - C_G \|\boldsymbol{\eta}_t\|_2^4 + \tilde{G}_3(\boldsymbol{\eta}_t, t/\varepsilon^2) \\ + \|\tilde{\Psi}\|_2 \|\boldsymbol{\eta}_t\|_2 K_G \|\mathbf{w}_t\|_1 + \sum_{j=1}^3 c_j \|\boldsymbol{\eta}_t\|_2^{4-j} \|\mathbf{w}_t\|_1^j.$$

Using Young's inequality we have for some $C_Y > 0$

$$(B.15) \quad \frac{1}{2} \frac{d}{dt} \|\boldsymbol{\eta}_t\|_2^2 < -\frac{1}{2} C_G \|\boldsymbol{\eta}_t\|_2^4 + C_Y \|\mathbf{w}_t\|_1^4 + C_Y + \tilde{G}_3(\boldsymbol{\eta}_t, t/\varepsilon^2).$$

Assume

$$(B.16) \quad \sup_{t \in [0, T]} \|\mathbf{w}_t\|_1^4 < R,$$

and let $\tilde{C} = C_Y(1 + R)$. Then

$$(B.17) \quad \frac{1}{2} \frac{d}{dt} \|\boldsymbol{\eta}_t\|_2^2 < -\frac{1}{2} C_G \|\boldsymbol{\eta}_t\|_2^4 + \tilde{C} + \tilde{G}_3(\boldsymbol{\eta}_t, t/\varepsilon^2).$$

Using comparison principle (theorem 6.1 on page 31 of [19]), we have that $\|\boldsymbol{\eta}_t\|_2^2 \leq \mathbf{v}_t$ where \mathbf{v} is governed by

$$(B.18) \quad \frac{d}{dt} \mathbf{v}_t = -C_G \mathbf{v}_t + C_G(\mathbf{v}_t - \mathbf{v}_t^2) + 2\tilde{C} + 2\tilde{G}_3(\boldsymbol{\eta}_t, t/\varepsilon^2), \quad \mathbf{v}_0 = \|\boldsymbol{\eta}_0\|_2^2.$$

Using variation-of-constants formula, the fact that $\mathbf{v} - \mathbf{v}^2 < 1$, and integration-by-parts, we find that

$$(B.19) \quad \mathbf{v}_t < \mathbf{v}_0 e^{-C_G t} + \frac{2\tilde{C} + C_G}{C_G} (1 - e^{-C_G t}) + 2 \int_0^t \tilde{G}_3(\boldsymbol{\eta}_s, s/\varepsilon^2) ds$$

$$- C_G \int_0^t e^{-C_G(t-s)} \left(2 \int_0^s \tilde{G}_3(\boldsymbol{\eta}_u, u/\varepsilon^2) du \right) ds.$$

Now we try to obtain some bounds on the last two terms of the above inequality.

Using the structure of G_3 (defined in (1.8)) and \tilde{G}_3 (defined in (B.14)) and that $\mathbb{T}(\tilde{G}_3(z, \cdot)) = 0$, it is easy to see that \tilde{G}_3 can be expressed as

$$\tilde{G}_3(\boldsymbol{\eta}, t) = \sum_{j=1}^4 (\alpha_j \cos(j\omega_c t) + \beta_j \sin(j\omega_c t))$$

where α_j and β_j are fourth order polynomials in the components of $\boldsymbol{\eta}$. Define

$$\mathbf{g}(z, t) := 2 \int_0^t \tilde{G}_3(z, s) ds.$$

Using the structure of \tilde{G}_3 it is easy to see that (note that \tilde{G}_3 is mean zero and periodic as a function of its second argument) there exists $C_{\mathbf{g}} > 0$ such that

$$|\mathbf{g}(\boldsymbol{\eta}, t)| \leq C_{\mathbf{g}}(1 + \|\boldsymbol{\eta}\|_2^4), \quad \left\| \frac{\partial \mathbf{g}}{\partial \boldsymbol{\eta}}(\boldsymbol{\eta}, t) \right\|_2 \leq C_{\mathbf{g}}(1 + \|\boldsymbol{\eta}\|_2^3).$$

Also, from (B.11), it is easy to see that $\exists C_* > 0$ such that $\|\dot{\boldsymbol{\eta}}\|_2 \leq C_*(1 + \|\boldsymbol{\eta} + \boldsymbol{\mathfrak{w}}\|_2^3)$. Since

$$\varepsilon^2 \mathbf{g}(\boldsymbol{\eta}_t, t/\varepsilon^2) - \varepsilon^2 \mathbf{g}(\boldsymbol{\eta}_0, 0) - \varepsilon^2 \int_0^t \frac{\partial \mathbf{g}}{\partial \boldsymbol{\eta}}(\boldsymbol{\eta}_s, s/\varepsilon^2) \dot{\boldsymbol{\eta}}_s ds = 2 \int_0^t \tilde{G}_3(\boldsymbol{\eta}_s, s/\varepsilon^2) ds,$$

and $\mathbf{g}(\boldsymbol{\eta}, 0) = 0$, we have

$$\begin{aligned} \left| 2 \int_0^t \tilde{G}_3(\boldsymbol{\eta}_s, s/\varepsilon^2) ds \right| &\leq \varepsilon^2 C_{\mathbf{g}}(1 + \|\boldsymbol{\eta}_t\|_2^4) + \varepsilon^2 C_{\mathbf{g}} C_* \int_0^t (1 + \|\boldsymbol{\eta}_s\|_2^3)(1 + \|\boldsymbol{\eta}_s + \boldsymbol{\mathfrak{w}}_s\|_2^3) ds \\ &\leq \varepsilon^2 C_{\mathbf{g}}(1 + \|\boldsymbol{\eta}_t\|_2^4) + \varepsilon^2 11 C_{\mathbf{g}} C_* \int_0^t (1 + \|\boldsymbol{\eta}_s\|_2^6 + \|\boldsymbol{\mathfrak{w}}_s\|_1^6) ds. \end{aligned}$$

Let

$$(B.20) \quad \tau^\varepsilon := \inf \left\{ t \geq 0 : \|\boldsymbol{\eta}_t\|_2 \geq \frac{1}{\varepsilon^{1/6}} \right\}.$$

Then, for $t \leq \min\{\tau^\varepsilon, T\}$ we have

$$\left| 2 \int_0^t \tilde{G}_3(\boldsymbol{\eta}_s, s/\varepsilon^2) ds \right| \leq \varepsilon^2 C_{\mathbf{g}} + \varepsilon^{4/3} C_{\mathbf{g}} + \varepsilon^2 11 C_{\mathbf{g}} C_* \int_0^t (1 + \varepsilon^{-1} + R^{3/2}) ds.$$

When $\varepsilon < 1$, we have (from the above inequality) that for $t \leq \tau^\varepsilon \wedge T$

$$(B.21) \quad \left| 2 \int_0^t \tilde{G}_3(\boldsymbol{\eta}_s, s/\varepsilon^2) ds \right| \leq \varepsilon^{4/3} 2 C_{\mathbf{g}} + \varepsilon \widehat{C} t$$

where $\widehat{C} = 22 C_{\mathbf{g}} C_*(1 + R^{3/2})$.

Using (B.21) in (B.19), we have⁶ for $\varepsilon < 1$ and $t \leq \tau^\varepsilon \wedge T$

$$(B.22) \quad \begin{aligned} \|\eta_t\|_2^2 &\leq \mathbf{v}_t < \mathbf{v}_0 e^{-C_G t} + \frac{2\tilde{C} + C_G}{C_G} (1 - e^{-C_G t}) + \varepsilon 2\widehat{C}t + \varepsilon^{4/3} 4C_g \\ &< \max\{\|\eta_0\|_2^2, \frac{2\tilde{C} + C_G}{C_G}\} + \varepsilon 2(\widehat{C}T + 2C_g). \end{aligned}$$

Hence, for $\varepsilon < 1$ and $t \leq \tau^\varepsilon \wedge T$

$$\|\eta_t\|_2 < \|\eta_0\|_2 + 1 + \sqrt{\frac{2\tilde{C}}{C_G}} + \sqrt{\varepsilon} \sqrt{2\widehat{C}T} + \sqrt{\varepsilon} \sqrt{4C_g}.$$

Using $\widehat{C} = 22C_g C_*(1 + R^{3/2})$ and that $\tilde{C} = C_Y(1 + R)$ we find that

$$\|\eta_t\|_2 < \|\eta_0\|_2 + (1 + \sqrt{\varepsilon} \sqrt{4C_g}) + \sqrt{\frac{2C_Y}{C_G}} \sqrt{1 + R} + \sqrt{\varepsilon} \sqrt{44C_g C_* T} (1 + R^{3/4}).$$

Note that $\exists \varepsilon_{(2)}$ such that $\forall \varepsilon < \varepsilon_{(2)}$ we have $\sqrt{\varepsilon} \sqrt{4C_g} < 1$. Also, $\exists \varepsilon_{(3)}$ such that $\forall \varepsilon < \varepsilon_{(3)}$ we have $\sqrt{\varepsilon} \sqrt{\frac{44C_g C_* T}{2C_Y/C_G}} < 1$. Hence, for $\varepsilon < \min\{1, \varepsilon_{(2)}, \varepsilon_{(3)}\} =: \varepsilon_{(4)}$ and $t \leq \tau^\varepsilon \wedge T$ we have⁷

$$(B.23) \quad \|\eta_t\|_2 < \|\eta_0\|_2 + 2 + 6\sqrt{\frac{2C_Y}{C_G}} \sqrt{1 + R^{6/4}}.$$

Hence, for $\varepsilon < \varepsilon_{(4)}$ if $\tau^\varepsilon \geq T$ we have (using $\|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 \leq \|\eta_t\|_2 + \|\mathbf{w}_t\|_1$)

$$\begin{aligned} \sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 &< \|\widehat{\mathfrak{z}}_0^\varepsilon\|_2 + 2 + 6\sqrt{\frac{2C_Y}{C_G}} \sqrt{1 + R^{6/4}} + R^{1/4} \\ &< \|\widehat{\mathfrak{z}}_0^\varepsilon\|_2 + 2 + 6\left(\frac{1}{6} + \sqrt{\frac{2C_Y}{C_G}}\right) \sqrt{1 + R^{6/4}} \\ &=: \|\widehat{\mathfrak{z}}_0^\varepsilon\|_2 + 2 + C_{YG} \sqrt{1 + R^{6/4}} \end{aligned}$$

Because the initial condition is deterministic $\exists C_0 > 0$ such that $\|\widehat{\mathfrak{z}}_0^\varepsilon\|_2 < C_0$. Hence

$$(B.24) \quad \sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 < C_0 + 2 + C_{YG} \sqrt{1 + R^{6/4}}.$$

Define C_R by

$$C_R \stackrel{\text{def}}{=} C_0 + 2 + C_{YG} \sqrt{1 + R^{6/4}}.$$

For $\varepsilon < \varepsilon_{(4)}$ and $t \leq \tau^\varepsilon \wedge T$, we have from (B.23) that $\|\eta_t\|_2 < C_R$. So, if we define $\varepsilon_R := (1/C_R)^6$, then for $\varepsilon < \min\{\varepsilon_{(4)}, \varepsilon_R\}$ we have that $\|\eta_t\|_2 < \frac{1}{\varepsilon^{1/6}}$ and hence $\tau^\varepsilon > T$.

⁶For the last term in RHS of (B.19) we have used that

$$|C_G \int_0^t e^{-C_G(t-s)} f(s) ds| \leq \left(\sup_{s \in [0, t]} |f(s)| \right) C_G \int_0^t e^{-C_G(t-s)} ds \leq \left(\sup_{s \in [0, t]} |f(s)| \right).$$

⁷We use $\sqrt{1 + R} + (1 + R^{3/4}) < 6\sqrt{1 + R^{6/4}}$.

Hence, from (B.24) we have that for $\varepsilon < \min\{\varepsilon_{(4)}, \varepsilon_R\}$

$$(B.25) \quad \sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 < C_R.$$

Recalling the definition of R from (B.16) we have for $\varepsilon < \min\{\varepsilon_{(4)}, \varepsilon_R\}$

$$(B.26) \quad \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 \geq C_R \right] \leq \mathbb{P} \left[\sup_{s \in [0, T]} \|\mathfrak{w}_t\|_1 \geq \left(\left(\frac{C_R - C_0 - 2}{C_{YG}} \right)^2 - 1 \right)^{1/6} \right].$$

Lets estimate the RHS of the above equation. Using the definition of \mathfrak{w} and then Markov and Burkholder-Davis-Gundy inequalities we have that

$$\begin{aligned} \mathbb{P} \left[\sup_{s \in [0, T]} \|\mathfrak{w}_t\|_1 \geq \rho \right] &\leq \sum_{j=1}^2 \mathbb{P} \left[\sup_{s \in [0, T]} \left| \int_0^t (e^{-sB/\varepsilon^2} \tilde{\Psi})_j \sigma dW_s \right| \geq \frac{1}{2} \rho \right] \\ &\leq \frac{2}{\rho} \sum_{j=1}^2 \mathbb{E} \left[\sup_{s \in [0, T]} \left| \int_0^t (e^{-sB/\varepsilon^2} \tilde{\Psi})_j \sigma dW_s \right| \right] \\ &\leq \frac{2C_{bdg}}{\rho} \sum_{j=1}^2 \mathbb{E} \sqrt{\int_0^t (e^{-sB/\varepsilon^2} \tilde{\Psi})_j^2 \sigma^2 ds} \leq \frac{2\sqrt{2}C_{bdg}}{\rho} \|\tilde{\Psi}\|_2 \sigma \sqrt{T} \end{aligned}$$

Hence, from (B.26) we have that for $\varepsilon < \min\{\varepsilon_{(4)}, \varepsilon_R\}$

$$(B.27) \quad \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 \geq C_R \right] \leq \frac{2\sqrt{2}C_{bdg} \|\tilde{\Psi}\|_2 \sigma \sqrt{T}}{\left(\left(\frac{C_R - C_0 - 2}{C_{YG}} \right)^2 - 1 \right)^{1/6}} \stackrel{\text{def}}{=} f(C_R).$$

Let $C_{R,q} > C_0 + 2 + C_{YG}$ be such that $f(C_R) < q$ for $C_R > C_{R,q}$. Such a $C_{R,q}$ exists because f is monotonically decreasing in C_R (for $C_R > C_0 + 2 + C_{YG}$). Choose $C_\varepsilon > C_{R,q}/0.99$ and $\varepsilon_* = \min\{\varepsilon_{(4)}, (C_{R,q})^{-6}\}$. Let E^ε be as defined in (3.7). Now, using (B.27) for $\varepsilon < \varepsilon_*$

$$\begin{aligned} \mathbb{P}[\Omega \setminus E^\varepsilon] &= \mathbb{P} \left[\sup_{t \in [0, T]} \|\Phi e^{tB/\varepsilon^2} \widehat{\mathfrak{z}}_t^\varepsilon\| \geq 0.99C_\varepsilon \right] \\ &\leq \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 \geq 0.99C_\varepsilon \right] \leq \mathbb{P} \left[\sup_{t \in [0, T]} \|\widehat{\mathfrak{z}}_t^\varepsilon\|_2 \geq C_{R,q} \right] \leq f(C_{R,q}) < q. \end{aligned}$$

Hence $\mathbb{P}[E^\varepsilon] \geq 1 - q$, and so (1.8) possesses the property $\mathcal{P}(T, q)$. As mentioned above, for any $q > 0$ it is possible to select ε_* and $C_{R,q}$ such that $f(C_{R,q}) < q$ and hence $\widehat{\mathfrak{z}}^\varepsilon$ possesses the property $\mathcal{P}(T, q)$ for arbitrary $q > 0$.

B.7. Proof of proposition 3.9. Using Ito formula, β_t^ε satisfies

$$d\beta_t^\varepsilon = \mathcal{B}_t dt, \quad \beta_0^\varepsilon = 0,$$

where

$$\mathcal{B}_t = \Gamma_t \left(G(\Phi e^{tB/\varepsilon^2} \tilde{\mathfrak{z}}_t^\varepsilon) - G(\Phi e^{tB/\varepsilon^2} \hat{\mathfrak{z}}_t^\varepsilon + \mathfrak{y}_t) \right), \quad \Gamma_t = \sum_{i=1}^2 \left(e^{-tB/\varepsilon^2} \tilde{\Psi} \right)_i (\tilde{\mathfrak{z}}_t^\varepsilon - \hat{\mathfrak{z}}_t^\varepsilon)_i.$$

Using the structure of e^{-tB} and $\tilde{\Psi}$, we have $|\Gamma_t| \leq \sqrt{\tilde{\Psi}^* \tilde{\Psi}} \sqrt{2\beta_t^\varepsilon}$. Writing $\tilde{\mathfrak{z}}_t^\varepsilon$ as $\hat{\mathfrak{z}}_t^\varepsilon + (\tilde{\mathfrak{z}}_t^\varepsilon - \hat{\mathfrak{z}}_t^\varepsilon)$ and expanding G in \mathcal{B}_t we get

$$|\mathcal{B}_t| \leq C \sqrt{\beta_t^\varepsilon} \left(\sqrt{\beta_t^\varepsilon} + \|\mathfrak{y}_t^\varepsilon\| + \sum_{j=1}^3 \|\Phi e^{tB/\varepsilon^2} \hat{\mathfrak{z}}_t^\varepsilon\|^{3-j} \left((\sqrt{\beta_t^\varepsilon})^j + \|\mathfrak{y}_t^\varepsilon\|^j \right) \right)$$

Because the nonlinearity is such that (3.8) is satisfied, by lemma 3.8, $\hat{\mathfrak{z}}^\varepsilon$ possesses property $\mathcal{P}(T, q)$ for arbitrary $q > 0$. Hence, it is possible to select $C_\varepsilon > 0$ so that $\exists \varepsilon_* > 0$ such that $\forall \varepsilon < \varepsilon_*$, we have $\mathbb{P}[E^\varepsilon] \geq 1 - q$ where E^ε is defined in (3.7). So, with probability at least $1 - q$ we have

$$|\mathcal{B}_t| \leq C \sqrt{\beta_t^\varepsilon} \left((1 + C_\varepsilon^2) \sqrt{\beta_t^\varepsilon} + C_\varepsilon (\sqrt{\beta_t^\varepsilon})^2 + (\sqrt{\beta_t^\varepsilon})^3 + (1 + C_\varepsilon^2) \sum_{j=1}^3 \|\mathfrak{y}_t^\varepsilon\|^j \right).$$

Let $\mathcal{C} := \|(1 - \pi)\Pi_0^\varepsilon X^\varepsilon\|$. As long as $\sqrt{\beta_t^\varepsilon} < 1$ we have

$$|\mathcal{B}_t| \leq C \sqrt{\beta_t^\varepsilon} \left((3 + 2C_\varepsilon^2) \sqrt{\beta_t^\varepsilon} + (1 + C_\varepsilon^2) \left(\sum_{j=1}^3 \mathcal{C}^j \right) e^{-kt/\varepsilon^2} \right).$$

Hence, as long as $\sqrt{\beta_t^\varepsilon} < 1$

$$d\sqrt{\beta_t^\varepsilon} \leq C(3 + 2C_\varepsilon^2) \sqrt{\beta_t^\varepsilon} + C(1 + C_\varepsilon^2) \left(\sum_{j=1}^3 \mathcal{C}^j \right) e^{-kt/\varepsilon^2}.$$

Using Gronwall, we get (as long as $\sqrt{\beta_t^\varepsilon} < 1$)

$$\begin{aligned} \sqrt{\beta_t^\varepsilon} &\leq \frac{\varepsilon^2 C(1 + C_\varepsilon^2) (\sum_{j=1}^3 \mathcal{C}^j)}{k + \varepsilon^2 C(3 + 2C_\varepsilon^2)} \left(e^{C(3+2C_\varepsilon^2)t} - e^{-kt/\varepsilon^2} \right) \\ &< \frac{\varepsilon^2 C(1 + C_\varepsilon^2) (\sum_{j=1}^3 \mathcal{C}^j)}{k} e^{C(3+2C_\varepsilon^2)T} \end{aligned}$$

Choose ε_{**} small enough so that the above expression is less than one. Set $\varepsilon_\circ = \min\{\varepsilon_*, \varepsilon_{**}\}$. Then, we have $\forall \varepsilon < \varepsilon_\circ$, $\sup_{t \in [0, T]} \beta_t^\varepsilon < C\varepsilon^4$ with probability at least $1 - q$. \square