Concentration Inequalities for Nonautonomous Stochastic Delay Differential Equations

Dissertation

Fakultät für Mathematik
Universität Bielefeld

Juni 2017

Daniel Altemeier

To my family

Maren, Hannah, Paul

Acknowledgment

First and foremost, I would like to express my greatest gratitude to my supervisor Prof. Dr. Barbara Gentz for patience in times of confusion, for guidance in time of disorientation, and for always having a timeframe in need.

Moreover, I thank a couple of extraordinary people that have always been, or have become part of my life through the recent years. First of all, I feel deeply grateful for my parents Hans and Christel for a lifetime of confidence and support.
Next to them, my younger brother Bastian, a truly wonderful and warmhearted person, who has managed to constantly encourage me in a variety of situations including my study and scientific work. My colleague Christian Wiesel, who is a magnificent person and a brilliant mind, and I will miss our daily conversation. Moreover, Diana Kämpfe, who regularly witnessed my cluelessness and despair. I would particularly like to thank her for her continuous patience.

Further, I feel deeply grateful to Dr. Katharina von der Lühe, who has been a close friend since we both started our studies. I thank her for the large amount of typos and inaccuracies, that she sorted out during the proofreading. What remains of those is alone on my responsibility.

Contents

List of Symbols

$\lvert\cdot\rvert$	The 2-norm in $\mathbb{R}^n$, absolute value in $n = 1$.
$\lVert\cdot\rVert_A, \lVert\cdot\rVert$	Supremum or uniform norm over A; when there is no ambiguity, the symbol A is suppressed.
$\mathcal{B}^b(A, B)$	Bounded and Borel-measurable functions from A to B.
J	$J = [-r, 0]$.
$\mathcal{C}(A, C)$	Continuous C-valued functions over A.
$\mathscr{L}^p(A, C)$	p-times Lebesgue-integrable functions $x : A \to C$ which means $\int_A \lvert x(u)\rvert^p du < \infty$,
$\mathscr{L}^p_{\text{loc}}(A, C)$	Locally p-times Lebesgue-integrable functions,
$\mathbb{N}$	Natural numbers not including zero
$\mathbb{C}$	Complex numbers
$\mathbb{R}$, $\mathbb{R}_+$	Real numbers, nonnegative real numbers.
$\mathcal{L}^p(A, C)$	Adapted C-valued processes x with $\mathbb{E}\left[\int_A \lvert x(u)\rvert^p du\right] < \infty$.
$\mathcal{C}^b_{\mathcal{F}_{t_0}}(J, \mathbb{R}^d)$	Bounded $\mathcal{F}_{t_0}$-measurable, $\mathcal{C}(J, \mathbb{R}^d)$-valued random variables
$\mathcal{L}^p_{\mathcal{F}_t}(J, \mathbb{R}^d)$	Family of all $\mathcal{F}_t$-measurable $\mathcal{C}(J, \mathbb{R}^d)$-valued random variables $\varphi = (\varphi(u) : u \in J)$ such that $\mathbb{E}[\lVert\varphi\rVert^p] < \infty$.
$\mathcal{C}^{p,q}_b$	Continuous functions with two arguments and continuous and bounded p-th derivative first and continuous and bounded q-th derivative in the second argument.
$\mathcal{O}(\cdot)$	Landau Symbol, $f \in \mathcal{O}(x)$ existence of a constant, $c_0, d_0 > 0$ such that $\lvert f(x)\rvert \le c\lvert x\rvert$ for all $\lvert x\rvert < d_0$.
$\mathcal{C}^{(h_0)}$	Set of continuous functions with initial segment h_0, i.e. $f \in \mathcal{C}^{(h_0)}\big([t_0, t_1], \mathbb{R}\big)$, if $f \in \mathcal{C}\big([t_0 - r, t_1], \mathbb{R}\big)$ and $f(t_0 + u) = h_0(u)$ for all $u \in J$.
S	Stability area: Set of coefficient combinations in autonomous SDDEs that provide stability, see Figure 4 in Chapter 4.
$\sim$	Proportionality, $f \sim g$ means $f/g \to 1$.
$\inf\emptyset$	We generally assume that the infimum of an empty set is $+\infty$.

1. Introduction

In rigorously establishing the notion of a stochastic integral K. Itô smoothed the way for applicationers to represent effects, for instance due to imperfect information or imprecise measurement, into their mathematical models by *noise* in a formally sound way. And in places where those applicationers had previously pursued stability results for ordinary differential systems, the spotlight fell on new concepts in order to compensate the unsatisfiable desire to bound processes pathwise for instance by the concept that certain bounds hold with high probability. And where observation times of ODEs had only played a minor role unless the system had been significantly changing with time, the stochastic integral established a sort of *inner clock* to the classical deterministic perturbation theory. Even for time-stationary systems it is no longer exhausting to ask *if* corresponding solutions feature *interesting* behavior, but it turns out naturally to ask how long it takes such systems to do something exciting. Exemplary, one might think of a particle movement driven by differential law due to the symmetric one-dimensional double-well potential. It is kind of hard to think of something interesting to ask, to observe or to say about that particle left only to the potential. But by adding only the slightest amount of white noise, the particle hops from one well to the other, regularly in terms of the *Kramer's times*. It is evident that classical stability concepts for deterministic systems are of fairly limited use in the study of noisy systems. Further, the introduction of a time-delayed argument in the formulation of a differential law reflects the idea that a system's evolution is influenced from a prior state of the system itself. Early motivation and has conveniently arisen in biology, chemistry, and mechanical engineering. There, a time-delayed argument has natural applications in the description of real-world systems which evolve depending on a prior state through memory, duration of signal transaction, reaction duration, minimal response time, or gestation period.

To describe the behavior of a system subject to stochastic perturbation there are several well-established techniques like the Fokker-Planck approach, which can provide insights about the stationary distributions and the transition probabilities of a system, [KS91], [SV06]. Also the large-deviation theory has approved as a powerful tool in various situations. It often provides sharp estimates of the probability of atypical or *rare* events of a solution path in terms of exponential rates, [DZ92], [Fre12]. Another main tool for the description of such differential system subject to stochastic noise, say with the solution denoted as $X = (X(t))_{t\in[0,T]}$, are concentration estimates of the form

$$\mathbb{P}\left\{ \sup_{t\in[0,T]} |X(t) - \mathbb{E}[X(t)]| > h \right\} \leq C(h,T), \tag{1.0.1}$$

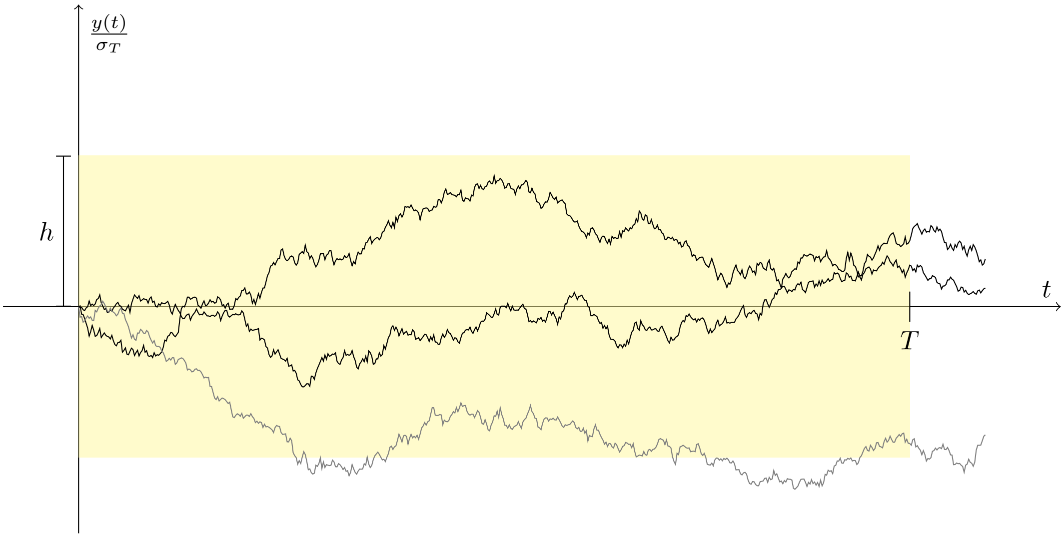

Figure 1: Suppose that $y(t) = X(t) - \mathbb{E}[X(t)]$, $t \in [0,T]$, is the deviation of some stochastic process from its deterministic counterpart and let us introduce the short notations for the variance $\sigma_T^2 := \operatorname{var} y(T)$, and $A := \{\sup_{s\in[0,T]} |y(s)| > h\sigma_T\}$ for the event that a deviation path leaves the interval $[-h\sigma_T, h\sigma_T]$ before time horizon $T > 0$ for some $h > 0$. The figure shows several paths of the stochastically perturbed deviation process y. The two black lines correspond to paths that satisfy A, while the gray line does not.

that provide upper bounds on the probability of an escape from an environment of the expectation process within a finite time horizon $T > 0$. Here $C(h,T)$ is some expression that depends on h and T. In the following we will refer to estimates of that, or closely related form as *concentration inequalities*. Figure 1 serves as an illustration. Typically, h is formulated as a multiple of the standard deviation of X. Concentration inequalities have been well-known for a long time; for instance concerning partial-sum processes in form of the *Dvoretzky–Kiefer–Wolfowitz inequality*, when increments are given by independent, identically distributed and bounded random variables, [DKW56]. And in the continuous-time case *Doob's* celebrated *maximal inequality*, [Kle14] has been available, when studied processes are submartingales. Due to the robustness of the Gaussian property, stochastic integrals, in case of an integrand that solely depends on time, are centered Gaussian processes, [Bau96]. And for such a process, say $(X(s))_{s\in[0,T]}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the Borel-TIS inequality, [AT07, Theorem 2.1.1], yields that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |X(s)| > \mathbb{E}\|X\|_{[0,T]} + h\right\} \leq \exp\left(-\frac{h^2}{2\mathbb{E}\|X\|_{[0,T]}^2}\right) \qquad \text{for } h > 0,$$

where $\|\cdot\|_{[0,T]}$ denotes the supremum norm over $[0,T]$. The Borel-TIS inequality is certainly one of the most valuable inequalities in the context of Gaussian processes. Its preciousness arises on the one hand from rather general validity, on the other hand from its simple, elegant structure. Let us mention one more type of concentration inequality, established by X. Fernique in 1964, which is applicable for a rather general class of Gaussian processes, and which is not explicitly given here due to a bit of notational bulkiness that is involved, but we present the original Fernique inequality in detail in Section 2.1.

1.1. Placement in the Literature and Aim

The predominant goal of this thesis is to establish a description of pathwise concentration results for stochastic delay differential equations (SDDEs) including the nonautonous case, and, at least in special cases, the more general stochastic functional differential equations (SFDEs) with additive noise. The book [BG06] by N. Berglund and B. Gentz serves as a paragon for our study. In particular, we aim for precisely-as-possible confined areas that solution paths do not leave with high probability, formulated in terms of concentration inequalities in the form (1.0.1) with $C(h,T) = C^{(1)}\exp(-C^{(2)}h^2)$. There are three points of particular importance, that delimit this work from the established results, that the literature provides so far:

- Paths stay in determined areas over finite time intervals with high probability, not asymptotically.
- Pathwise properties hold for specified sizes of respective parameters rather than solely in the small-noise limit.
- Special emphasis lies on estimates on the constants $C^{(1)}$ and $C^{(2)}$ regarding their dependence on the underlying set of parameters to track the role that the delay term as well as other involved quantities play.

Striving for pathwise properties of processes, distributional properties, such as we may obtain from the Fokker-Planck approach or generally concerning stationary distributions, do not suffice, because, even in case that the distribution density can be satisfactorily obtained, it only provides the one-dimensional distributions, [Lon10]. And in general, there is no way to gain insight on the level of paths from that. Regarding large deviations, the first one to study SFDEs in the white noise case was M. Scheutzow, [Sch84]. Further results have been contributed e.g. by R. Langevin, W. M. Oliva and J. C. F. de Oliveira in [LODO91]. An extension to more general diffusion terms has been achieved by S.-E. A. Mohammed and T. Zhang in 2006, see [MZ06]. Furthermore, Lévy noise was considered by K. Liu and T. Zhang in [LZ14] for the retarded type, and by J. Bao and C. Yuan in [BY15]. One part in the derivation of a large-deviation result is typically based on concentration inequalities. For instance, we follow the presentation in [LODO91], where $X^{(\varepsilon)}$ solves $\dot{X}^{(\varepsilon)}(t) = b(X^{(\varepsilon)}(t)) + \varepsilon\dot{W}(t)$, and x solves $\dot{x} = b(x(t))$. Then, the authors show that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \|X^{(\varepsilon)}(t) - x(t)\| > \delta\right\} \le C_1 \exp\left(-\frac{C_2}{\varepsilon^2}\right),$$

which serves to reason that

$$\lim_{\varepsilon\to 0} \mathbb{P}\left\{\sup_{t\in[-1,T]} |X^{(\varepsilon)}(t) - x(t)| > \delta\right\} = 0.$$

But unfortunately, apart from the missing relation between δ and ε, the concentration inequality bears the unknown constant C_1 and prefactor C_2 in the exponent, which is why is does not suit our needs. Moreover, there are excellent results available on the asymptotic maxima of Gaussian processes, in particular in the stationary case, e.g. [Pic67] [Mar72], see also [AMW10]. For instance, in fairly general situations, we know that there is a process ρ

such that

$$\mathbb{P}\left\{\limsup_{t\to\infty}\frac{X(t)}{\sqrt{\rho(t)}}=1\right\}=1,$$

in which case $\rho(\cdot)$ is called the *essential growth rate*, or *running maximum*, which is explicitly known in many cases. Such essential-growth rate results do not provide any insight for finite time horizons, but they will serve as orientation, even when only formulated with "$\leq$" inside the braces. Concerning growth rates for SFDEs, recent studies have been performed in [Sch05], [AGR06], [Sch13], [AGR11], [AP15], [AP17], see also [HP14]. In the context of stochastic processes, a whole zoo of notions of stability is well-established in the literature. Among them let us mention the concept of almost-sure exponential stability. For example, the work [Mao07] of X. Mao provides an introduction and overview. A process $(X(t))_{t\geq t_0}$ is said to be almost-surely exponentially stable, if

$$\lim_{t\to\infty}\frac{1}{t}\log|X(t)|<0 \quad \mathbb{P}\text{-almost surely.}$$

If it exists, the left-hand side is called the *Lyapunov exponent.* Results for SFDEs are due to [MS90], [MS96], [MS97], [Els99], [Sch05], [Sch13]. In the same spirit as the essential-growth rate results, the concept provides a picture of the long-term behavior of a process. This picture is a rather crude one in the sense that constants and subexponential correction terms are lost in the statement. Regarding the mentioned concentration inequalities, solutions of SFDEs lack the martingale property, which in turn implies that the Bernstein-type inequality for stochastic integrals, a former valuable tool, is not straightly applicable here, [BG06]. The Borel-TIS inequality requires knowledge of the first two moments of the running supremum of the process, which are not easily available. Therefore, all but the Fernique inequality of the mentioned representatives of concentration inequalities will only be of limited use, and some of no use at all. In fact, we will build our analysis on a combination of the Bernstein-type and the Fernique inequality.

More than that, the broad field of SFDEs generally has remained under constant intense scientific study for several decades. This includes the classical rather abstract research areas like the question for existence and uniqueness For example [vRS10] considers fairly general conditions on the coefficient functions, [WYM17] treats a setting with infinite-delay, [BM16] and [ZAL$^+$17] provide results for the fractional-derivative formulation. The numerics of stochastic functional differential equations have for example been studied in [BB00], [Mao03], [HMY04], [Buc04], [Buc06], [Mao07], [BKMS08], [FN09], [AB10], [KS12], [Kim16]. Deterministic systems have experienced a tremendous amount of scientific research with regard to stability issues; [vC15] provides an overview. Regarding the large field natural-scientific research, time-delayed differential laws have found a variety of applications like for instance optical devices (e.g. [HKGS82]), chemical dynamics (e.g. [Rou96]), traffic flow models (e.g. [SN10], [Hel01]), mechanical engineering (e.g. [DK92]), neural networks [BMS01], or finance (e.g. [KP07], [AI05], [AIK05], [AHMP07], [ARS13], [Zhe15], [TKBM15]). For an introduction and survey see [Ern09], [Lak11], for difference equations [IKS03]. A beautiful introduction to the applications is provided by *T. Erneux* in [Ern09].

1.2. Structure and Progress

The actual content of this work starts with Chapter 2 with a review of a concentration inequality for a rather general class of Gaussian processes. The main result is due to X. Fernique and was originally established in 1964, [Fer64], [Fer90], which is why we refer to it as the *Fernique inequality*. We will basically repeat the arguments except for negligible modifications. Save for technical assumptions, the requirements of the Fernique inequality consist of simply one integrability condition on the covariance structure. It is not restricted to stationarity or autonomy. Many of the concentration inequalities in this work including solutions of autonomous linear functional differential equations with additive noise, constant-coefficient SDDEs and linearizations of a special kind of nonautonomous nonlinear functional differential equations subject to noise, will be based on the Fernique inequality.

The subsequent Chapter 3 provides a short introduction to stochastic retarded functional differential equations (SRFDEs), mainly consisting of existence and uniqueness results and solution representations. The purpose is a review of fundamental-solution concept and the variation-of-constants formula, which is the reason why most of the details are basically taken from the literature. The very core, and the benefit of this Chapter, is the variation-of-constants formula for nonautonomous linear stochastic retarded functional differential equations. This one plays a crucial role, especially in Chapter 5, when we consider a retarded differential equations in a scenario where stability is slowly vanishing. As a first application of the Fernique inequality, a concentration inequality for autonomous retarded functional differential equations will be established. Due to recent work [AMW10] of Appleby, Mao and Wu the essential growth rate is explicitly known here in the stable regime. Their result and the concentration result, that we will achieve, suit each other. The generality comes with a price, there are constants involved on which we know almost nothing.

In Chapter 4 the generality is tremendously weakened in order to provide transparency on the respective impact of the underlying parameters involved in the formulation of the concentration inequality. We will consider stochastic delay differential equations (SDDE), which means systems of the form

$$\begin{cases} dx(t) = -ax(t)dt + bx(t-r)dt + \sigma dW(t) & \text{for } t \geq 0, \\ x(t) = \Upsilon(t) & \text{for } t \in [-r, 0], \end{cases}$$

where $\Upsilon \in \mathcal{C}([-r,0],\mathbb{R})$ and $a \in \mathbb{R}$, $b, \sigma > 0$, and W denotes a standard Brownian motion. As a central result we will show that for every $a = b > 0$ and every $\Upsilon \in \mathcal{C}([-r,0],\mathbb{R})$ the corresponding solution converges to a non-trivial limit in the deterministic case, i.e. $\sigma = 0$. We will provide the exact limit as well as a lower bound for the rate of convergence. This provides concrete knowledge adding to the presentation in [ARS13], [DvGVLW95], who were able to acquire the asymptotic limit for a general class of time-delayed feedback and in SDDE case at least for certain parameter combinations a and b. We will provide a self-contained presentation of the convergence result as well as a lower bound for the rate of convergence for the fundamental solutions. Knowledge on the convergence rate is crucial in the computation of concentration inequalities. In particular, we will show that

$$\mathbb{P}\left\{\sup_{s\in[0,T]}|x(s)-\mathbb{E}[x(s)]|>h\frac{\sqrt{T}}{1+ar}\right\}\leq\frac{5}{2}T^2\exp\left(-\frac{h^2}{2\sigma^2}(1+\mathcal{O}(T^{-\frac{1}{2}}))\right)\text{ for big }T,$$

under irrestrictive conditions on $h>0$. This result is due to an application of the Fernique inequality. We will show how the convergence result can be taken over to arbitrary parameter combinations with $b>0$, and provide concentration inequalities in a variety of regimes. A careful study of small-ball probabilities further reveals that the first-exit-time distribution of $x(t)-\mathbb{E}[x(t)]$, $t\geq 0$, is to some extend very similar to the one of the rescaled Brownian motion $\frac{W(t)}{1+ar}$, $t\geq 0$.

In Chapter 5 we will consider a particular nonautonmous system that features delay-feedback and nonlinearity. We consider systems of the form

$$dx(t)=f(x(t),\nu t)dt+g(x(t-r),\nu t)dt+\sigma dW(t),\quad t\geq 0. \tag{1.2.1}$$

where f and g are potential gradients that slowly change with time due to the small parameter ν. This formulation of an SRFDE, consisting of two possibly different potentials acting on the current value and on the delayed term, has been inspired by the work [FI05] of P. Imkeller and M. Fischer, who study the effective dynamics of a bistable system featuring stochastic resonance. There, $f(t,\cdot)=f(\cdot)=V'(\cdot)$ where V is a symmetric one-dimensional double-well potential, and $g(t,\cdot)=g(\cdot)=U'(\cdot)$, where U is a quadratic potential. Due to an analysis of residence times in a two-state model, and corresponding limiting distributions, they establish an instance of stochastic resonance.

The analysis, that we present, includes concentration results in rather general situations in a uniformly stable environment. Those are actually applicable for the model in [FI05]; there we can provide a lower bound on residence times that hold with high probability. The actual transition, i.e. an upper bound on residence times, is not included.

The procedure, which means the way the system changes with time, is inspired by [BG06, Chapter 3] where no delayed feedback is involved. We will present several methods to achieve concentration inequalities, one of them again inspired by the just mentioned work. Part of the description crucially relies on the nonautonomous variation-of-constants formula that is derived in Chapter 3 of this work. Without that particular variation-of-constants formula, a pathwise description of the transition from stability to instability is hardly thinkable. A significant role is taken by an appropriately chosen reference system that substitutes the lack of a conveniently defined equilibrium-branch concept. Furthermore, the transition to instability will either occur through a certain type of symmetric pitchfork bifurcation, or the system will be assumed to be linear. Denoting the time-speed parameter by $\nu>0$, then under the assumption that $\sigma<\nu/|\log\nu|$, the predominant achievements regarding nonautonomous systems (1.2.1) are the following.

- Uniformly stable branches attract solution paths into a neighborhood of order ν, when these have been initiated at a distance of order 1 within a time of order $|\log\nu|/\sqrt{\nu}$.
- A solution path, that is initiated close to a uniformly stable branch, remains in a neighborhood of order ν for an exponential amount of time.
- With regard to residence-time results with respect to neighborhoods around destabi-

lizing branches from [BG06, Chapter 3], we manage to carry over established results to the delayed-feedback case. Compare [BG06, Figure 3.12, Section 3.4] and Figure 8, and 12 which applies under stricter conditions.

- Ignoring nonlinear terms, we will show that solutions paths typically leave unstable branches in a time that is comparable to the delay-free case.

The above statements hold with high probability, formulated in terms of concentration inequalities. To the best of our knowledge there is no attempt to the pathwise analysis of SDDEs in terms of concentration inequalities anywhere in the literature, not even in the simplest constant scalar case, and results are generally scarce for nonautonomous systems. Concerning (stochastic) delay differential equations provides plenty of details regarding bifurcation diagramms, e.g. [YB11], [BC94], [CYB05], [GFF17] but the author has not seen any evidence of an approach of the kind that will be presented in this work. We will constantly work out explicit-as-possible conditions on the size of ν, that are necessary for our results to hold. And that is also the reason why we focus on basically simple settings and tend to avoid building on asymptotic spectral-theoretic results.

2. The Fernique Inequality: A Concentration Inequality for Gaussian Processes

This section is devoted to a review of a famous result on concentration inequalities of continuous real-valued Gaussian process $(X(s))_{s\in\mathcal{T}}$ over a multidimensional time-index set $\mathcal{T} = [a,b]^n$ where $a < b \in \mathbb{R}$ and $n \in \mathbb{N}$. The result has originally been established by X. Fernique in 1964, see [Fer90] or [Fer64]. The proof is comparably straightforward using rather basic estimates concerning normally distributed random variables. The profit, and to some extend the real power of the Fernique inequality, is its robustness to apply in quite general situations. We will formulate and prove the concentration inequality originally stated as *Théorème 4.1.1* from the above-mentioned reference. In Corollary 2.4 we will present an upper bound on the essential growth rate, that was established by M. B. Marcus [Mar70] based on a variante of the Fernique inequality. We will provide an own proof based on the version that we present below.

2.1. The Fernique Inequality

Let $X = (X(s))_{s\in\mathcal{T}}$ be some centered continuous $\mathbb{R}$-valued Gaussian process over some time-index set $\mathcal{T} \subset \mathbb{R}^n$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with covariance structure $\Gamma(s,t) = \mathbb{E}[X(t)X(s)]$ for $s,t \in \mathcal{T}$ and start in $X(0) = 0$. Dealing with finite-dimensional objects, we denote the maximum norm by $\|\cdot\|_{\max}$, e.g.

$$\|t\|_{\max} := \max_{i\in\{1,\ldots,n\}} |t_i| \quad \text{for all } t = (t_1,\ldots,t_n) \in \mathcal{T},$$

only if we want to emphasize the finite dimensionality. Otherwise, and if ambiguity can be excluded, we simply write $\|\cdot\|$ for the sup-norm as well as for the max-norm. We define

$$\begin{aligned}\varphi(h) &= \sup_{\substack{s,t\in\mathcal{T}\\ \|t-s\|_{\max}\le h}} \sqrt{\Gamma(s,s) - 2\Gamma(s,t) + \Gamma(t,t)}\\ &= \sup_{\substack{s,t\in\mathcal{T}\\ \|t-s\|_{\max}\le h}} \sqrt{\mathbb{E}\Big[\big(X(t)-X(s)\big)^2\Big]} \quad \text{for all } h > 0.\end{aligned}$$

In particular, by the Cauchy–Schwarz inequality, we have that

$$\sup_{(s,t)\in\mathcal{T}\times\mathcal{T}} \Gamma(s,t) \le \sup_{s\in\mathcal{T}} \sqrt{\Gamma(s,s)}\, \sup_{t\in\mathcal{T}} \sqrt{\Gamma(t,t)} = \sup_{s\in\mathcal{T}} \Gamma(s,s) \le \sup_{s\in\mathcal{T}} \sup_{t\in\mathcal{T}} \Gamma(s,t), \tag{2.1.1}$$

and so there must be equality in every step in (2.1.1), i. e.

$$\sup_{(s,t)\in\mathcal{T}\times\mathcal{T}} \Gamma(s,t) = \sup_{s\in\mathcal{T}} \Gamma(s,s) = \|\Gamma\|. \tag{2.1.2}$$

As a matter of fact, in 1964 Fernique formulated the following concentration inequality together with a sufficient condition on Gaussian processes to be continuous. The continuity part of the theorem aroused much more attention in the literature than the actual concentration inequality, and while solutions of SRFDEs are required to be continuous anyway, the converse is true in this work. The proof is almost the same save a tiny alteration for clearity sake that comes up as an additional factor of 2 within the formulation of φ.

Theorem 2.1 (The Fernique inequality, Théorème 4.1.1 in [Fer90]). *Let $\mathcal{T} = [0,1]^n$, $n \in \mathbb{N}$, and let $X = (X(s))_{s\in\mathcal{T}}$, a separable, real-valued, centered Gaussian process with covariance structure $\Gamma(s,t) = \mathbb{E}[X(s)X(t)]$ for $s,t \in \mathcal{T}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Further, we assume $\int_0^\infty \varphi\left(\exp\left(-u^2\right)\right) du$ to be finite. Then:*

a) The process has continuous paths almost surely.

b) For all $p \in \mathbb{N}$, $p \geq 2$ and real $h \geq \sqrt{1+4n\log p}$, we have

$$\mathbb{P}\left\{\sup_{s\in\mathcal{T}} |X(s)| \geq h\left(\sqrt{\|\Gamma\|} + \left(2+\sqrt{2}\right)\int_1^\infty \varphi\left(p^{-u^2}\right) du\right)\right\} \leq \frac{5}{2}p^{2n}\int_h^\infty \exp\left(-\frac{u^2}{2}\right) du.$$

Proof. a) This part is of no particular interest for us, it is stated for completeness sake.

b) Let $m \in \mathbb{N}\backslash\{0\}$ be arbitrarily given. By I_m we denote the collection of multi-indices $I_m := \{0,\dots,m-1\}^n$. Further, we let

$$t_i^{(m)} := \frac{1}{m}\left(i_1,\dots,i_n\right) \quad \text{for all } i = (i_1,\dots,i_n) \in I_m$$

denote what we may think of as *lattice points* of $\mathcal{T}$ due to fineness m^{-1}. The collection of those lattice points for fineness m^{-1} is denoted by $\mathcal{T}^{(m)}$, i. e., with a slight abuse of notations,

$$\mathcal{T}^{(m)} := \left\{t_i^{(m)} : i \in I_m\right\} = \frac{1}{m}I_m.$$

And we denote by

$$B_i^{(m)} := \left\{t \in [0,1)^n \ : i_j \leq mt_j < i_j + 1 \text{ for all } j \in \{1,\dots,n\}\right\} = \bigtimes_{j=1}^n \left[\frac{i_j}{m}, \frac{i_j+1}{m}\right) \quad \text{for all } i \in I_m,$$

those boxes in the time-index set $\mathcal{T}$ that are canonically associated with the lattice set $\mathcal{T}^{(m)}$. The partition $B^{(m)} := \{B_i^{(m)} : i \in I_m\}$ serves as container for all those boxes. Figure 2 serves as an illustration.

Remember that for two random variables ξ_1, ξ_2 that are normally distributed with respect to $\mathbb{P}$ with mean 0 and standard deviations $\sqrt{\operatorname{var}\xi_1} < \sqrt{\operatorname{var}\xi_2}$, we have that

$$\mathbb{P}\left\{|\xi_1| > h\right\} \leq \mathbb{P}\left\{|\xi_2| > h\right\} \quad \text{for all } h \geq 0. \tag{2.1.3}$$

Note that, by the simple fact that

$$\text{for all } m \in \mathbb{N}\backslash\{0\} \text{ and } i \in I_m \text{ there is unique } \hat{t} \in \mathcal{T}^{(m)} \text{ such that } \hat{t} \in B_i^{(m)},$$

which is of course given by $\hat{t} = t_i^{(m)}$, there is a one-to-one corresponence between boxes and lattice points. In order to define an appropriate sequence of approximations $\left(X^{(m)}(\cdot)\right)_{m\in\mathbb{N}}$ of X, we observe the values X takes at the m^n lattice points of $\mathcal{T}^{(m)}$ and endow $X^{(m)}$ in every point in a given tile $B_i^{(m)}$ with the value $X\left(t_i^{(m)}\right)$, where $i \in I_m$. For an illustration, see Figure 3. Formally, for all $m \in \mathbb{N}$, we define $X^{(m)}$ for all $t \in \mathcal{T}$ by

$$X^{(m)}(t) := X(t_i^{(m)}) \quad \text{if } t \in B_i^{(m)} \text{ for } i \in I_m. \tag{2.1.4}$$

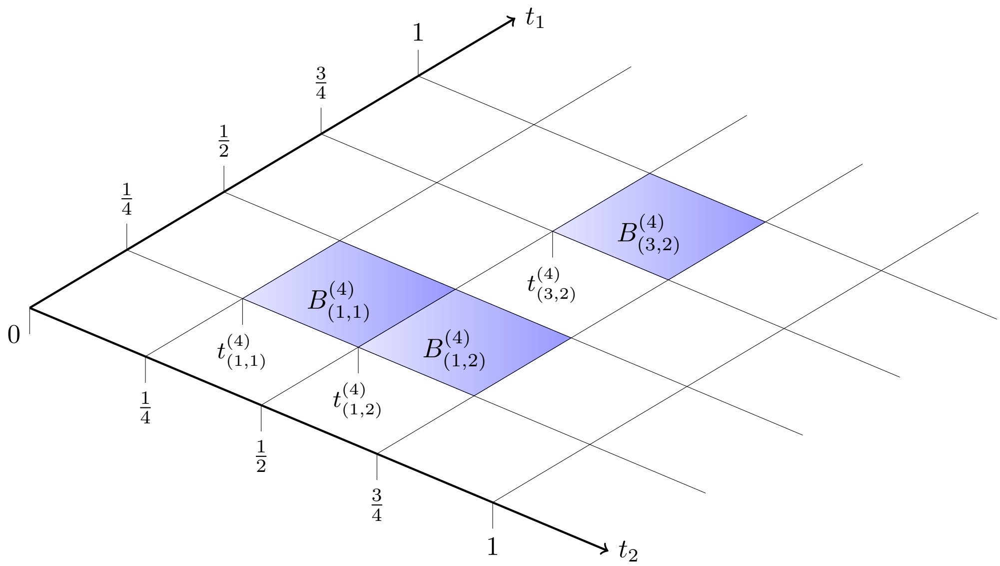

Figure 2: Illustration of the partition of $\mathcal{T} = [0,1]^2$ proposed in [Fer90] for $m = 4$. Line crossings refer to the elements of $\mathcal{T}^{(4)}$ and tiles correspond to elements of $B^{(4)}$.

Well-definedness is then due to the one-to-one correspondence of lattice points and boxes.

Of course, when studying $\|X^{(m)}\|$, it suffices to restrict the attention to the lattice points $\mathcal{T}^{(m)}$. Formally,

$$\begin{aligned}\|X^{(m)}\| = \sup_{t\in\mathcal{T}} |X^{(m)}(t)| &= \sup_{i\in I_m} \sup_{t\in B_i^{(m)}} |X(t)| \\ &= \sup_{i\in I_m} |X^{(m)}(t_i^{(m)})| = \sup_{i\in I_m} |X(t_i^{(m)})| = \sup_{t\in\mathcal{T}^{(m)}} |X(t)| \qquad (2.1.5)\end{aligned}$$

is the maximum of m^n absolute values of (possibly correlated) Gaussian random variables. In other words, the probability $\mathbb{P}\left\{\|X^{(m)}\| \geq h\sqrt{\|\Gamma\|}\right\}$ depends on an m^n-dimensional marginal distribution of X which is, of course, a Gaussian distribution, but still it is not too handy. The following provides a way to deduce an estimate that actually only relies upon the one-dimensional marginal distributions:

$$\begin{aligned}\mathbb{P}\left\{\|X^{(m)}\| \geq h\sqrt{\|\Gamma\|}\right\} &= \mathbb{P}\left\{\sup_{t\in\mathcal{T}^{(m)}} |X(t)| \geq h\sqrt{\|\Gamma\|}\right\} \qquad (2.1.6)\\ &= \mathbb{P}\left(\bigcup_{t\in\mathcal{T}^{(m)}} \left\{|X(t)| \geq h\sqrt{\|\Gamma\|}\right\}\right) \\ &\leq \sum_{t\in\mathcal{T}^{(m)}} \mathbb{P}\left\{|X(t)| \geq h\sqrt{\|\Gamma\|}\right\} \\ &\leq m^n \sup_{t\in\mathcal{T}^{(m)}} \mathbb{P}\left\{|X(t)| \geq h\sqrt{\|\Gamma\|}\right\} \quad \text{for all } h > 0. \qquad (2.1.7)\end{aligned}$$

As X is Gaussian, for arbitrary $t \in \mathcal{T}$ the random variable $X(t)$ is normally distributed with mean 0 and its standard deviation is dominated by $\sqrt{\|\Gamma\|}$. Then $\frac{X(t)}{\sqrt{\|\Gamma\|}}$ has mean 0 and its standard deviation is dominated by 1. Let $\mathcal{X}$ be a normally distributed random variable (with respect to $\mathbb{P}$) with mean 0 and standard deviation $\sigma_{\mathcal{X}} = 1$. Then by (2.1.3), we may

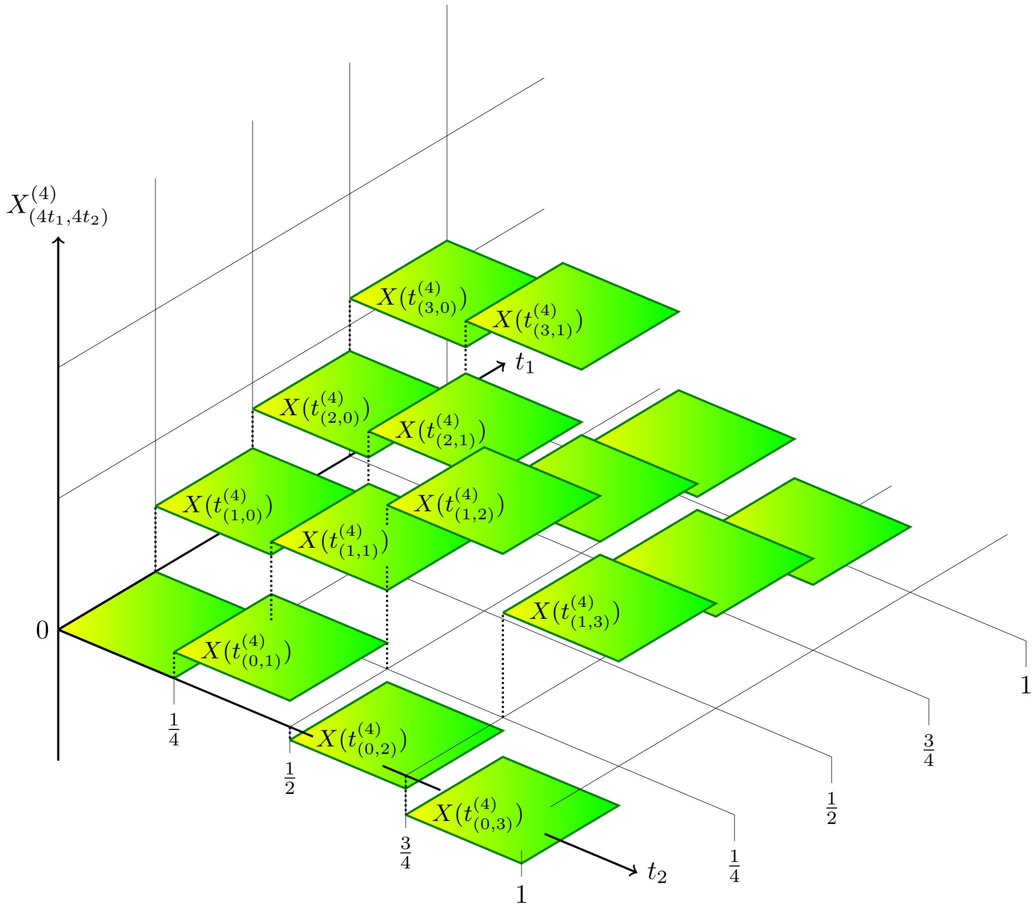

Figure 3: An illustration of the approximation $X^{(4)}$ taking the value $X(t_i^{(4)})$ on every $B_i^{(4)}$ for $i \in I_4 = \{0,1,2,3\}^2$. The original process X is not included in the figure. One can imagine the process X as wavering plain that coincides with the floating tiles in that very point of a tile that is closest to the origin.

deduce

$$\mathbb{P}\left\{|X(t)| \geq h\sqrt{\|\Gamma\|}\right\} \leq \mathbb{P}\left\{|\mathcal{X}| \geq h\right\} = \frac{2}{\sqrt{2\pi}} \int_h^\infty \exp\left(-\frac{u^2}{2}\right)\, du \quad \text{for all } t \in \mathcal{T}, h > 0. \tag{2.1.8}$$

Then (2.1.7) and (2.1.8) yield

$$\mathbb{P}\left\{\|X^{(m)}\| \geq h\sqrt{\|\Gamma\|}\right\} \leq m^n \sqrt{\frac{2}{\pi}} \int_h^\infty \exp\left(-\frac{u^2}{2}\right)\, du \quad \text{for all } h > 0. \tag{2.1.9}$$

The above inequality constitutes an upper-bound estimate for the probability that $X(\cdot)/\|\Gamma\|$ exceeds h, when only observed at the m^n lattice points of $\mathcal{T}^{(m)}$, where $m \in \mathbb{N}$ is arbitrary. In the next step, we work out how the probability on the left-hand side of (2.1.9) evolves when we put more and more points into observation. To this end, let the sequence $(m_i)_{i\in\mathbb{N}}$ be successively divisible, i. e. $m_{i+1}/m_i \in \{2,3,4,\ldots\}$ for all $i \in \mathbb{N}$. Then for $k < l$, we have that $I_{m_k} \subset I_{m_l}$ and the partition $B^{(m_l)}$ is a refinement for $B^{(m_k)}$. Furthermore, for all $k \in \mathbb{N}$, the random variable $X^{(m_{k+1})} - X^{(m_k)}$ is Gaussian again, because it is the image of a linear mapping of Gaussian variables. It is centered and its supremum is determined over the lattice points $\mathcal{T}^{(m_{k+1})}$ in the sense that

$$\|X^{(m_{k+1})} - X^{(m_k)}\| = \sup_{t\in\mathcal{T}} |X^{(m_{k+1})}(t) - X^{(m_k)}(t)| = \sup_{t\in\mathcal{T}^{(m_{k+1})}} |X^{(m_{k+1})}(t) - X^{(m_k)}(t)| \tag{2.1.10}$$

is the maximum of m_{k+1}^n normally distributed random variables. Consider an arbitrary fixed $t_0 \in \mathcal{T}$. By the correspondence between lattice points and tiles there must be unique $i_0 \in I_{m_k}$ such that $t_0 \in B_{i_0}^{(m_k)}$, and so $|t_0 - t_{i_0}^{(m_k)}| \leq m_k^{-1}$ holds true for all $k \in \mathbb{N}$. Therefore, the variance of $X^{(m_{k+1})}(t_0) - X^{(m_k)}(t_0)$ is dominated by

$$\begin{aligned}\mathbb{E}\left[\left(X^{(m_{k+1})}(t_0) - X^{(m_k)}(t_0)\right)^2\right] &= \mathbb{E}\left[\left(X^{(m_{k+1})}(t_0) - X^{(m_k)}(t_{i_0}^{(m_k)})\right)^2\right] \\ &= \mathbb{E}\left[\left(X^{(m_{k+1})}(t_0) - X(t_{i_0}^{(m_k)})\right)^2\right] \\ &\leq \sup_{s,t:\|s-t\|\leq 1/m_k} \mathbb{E}\left[\left(X(s) - X(t)\right)^2\right].\end{aligned}$$

And therefore,

$$\mathbb{E}\left[\left(X^{(m_{k+1})}(t_0) - X^{(m_k)}(t_0)\right)^2\right] \leq \varphi^2\left(\frac{1}{m_k}\right) \quad \text{for all } t_0 \in B_{i_0}^{(m_k)}. \tag{2.1.11}$$

Then, applying the same ideas as between (2.1.6) and (2.1.7), together with (2.1.10), we may deduce that

$$\begin{aligned}&\mathbb{P}\left\{\|X^{(m_{k+1})} - X^{(m_k)}\| \geq h\varphi\left(\frac{1}{m_k}\right)\right\} \\ &\leq \mathbb{P}\left(\bigcup_{i\in I_{m_k}}\left\{\sup_{s\in B_i^{(m_k)}} \left|X^{(m_{k+1})}(s) - X^{(m_k)}(s)\right| \geq h\varphi\left(\frac{1}{m_k}\right)\right\}\right) \\ &\leq \sum_{i\in I_{m_k}} \mathbb{P}\left\{\sup_{t\in\mathcal{T}^{(m_{k+1})}\cap B_i^{(m_k)}} \left|X(t) - X(t_i^{(m_k)})\right| \geq h\varphi\left(\frac{1}{m_k}\right)\right\} \\ &\leq \sum_{i\in I_{m_k}} \sum_{t\in\mathcal{T}^{(m_{k+1})}\cap B_i^{(m_k)}} \mathbb{P}\left\{\left|X(t) - X(t_i^{(m_k)})\right| \geq h\varphi\left(\frac{1}{m_k}\right)\right\} \\ &\leq m_{k+1}^n\sqrt{\frac{2}{\pi}}\int_h^\infty \exp\left(-\frac{u^2}{2}\right)du \quad \text{for all } h > 0,\end{aligned} \tag{2.1.12}$$

where in the last step we have used (2.1.9) and the fact that $\mathcal{T}^{(m_{k+1})}$ contains $\left(\frac{m_{k+1}}{m_k}\right)^n$ times so many lattice points over $B_i^{(m_k)}$ as $\mathcal{T}^{m_k}$ for all $i \in I_{m_k}$. Combining (2.1.9) and (2.1.12), and using the fact that the probability, that a sum overcomes a given threshold, is dominated by the probability that, informally, at least one addend overcomes its share of the threshold, leads to the following estimate,

$$\begin{aligned}
&\mathbb{P}\left\{\|X^{(m_1)}\| + \sum_{k=1}^{\infty}\|X^{(m_{k+1})} - X^{(m_k)}\| \geq h_0\sqrt{\|\Gamma\|} + \sum_{k=1}^{\infty} h_k\varphi\left(\frac{1}{m_k}\right)\right\}\\
&\quad\leq \mathbb{P}\left(\{\|X^{(m_1)}\| \geq h_0\sqrt{\|\Gamma\|}\} \cup \bigcup_{k=1}^{\infty}\left\{\|X^{(m_{k+1})} - X^{(m_k)}\| \geq h_k\varphi\left(\frac{1}{m_k}\right)\right\}\right)\\
&\quad\leq \sqrt{\frac{2}{\pi}}\sum_{k=0}^{\infty} m_{k+1}^n \int_{h_k}^{\infty}\exp\left(-\frac{u^2}{2}\right)du \qquad \text{for all } h_k > 0 \text{ for all } k \in \mathbb{N}\cup\{0\}.
\end{aligned}$$

We let $\tilde{\mathcal{T}} := \cup_{k\in\mathbb{N}}\mathcal{T}^{(m_k)}$ which is a countable dense subset of $[0,1]^n$. Therefore, $\|X\|$ has the same law as $\sup_{s\in\tilde{\mathcal{T}}}|X(s)|$ by continuity. And as $X^{(0)} = X(0) = 0$, that one is dominated by

$$\sup_{s\in\tilde{\mathcal{T}}}|X(s)| \leq \|X^{(m_1)}\| + \sum_{k=1}^{\infty}\|X^{(m_{k+1})} - X^{(m_k)}\| = \sum_{k=0}^{\infty}\|X^{(m_{k+1})} - X^{(m_k)}\|.$$

We deduce, formally by applying monotone convergence on both sides, that

$$\mathbb{P}\left\{\|X\| \geq h_0\sqrt{\|\Gamma\|} + \sum_{k=1}^{\infty} h_k\varphi\left(\frac{1}{m_k}\right)\right\} \leq \sqrt{\frac{2}{\pi}}\sum_{k=0}^{\infty} m_{k+1}^n\int_{h_k}^{\infty}\exp\left(-\frac{u^2}{2}\right)du \quad \text{for all } h_k > 0 \text{ for all } k \in \mathbb{N}\cup\{0\}. \tag{2.1.13}$$

The remainder is due to a neat choice of the sequences $(m_k)_{k\in\mathbb{N}}$ and $(h_k)_{k\in\mathbb{N}\cup\{0\}}$. For an arbitrary integer $p \in \mathbb{N}$, $p \geq 2$ and $h > 0$ we let

$$m_k = p^{2^k}, \quad x_k = 2^{\frac{k}{2}}, \quad h_0 = h, \quad h_k = 2^{\frac{k}{2}}h = x_k h \quad \text{for all } k \in \mathbb{N}. \tag{2.1.14}$$

Then for all $k \geq 1$, as φ is increasing and

$$\begin{aligned}
x_k - x_{k-1} = 2^{\frac{k}{2}} - 2^{\frac{k-1}{2}} = 2^{\frac{k}{2}}\left(1 - \frac{1}{\sqrt{2}}\right) = \frac{2^{\frac{k}{2}}}{2+\sqrt{2}} = \frac{x_k}{2+\sqrt{2}}\\
\Leftrightarrow \quad x_k = (x_k - x_{k-1})(2+\sqrt{2}),
\end{aligned} \tag{2.1.15}$$

for the series on the left-hand side in (2.1.13), we apply the definition of h_k to achieve

$$h_k\varphi\left(\frac{1}{m_k}\right) = h(2+\sqrt{2})(x_k - x_{k-1})\varphi\left(p^{-x_k^2}\right) \leq h(2+\sqrt{2})\int_{x_{k-1}}^{x_k}\varphi\left(p^{-u^2}\right)du. \tag{2.1.16}$$

Iterated applications of (2.1.16) directly lead to

$$\sum_{k=1}^{\infty} h_k\varphi\left(\frac{1}{m_k}\right) \leq h(2+\sqrt{2})\int_1^{\infty}\varphi\left(p^{-u^2}\right)du,$$

which is finite due to the assumption $\int_0^{\infty}\varphi\left(\exp\left(-x^2\right)\right)dx < \infty$. And for the series on the right-hand side of equation (2.1.13), using the substitution $u = v2^{\frac{k}{2}}$, we find that

$$\begin{aligned} m_{k+1}^n \int_{h_k}^\infty \exp\left(-\frac{u^2}{2}\right) du &= \left(p^{2^{k+1}}\right)^n \int_{h_k}^\infty \exp\left(-\frac{u^2}{2}\right) du \\ &= p^{n2^{k+1}} 2^{\frac{k}{2}} \int_h^\infty \exp\left(-\frac{v^2}{2}2^k\right) dv \\ &= \int_h^\infty \exp\left(n2^{k+1}\log(p) + \frac{k}{2}\log(2) - \frac{v^2}{2}2^k\right) dv \quad \text{for all } k \in \mathbb{N}. \end{aligned} \tag{2.1.17}$$

To find an upper bound for the exponent that appears in (2.1.17), we apply in particular that for all $v \geq h \geq \sqrt{1 + 4n\log(p)}$, we have that

$$\begin{aligned} n2^{k+1}\log(p) + \frac{k}{2}\log(2) - \frac{v^2}{2}2^k &= 2^k\left(2n\log(p) - \frac{v^2}{2}\right) + \frac{k}{2}\log 2 \\ &= \frac{1}{2}2^k\left(4n\log(p) - v^2\right) + \frac{k}{2}\log 2 \\ &= \frac{1}{2}\left(4n\log(p) - v^2\right) + \frac{1}{2}(2^k - 1)\big(4n\log(p) - v^2\big) + \frac{k}{2}\log 2 \\ &\leq -\frac{v^2}{2} + 2n\log(p) + \frac{1}{2}\left(k\log(2) + 1 - 2^k\right) \quad \text{for all } k \in \mathbb{N}, \end{aligned}$$

where in the last step, we have applied that $4n\log(p) - v^2 \leq -1$ due to the assumption above. And therefore, we find that

$$\sum_{k=0}^\infty m_{k+1}^n \int_{h_k}^\infty \exp\left(-\frac{u^2}{2}\right) du \leq p^{2n} \sum_{k=0}^\infty 2^{\frac{k}{2}} \exp\left(-\frac{2^k - 1}{2}\right) \int_h^\infty \exp\left(-\frac{u^2}{2}\right) du.$$

To derive the claimed estimate, it suffices to plug in this estimates into (2.1.13) and calculate

$$\sum_{k=0}^\infty 2^{\frac{k}{2}} \exp\left(-\frac{2^k - 1}{2}\right) \leq \frac{5}{2}, \tag{2.1.18}$$

which is done in [Del65] or [Fer75], and this part of the proof is complete. □

Remark 2.2.

- *It is not too hard to derive an even slightly better estimate in* (2.1.18) *using sharp-as-possible estimates of the first addends (those with significant contribution) and then find an upper-bound estimate with the help of an appropriate geometric series.*

- *The original reference brings up an interesting fact concerning the integrability assumption of φ. According to that the only functions φ (increasing, positive) for which there exists an appropriate nonnegative sequence $(h_k)_{k\in\mathbb{N}}$ and a sequence $(m_k)_{k\in\mathbb{N}}$ of (divisible) integers such that the two series converge, are those for which the integral $\int_0^\infty \varphi(\exp(-u^2))du$ converges, see [Fer90]. For that reason, the particular choice of $(h_k)_{k\in\mathbb{N}}$ and $(m_k)_{k\in\mathbb{N}\cup\{0\}}$ within the proof does not raise any not strictly necessary assumptions on φ.*

- *In [Mar70] one finds a modification of the presented inequality. It leads to a prefactor improvement that can be noteworthy in special cases. A discussion can be found in the stated reference.*

Corollary 2.3 (*Lemme* subsequent to *Théorème 4.1.1* in [Fer90]). *Consider* $0 \leq a < b$ *and let* X *be a separable Gaussian process on* $T = [a,b]^n$*, then with the notations from above, define for* $l > 0$,

$$Q(l) := \left(2+\sqrt{2}\right)\int_1^\infty \varphi\left(lp^{-u^2}\right) du. \tag{2.1.19}$$

Then

$$\mathbb{P}\left\{\sup_{s\in T}|X(s)| \geq h\left(\sqrt{\|\Gamma\|} + Q(b-a)\right)\right\} \leq \frac{5}{2}p^{2n}\int_h^\infty \exp\left(-\frac{u^2}{2}\right) du. \tag{2.1.20}$$

Proof. Simple; we only set $m_k = \frac{p^{2^k}}{b-a}$ instead of p^{2^k} in the proof of Theorem 2.1, one gets the desired result. Notice that none of the assumed properties is affected. □

Let us restrict to the situation of a single time dimension, $n = 1$, where $(X(t))_{t\in[0,\infty)}$ is a real-valued Gaussian process. If $\Gamma(\cdot)$ is bounded by some finite $\overline{\Gamma}$, the assumptions of the Fernique inequality imply a lower bound on the essential growth rate due to [Mar70]. The important properties in this situation are

$$\mathbb{E}\left[(X(t)-X(s))^2\right] \leq \varphi(|t-s|), \quad \mathbb{E}X^2(t) \leq \overline{\Gamma} \quad \text{and} \quad \int_1^\infty \varphi(e^{-u^2})du < \infty.$$

Corollary 2.4 (Upper bound of the essential growth rate, [Mar70]). *Suppose that* $\mathbb{E}[X^2(t)] < \overline{\Gamma}$ *for all* $t \in [0,\infty)$*, and that* $\int_1^\infty \varphi(e^{-u^2})du < \infty$*. Then*

$$\mathbb{P}\left\{\limsup_{t\to\infty}\frac{|X(t)|}{\sqrt{2\log t}} \leq \sqrt{\overline{\Gamma}}\right\} = 1. \tag{2.1.21}$$

Proof. Following the presentation in [Mar70], this can be seen by denoting $Y_k(t) = X(k+t)$ for all $t \in [0,1]$ and $k \in \mathbb{N}$. Observe that

$$\int_1^\infty \varphi\left(p^{-u^2}\right) du = \frac{1}{\sqrt{\log p}}\int_{\sqrt{\log p}}^\infty \varphi\left(e^{-u^2}\right) du \quad \text{for all } p \in \mathbb{N},\ p \geq 2.$$

Through the integrability condition that shows that for given $\varepsilon > 0$, there is p sufficiently large such that

$$\frac{2+\sqrt{2}}{\sqrt{\overline{\Gamma}}}\int_1^\infty \varphi\left(p^{-u^2}\right) du < \frac{\varepsilon}{2}, \quad \text{which implies} \quad \sqrt{\|\Gamma\|} + Q(1) \leq \left(1+\frac{\varepsilon}{2}\right)\sqrt{\overline{\Gamma}}. \tag{2.1.22}$$

Given such p an application of the Fernique inequality yields

$$\begin{aligned}\mathbb{P}\left\{\sup_{t\in[0,1]}|Y_k(t)| > \max\left\{\sqrt{2\log k}, \sqrt{1+4\log p}\right\}(1+\varepsilon)\sqrt{\overline{\Gamma}}\right\} &\leq \frac{5p^2}{2}e^{-\frac{2\log k}{2}(1+\varepsilon)} \\ &= \frac{5p^2}{2}k^{-(1+\varepsilon)}.\end{aligned}$$

Then, the proof is completed by an application of the Borel–Cantelli Lemma, because for arbitrarily small $\varepsilon > 0$

$$\sum_{k=2}^{\infty} \mathbb{P}\left\{\sup_{t\in[0,1]} \frac{|X(k+t)|}{\sqrt{2\log t}} > (1+\varepsilon)\sqrt{\Gamma}\right\} \leq \sum_{k=2}^{\infty} \mathbb{P}\left\{\sup_{t\in[0,1]} \frac{|X(k+t)|}{\sqrt{2\log k}} > (1+\varepsilon)\sqrt{\Gamma}\right\} < \infty.$$

For the second series the Fernique inequality is not applicable for only finitely many initial addends due to the insufficient size of $\sqrt{2\log k}$. Every one of them is bounded by 1, so together they only have finite contribution to the series. □

Remark 2.5. *In the course of this work, boundedness of* $\Gamma(\cdot)$ *will occur as a phenomenon that comes with a proper notion of stability of the studied process* X*.*

Let us denote $\rho(t,s) = \mathbb{E}[X(t)X(s)]$ for all $s,t \in [0,\infty)$ and consider the special case where $\|\Gamma\| = v$ is a constant and additionally $\lim_{T\to\infty} \sup_{|t-s|>T} \rho(t,s) \leq 0$. In that case the work of M. Nisio [Nis67] provides that the almost-sure upper bound in (2.1.21) actually is the limit, i.e.

$$\mathbb{P}\left\{\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2\log t}} = \sqrt{\|\Gamma\|}\right\} = 1.$$

An alternative condition can be found in [Mar72].

Remark 2.6. *In the case of a stationary Gaussian process* $Z = (Z(t))_{t\in[0,\infty)}$ *there are much more results available concerning the asymptotic limit:*

- *Exact essential growth rates. Under reasonable assumptions, the result of Theorem (2.4) can be shown to hold as a convergence result, i.e. with* $\Gamma = \Gamma(\cdot)$,

$$\mathbb{P}\left\{\limsup_{t\to\infty} \frac{|Z(t)|}{\sqrt{2\log t}} = \sqrt{\Gamma}\right\} = 1.$$

  Sufficient conditions for the growth-rate result are e.g. provided by Pickands [Pic67] as a generalization of prior work of Simeon M. Berman [Ber64] in a discrete time setting

- *Extremal distributions. Given that the covariance function*

$$r(t) := \Gamma(s, s+t) = \mathbb{E}[Z(s)Z(s+t)] \qquad \textit{for all } s,t \in \mathbb{R} \textit{ (stationary case)}$$

  vanishes fast enough, i.e. either $\lim_{t\to\infty} r(t)\log t = 0$ *or* $\int_{\mathbb{R}} r^2(u)du < \infty$*, the exact asymptotic distribution of a properly rescaled version of* Z *can be received. See e.g. [Wat54], [Gum67], [Ber64], [Pic67], [Pic69].*

- *Concentration inequalities. Due to the work of Marcus and L. A. Shepp [MS72], for given finite time horizon* $T = 1$, $\|\Gamma = 1\|$*, and every* $\varepsilon > 0$*, there is* $\beta > 0$ *sufficiently large such that*

$$P\left\{\sup_{t\in[0,1]} |X(t)| > \beta\right\} \leq \exp\left(-\frac{\beta^2}{2(1+\varepsilon)}\right).$$

- *Distributions of high-level excursions. An involved treatment can be found in [Ber71a], [Ber71b], and for the case of stationary increments in [Ber72a], [Ber72b]. Exact results depend on intricate functions of the covariance.*

The essential growth rate in the formulation of Corollary (2.4) provides an elegant picture of the long-term behavior of such a process. It it worth emphasizing that Corollary 2.4 is deduced from neatly chosen concentration inequalities on finite time interval. But actually, from the essential growth rate, there is nothing left to be learned about a concentration in finite time. In this case the application of the Borel–Cantelli lemma has erased any information on finite time intervals.

3. On Stochastic Functional Differential Equations

The present part on general stochastic functional differential equations (SFDEs) is based upon the books of J. Hale and S. M. Verduyn Lunel [HVL93], and of X. Mao [Mao08]. We will present a brief review on existence and uniqueness of strong solutions based on the two references. In the second half we turn to the concept of fundamental matrix solutions and solution representations which is again based on the book of Hale and Lunel for the deterministic case. Further, we provide a generalization of the stochastic variation-of-constants formula in a nonautonomous setting. For the autonomous case the formula can be found in [Moh84] for instance.

3.1. Definitions and Conventions

We will always consider finite constant time delay $r > 0$ throughout this work. In order not to be overwhelmed by notations, we follow the established literature that commonly employs a handful of convenient short-hand notations. We will abbreviate $J := [-r, 0]$, and given any $n \in \mathbb{N}$ and any $\mathbb{R}^n$-valued process $(x(t))_{t\in[-r,T]}$, we will refer to its segment process by $(x_t)_{t\in[0,T]}$. That means that for arbitrary $t_0 \in [0, T]$ we write $x_{t_0} := (x(u) : u \in [t_0 - r, t_0])$ and reserve to write $x(t_0)$ if we consider the process's $\mathbb{R}^n$-valued evaluation at t_0. For any subsets $A \subset \mathbb{R}$, $B \subset \mathbb{R}^n$, we let $\mathcal{C}(A, B)$ denote the set of functions from A to B that are continuous with respect to the sup-norm $\|\cdot\|$. Then for $H : [t_0, \infty) \times \mathcal{C}(J, \mathbb{R}^n) \to \mathbb{R}^n$, $\Upsilon \in \mathcal{C}(J, \mathbb{R}^n)$ and $\sigma : [0, \infty) \times \mathcal{C}(J, \mathbb{R}^n) \to \mathbb{R}^{n\times m}$ and a given m-dimensional Brownian motion $(W(t))_{t\geq 0}$ on a filtered and completed probability space $(\Omega, \{\mathcal{F}_t\}_{t\geq t_0}, \mathcal{F}, \mathbb{P})$ it makes at least syntactically sense to consider the SFDE

$$\begin{cases} dx(t) = H(t, x_t)dt + \sigma(t, x_t)dW(t) & \text{for } t \geq t_0, \\ x_{t_0} = \Upsilon. \end{cases} \tag{3.1.1}$$

We will generally consider *mild* solutions, which means that the "$dx(t) = \ldots$"–notation formally must be taken as an integral equation. This is inevitable in the case $\sigma(\cdot) \neq 0$, and is also necessary, for instance, when considering deterministic differential equations with involved inhomogeneity that is only integrable. Further, solutions of differential systems are generally supposed to be continuous. In contrast to the formulation of *neutral* functional differential equations, in the literature the formulation in (3.1.1) is commonly called a *retarded* functional differential equation which we will abbreviate as RFDE, or SRFDEs respectively when considering RFDE subject to noise. We will frequently compare a stochastic system, for example the system (3.1.1), with its *deterministic version* or *deterministic counterpart* which simply means that we consider the system without noise, formally letting $\sigma = 0$.

3.2. General Existence and Uniqueness of Solutions

We will say that a mapping $x : [t_0 - r, \infty) \times \Omega \to \mathbb{R}^n$ is a *solution* of (3.1.1), if the following three conditions a), b), c) are satisfied:

a) The process x is continuous and $\{\mathcal{F}_t\}_{t\in[t_0,\infty)}$-adapted.

b) For every finite $T > t_0$ the coefficient processes are reasonably defined, which means that $(H(t, x_t))_{t\in[t_0,T]} \in \mathcal{L}^1_0([t_0, T], \mathbb{R}^n)$ and $(\sigma(t, x_t))_{t\in[t_0,T]} \in \mathcal{L}^2_0([t_0, T], \mathbb{R}^{n\times m})$, where $\mathcal{L}^p_0(A, B)$ denotes the measurable functions $f : A \times \Omega \to B$ with $\int_A |f(u)|^p du < \infty$ $\mathbb{P}$-a.s.

c) The initial condition holds, and the differential law of (3.1.1), interpreted as integral equation, holds $\mathbb{P}$-almost surely for all $t \in [t_0, \infty)$.

Due to [Mao08, Chapter 5] existence and uniqueness of solutions can be achieved by assuming that:

- The coefficients H and σ are *locally Lipschitz* in the second argument uniformly on compacts with respect to the first argument, i.e. for every $T \in (t_0, \infty)$ there is a family of constants $(K_{T,n})_{n\in\mathbb{N}}$ such that for those $\varphi, \psi \in \mathcal{C}(J, \mathbb{R}^n)$ with $\max\{\|\varphi\|, \|\psi\|\} \le n$

$$\max\Big\{\big|H(t,\varphi) - H(t,\psi)\big|, \big|\sigma(t,\varphi) - \sigma(t,\psi)\big|\Big\} \le K_{T,n}\|\varphi - \psi\| \quad \text{for all } t \in [t_0, T],$$

- H and σ satisfy the following *linear growth condition*: For every $T \in (t_0, \infty)$ there is a constant $K_T < \infty$ such that

$$\max\Big\{|H(t,\varphi)|, |\sigma(t,\varphi)|\Big\} \le K_T(1 + \|\varphi\|) \quad \text{for all } (t,\varphi) \in [t_0, T] \times \mathcal{C}(J, \mathbb{R}^n).$$

Then (3.1.1) admits a unique global continuous solution; the solution belongs to $\mathcal{L}^2_{\text{loc}}([t_0 - r, \infty), \mathbb{R}^d)$ and so uniqueness means up to indistinguishability. Implicitly the deterministic case is covered by those assumptions. Roughly speaking, the conditions restricted to the drift coefficient H imply the general Carathéodory conditions in [HVL93, Chapter 2.6] providing local existence; the local Lipschitz property yields uniqueness, and global existence is due to the local linear growth condition. In both cases, the stochastic and the deterministic case, proofs rely on techniques that are well-known from the classical theory of ODEs: In the deterministic case solutions are located in $\mathcal{C}([t_0 - r, \infty), \mathbb{R}^n)$; here an application of the Schauder fixed-point theorem with lower-bounded continuation-step sizes on each compact ensures global existence [HVL93, Theorem 2.1], and a Gronwall-type argument [HVL93, Theorem 2.3] provides uniqueness. Noisy solutions are located in $\mathcal{L}^2_{\text{loc}}([t_0 - r, T], \mathbb{R}^n)$, and Mao's proof uses Picard iterates for existence and again a Gronwall-type argument for uniqueness, see e.g. [Mao08, Chapter 5, Theorem 2.2].

In general, due to the dependence on the last segment of the solution paths, solutions of SRFDEs can not have the Markov property in the sense of an $\mathbb{R}^n$-valued process. But, they actually have that property on the segment level. This property is of particular importance in Chapter 5 and will be tacitly applied. A full grown result can e.g. be found in [Moh84, Chapter 3] or be adapted from the argument in [Sch84].

3.3. Representations for Linear RFDEs with Additive Noise

As linear SRFDEs we refer to systems where the drift coefficient $H(t,\psi)$ is affine linear in ψ for each t, which means $H(t,\psi) = L(t,\psi) + h(t)$ for some operator $L : \mathbb{R} \times \mathcal{C}(J, \mathbb{R}^n) \to \mathbb{R}^n$ that is linear with respect to the second argument, and an *inhomogeneity map* $h : [t_0, \infty) \to \mathbb{R}^n$. As it is common practice we will use the notations $L(t,\psi) = L(t)(\psi) = L(t)\psi$, and we will occasionally refer to L as a family of operators, e.g. $(L(t))_{t\in[t_0,\infty)}$. For later referencing we put this special case of (3.1.1) in display:

$$\begin{cases} dx(t) = L(t)x_t dt + h(t)dt + \sigma(t)dW(t) \quad \text{for } t \ge t_0, \\ x_{t_0} = \Upsilon. \end{cases} \tag{3.3.1}$$

Remember that, by the Riesz representation theorem, the linear operators $(L(t))_{t\in[t_0,\infty)}$ may uniquely be extended from $\mathcal{C}(J,\mathbb{R}^n)$ to $\mathcal{B}^b(J,\mathbb{R}^n)$, where $\mathcal{B}^b$ denotes the measurable and bounded mappings. This unique extension will be tacitly applied when needed, and the extended family of linear operators will also be denoted by the same symbols $(L(t))_{t\in[t_0,\infty)}$. Regarding the deterministic version, the conditions for existence and uniqueness are carried over from the account of Hale and Lunel, [HVL93, chapter 6], thereby fixing the related notations to have them at hand later on.

Assumption 3.1 (Hale–Lunel conditions for global existence and uniqueness)**.** *There is an* $m \in \mathscr{L}^1_{\text{loc}}\big([t_0,\infty)\times\mathbb{R},\mathbb{R}^{n\times n}\big)$*, which means locally Lebesgue-integrable,* $n\times n$ *matrix-valued function* $\eta(t,u)$*, measurable in* $(t,u)\in\mathbb{R}\times\mathbb{R}$*, so that*

$$\eta(t,u) = \begin{cases} 0 & \text{for } u \geq 0, \\ \eta(t,-r) & \text{for } u \leq -r, \end{cases} \tag{3.3.2}$$

continuous from the left in all $u \in (-r,0)$ *and has bounded variation in* u *on* $[-r,0]$ *for each* t*. And the variation with respect to* u *is bounded through*

$$\mathrm{Var}_{[-r,0]}\,\eta(t,\cdot) \leq m(t) \qquad \text{for all } t \geq t_0, \tag{3.3.3}$$

and the linear mapping $L(t) : \mathcal{C}(J,\mathbb{R}^n) \to \mathbb{R}^n$ *is given by*

$$L(t)\psi = \int_{-r}^{0} \psi(u)d_u\eta(t,u) \qquad \text{for all } t\in(-\infty,\infty),\ \psi\in\mathcal{C}(J,\mathbb{R}^n),$$

where d_u *indicates that the Lebesgue–Stieltjes integration is carried out with respect to the* u*-argument of the integrator, and* t *is fixed. In particular,* $|L(t)\psi| \leq m(t)\|\psi\|$*.*

Together with the $\mathscr{L}^1_{\text{loc}}$-assumption on h Assumption 3.1 ensures existence and uniqueness of global solutions in the deterministic case. These Hale–Lunel conditions are satisfied if we, for instance, assume the family L to be continuous with repect to the sup-norm on $[t_0,T]\times\mathcal{C}(J,\mathbb{R}^n)$, given by

$$\|(t,\psi)\|_{[t_0,T]\times\mathcal{C}(J,\mathbb{R}^n)} := \max\{|t|,\|\psi\|\} \qquad \text{for all } t\geq t_0, \psi\in\mathcal{C}(J,\mathbb{R}^n). \tag{3.3.4}$$

Fundamental Solutions. The concept of fundamental solutions, which is a generalization from classical theory of ordinary differential equations, will be of vital importance for this work due to its crucial role in the variation-of-constants formula. This extract from the book [HVL93] outlines a formal definition of the fundamental matrix solution, and reviews the solution representation through the variation-of-constants formula in the nonautonomous deterministic case; we will not present every detail, but mainly follow the main ideas from the introduction of an appropriate resolvent kernel in order to rigorously define fundamental solutions to solution representations. All details can be found in [HVL93, chapter 6]. Informally speaking the variation-of-constants formula originates from the linear differential law and does not get in the way of the retarded feedback mechanism.
First, we rewrite the solution of the deterministic version of (3.3.1) with an application of the integration by parts formula, which is applicable due to absolute continuity of the solution x, and where we write the formal weak derivative of x as $\dot{x}$. As we mentioned before,

related differential formulas have to be understood as integrated equations. We obtain that

$$\begin{aligned}\dot{x}(t) &= \int_{t_0}^{t} x(u)d_u\eta(t,u-t) + \int_{-r}^{t_0-t} \Upsilon(t-t_0+u)d_u\eta(t,u) + h(t) \\ &= -\eta(t,t_0-t)x(t_0) - \int_{t_0}^{t} \eta(t,u-t)\dot{x}(u)du + \int_{-r}^{t_0-t} \Upsilon(t-t_0+u)d_u\eta(t,u) + h(t) \\ &\qquad \text{for all } t \in [t_0,\infty).\end{aligned} \tag{3.3.5}$$

We define $k(t,s) := \eta(t,s-t)$, $s,t \in [t_0,\infty)$, a kernel of type $\mathcal{L}^1_{\text{loc}}$ on $[t_0,\infty)$, in order to reformulate (3.3.5) with $y(t) = \dot{x}(t)$ as a *Volterra equation* of the second kind,

$$y(t) = \int_{t_0}^{t} k(t,u)y(u)du + g(t) \quad \text{for Lebesgue-a.e. } t \in [t_0,\infty), \tag{3.3.6}$$

where $g \in \mathscr{L}^1_{\text{loc}}([t_0,\infty),\mathbb{R}^n)$ is given by the collection of terms from inhomogeneity and initial-segment influence, namely

$$g(t) := -\eta(t,t_0-t)\Upsilon(0) + \int_{-r}^{t_0-t} \Upsilon(t-t_0+u)d_u\eta(t,u) + h(t) \quad \text{for all } t \geq t_0.$$

From the corresponding theory of Volterra equations, we conclude that there is a *Volterra resolvent* R satisfying

$$R(t,s) = -\eta(t,s-t) + \int_{s}^{t} R(t,u)\eta(u,s-u)du \quad \text{for all } t \geq s, s \in [t_0,\infty), \tag{3.3.7}$$

and it is unique in the $\mathscr{L}^1$-sense on every finite time horizon. By means of a Gronwall-type argument, the variation condition (3.3.3) implies

$$|R(t,s)| \leq m(t)\exp\left(\int_{s}^{t} m(u)du\right) \quad \text{for all } t \geq s, s \in [t_0,\infty). \tag{3.3.8}$$

We define the fundamental matrix solution $\breve{x}$ as

$$\breve{x}(t,s) := I_n - \int_{s}^{t} R(u,s)du \quad \text{for all } s \in [t_0,\infty),\ t \geq s, \tag{3.3.9}$$

where I_n denotes the n-dimensional unit matrix. We may interpret the fundamental solution $(\breve{x}(t,u) : u \in [t_0,\infty), t \geq u-r)$ as the family of matrix solutions of the homogeneous deterministic systems

$$\begin{cases} dx(t) = L(t)x_t\,dt & \text{for } t \geq u, \\ x(t) = \mathbb{1}_{\{u\}}(t)I_n & \text{for } t \in [u-r,u], \end{cases} \tag{3.3.10}$$

where the differential law $L(t)$ is taken as separately acting on the column vectors. As we have pointed out before, the existence of solutions of the deterministic version of (3.3.1) follows from an application of the Schauder fixed-point theorem, and crucially relies on the continuity of the initial segment Υ, which means that (3.3.10) is not covered through that approach due to its discontinuous initial segment. The slight detour to the Volterra resolvent provides a rigorous definition of the fundamental solution. In the first argument the fundamental solution is absolutely continuous, solves the integral equation and its differential law applies almost everywhere with respect to the Lebesgue measure. Continuing from (3.3.8), we can conclude that

$$|\check{x}(t,s)| \leq \exp\left(\int_s^t m(u)du\right) \quad \text{for all } s \in [t_0,\infty), t \geq s, \tag{3.3.11}$$

and for any finite time horizon $T>0$, due to boundedness of the resolvent in (3.3.8), there is $c_R = c_R(T) > 0$ such that for all $\Delta \in \mathbb{R}$ with $t+|\Delta| \leq T$ and $t-|\Delta| \geq u$

$$|\check{x}(t+\Delta,u) - \check{x}(t,u)| \leq c_R|\Delta| \quad \text{for all } u \in [0,T],\ t \in [u,T]. \tag{3.3.12}$$

That means the fundamental solution is locally uniformly Lipschitz in the first argument with respect to compacts of the second argument. The general existence and uniqueness result for solutions of the deterministic version of (3.3.1) also covers the corresponding homogeneous system started at any intermediate time point $s \in [t_0,T]$ initiated with some $\psi \in \mathcal{C}(J,\mathbb{R}^n)$, formally given by

$$\begin{cases} dx(t) = L(t)x_t\,dt & \text{for } t \geq s, \\ x_s = \psi. \end{cases} \tag{3.3.13}$$

That means that there is a solution semi group $(T^{\det}_{t,s} : s \in [t_0,\infty), t \geq s)$ that shoves segments from $\mathcal{C}(J,\mathbb{R}^n)$ along the solution path into $\mathcal{C}(J,\mathbb{R}^n)$ according to the deterministic differential law. In other words, if we denote $(z(t) : t \geq t_0)$ the solution of (3.3.13) for $s=t_0$, then $z_t = T^{\det}_{t,t_0}\psi$ for all $t \geq t_0$. Due to [HVL93, Chapter 6.1, 6.2] the unique solution of the inhomogeneous system is then given by

$$x(t) = T^{\det}_{t,t_0}\Upsilon(0) + \int_{t_0}^t \check{x}(t,u)h(u)du \quad \text{for all } t \geq t_0. \tag{3.3.14}$$

Example 3.2. *a) This special case is taken from [HVL93]. For arbitary $N \in \mathbb{N}$ and $r>0$ let $A_k \in \mathbb{R}^{n\times n}$, $k \in \{1,\dots,N\}$ be a family of constant matrices, and $r_k \in (0,r)$, $k \in \{1\dots N\}$, a collection of delay lengths. Assume further some $A : \mathbb{R}\times\mathbb{R} \to \mathbb{R}^{n\times n}, (t,u) \mapsto A(t,u)$, that is integrable in u for every t, and that there is some function $a \in \mathscr{L}^1_{\mathrm{loc}}(\mathbb{R},\mathbb{R})$ such that*

$$\left|\int_{-r}^0 A(t,u)\psi(u)du\right| \leq a(t)\|\psi\| \quad \textit{for all } t \in \mathbb{R},\ \psi \in \mathcal{C}(J,\mathbb{R}^n).$$

If we moreover assume that $h \in \mathscr{L}^1_{\mathrm{loc}}$, and let $t_0 \in \mathbb{R}$, and $\Upsilon \in \mathcal{C}(J,\mathbb{R}^n)$ arbitrary, then the system

$$\begin{cases} dx(t) = \sum_{i=1}^N A_i x(t-r_i)dt + \int_{-r}^0 A(t,u)x(t+u)\,du\,dt + h(t)dt & \textit{for } t \geq t_0, \\ x_{t_0} = \Upsilon, \end{cases} \tag{3.3.15}$$

satisfies Assumption 3.1 and therefore, there is a unique solution and it may be represented in the form (3.3.14). The reason for bringing up this particular example is that J. Hale and S. Verduyn Lunel refer to it as the most common type of linear systems with finite lag which is known to be useful in applications, see [HVL93, Chapter 6.1].

b) This one is a modification of the above example. It is an instance of a continuous family of continuous linear operators, which is to say that $(L(t))_{t\in[t_0,T]}$*, as a mapping from* $[t_0,T]\times\mathcal{C}(J,\mathbb{R}^n)$*, is continuous with respect to* $\|\cdot\|_{[t_0,T]\times\mathcal{C}(J,\mathbb{R}^n)}$*, see* (3.3.4)*. This example keeps jump positions fixed, but allows time dependence for the height of jumps. For arbitary* $N\in\mathbb{N}$ *and* $r>0$ *let* $A_k:[t_0,\infty)\to\mathbb{R}^{n\times n}$*,* $k\in\{1,\dots,N\}$*, be a family of continuously differentiable* $\mathbb{R}^{n\times n}$*-valued functions, and* $r_k\in(0,r)$*,* $k\in\{1\dots N\}$*, a collection of delay lengths. Assume further some* $A:[t_0,\infty)\times\mathbb{R}\to\mathbb{R}^{n\times n},(t,u)\mapsto A(t,u)$ *that is integrable in* u *for every* t *and that there is some function* $a\in\mathscr{L}^1_{\mathrm{loc}}(\mathbb{R},\mathbb{R})$ *such that*

$$\left|\int_{-r}^{0}A(t,u)\psi(u)du\right|\le a(t)\|\psi\|\quad \textit{for all } t\in[t_0,\infty),\ \psi\in\mathcal{C}(J,\mathbb{R}^n).$$

We additionally assume that $A(t,u)$ *is continuously differentiable in* t*. Then, for* $h\in\mathscr{L}^1_{\mathrm{loc}}$*, and* $\Upsilon\in\mathcal{C}$*, the system*

$$\begin{cases} dx(t)=\sum_{i=1}^{N}A_i(t)x(t-r_i)dt+\int_{-r}^{0}A(t,u)x(t+u)\,du\,dt+h(t)dt & \textit{for } t\ge t_0,\\ x_{t_0}=\Upsilon, \end{cases}\tag{3.3.16}$$

satisfies condition (3.1) from above with

$$\eta(t,u)=-\int_{u}^{0}A(t,v)dv-\sum_{i=1}^{N}A_i(t)\mathbb{1}_{\{u\le -r_i\}}\quad \textit{for } t\in\mathbb{R},\ u\in J.\tag{3.3.17}$$

It is generally true that systems of this form admit fundamental solutions that are Lipschitz-continuous in both arguments, see Lemma A.3 in the appendix. Further, this special case contains systems of the form

$$\begin{cases} dx(t)=-a(t)x(t)dt+b(t)x(t-r)dt+h(t)dt & \textit{for } t\in[t_0,T],\\ x_{t_0}=\Upsilon, \end{cases}\tag{3.3.18}$$

if we assume the coefficients $a,b\in\mathcal{C}^1([t_0,T],\mathbb{R})$*, i.e. to be continuously differentiable. Those systems play a crucial role in the second part of this work.*

In case of an autonomous drift coefficient $L(\cdot)=L$, the local Lipschitz property simplifies to ordinary continuity of L. In case of additive noise the stochastically perturbed system can also be described with the help of the fundamental solution by means of a stochastic variation-of-constants formula. Especially, for systems of the form

$$\begin{cases} dx(t)=Lx_tdt+\sigma(t)dW(t) & \text{for } t\ge t_0,\\ x_{t_0}=\Upsilon, \end{cases}\tag{3.3.19}$$

we cite a representation result from the book of S.-E. A. Mohammed, [Moh84]. For the deterministic version of (3.3.19), the solution semi group $(T^{\mathrm{det}}_{t,u}:u\ge t_0,t\in[u,\infty))$ from $\mathcal{C}(J,\mathbb{R}^n)$ to $\mathcal{C}(J,\mathbb{R}^n)$ does only depend on $t-u$ which motivates us to write

$$T^{\mathrm{det}}_{t-u}:=T^{\mathrm{det}}_{t,u}\quad \text{for all } u\in[t_0,\infty),\ t\ge u.$$

And analogously for the fundamental solution $\check{x}(t-u) := \check{x}(t,u)$ for $u \in [t_0,\infty),\ t \geq u-r$.

Proposition 3.3 ([Moh84], Chapter 4, Theorem (4.1), Remark (4.2))**.** *Suppose that* $(T_s^{\mathrm{det}})_{s\geq 0}$ *denotes the solution semi group of the deterministic version of* (3.3.19) *where* $L : \mathcal{C}(J,\mathbb{R}^n) \to \mathbb{R}^n$ *is continuous linear,* $\sigma : [t_0,\infty) \to \mathbb{R}^{n\times m}$ *is locally square integrable,* $\Upsilon \in \mathcal{C}(J,\mathbb{R}^n)$ *and* $(W(u))_{u\in[t_0,\infty)}$ *is an* m*-dimensional Brownian motion. Then there is a unique strong solution* $x = (x(t))_{t\in[t_0,\infty)}$ *of the SRFDE* (3.3.19) *and it admits the representation*

$$x(t) = T^{\mathrm{det}}_{t-t_0}\Upsilon(0) + \int_{t_0}^{t} \check{x}(t-u)\sigma(u)dW(u) \qquad \textit{for all } t \geq t_0 \ \mathbb{P}\textit{-a.s.} \tag{3.3.20}$$

The proof that is presented in [Moh84, Lemmas 4.3, 4.4, Theorem 4.1] uses relatively strong assumptions due to ensure a formula for the differential of a stochastic integral. We will generalize the result by closely related ideas using absolute continuity of the fundamental solution in the first argument and the stochastic Fubini theorem, which one can find in [Jac79, Théorème 5.44] for the finite-dimensional case in french language, or in a rather general Hilbert-space setting in [DPZ14, Theorem 4.33]. Our first objective is to show that our candidate solution has a (Hölder)-continuous modification.

Lemma 3.4. *If we denote the fundamental solution of* (3.3.1) *by* $(\check{x}(t,u) : u \in [t_0,T], t \in [u-r,T])$ *and assume that* $\sigma \in \mathcal{B}^b([t_0,T],\mathbb{R}^{n\times m})$*, i.e. bounded and Borel-measurable, with* $\sup_{u\in[t_0,T]}|\sigma(u)| =: \sigma_+$*, in case of the Hale–Lunel conditions 3.1 the process* z*, defined by*

$$z(t) := \int_{t_0}^{t} \check{x}(t,u)\sigma(u)dW(u) \qquad \textit{for all } t \in [t_0,T]$$

has a Hölder-continuous version of order $\gamma \in (0,1/2)$.

Proof. This can be seen by an application of the Kolmogorov continuity criterion applied to

$$\int_{t_0}^{t} \sigma(u)dW(u) - \int_{t_0}^{t} \check{x}(t,u)\sigma(u)dW(u) \qquad \text{for } t \in [t_0,T].$$

Due to the local Lipschitz continuity of $\check{x}$ in the first argument, see (3.3.12), we find that for $\Delta > 0$

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t+\Delta}\Big(I_n - \check{x}(t+\Delta,u)\Big)\sigma(u)dW(u) - \int_{t_0}^{t}\Big(I_n - \check{x}(t,u)\Big)\sigma(u)dW(u)\right|^2\right] \\
&= \mathbb{E}\left[\left|\int_{t_0}^{t}\Big(\check{x}(t+\Delta,u) - \check{x}(t,u)\Big)\sigma(u)dW(u)\right|^2\right] \\
&\quad + \mathbb{E}\left[\left|\int_{t}^{t+\Delta}\Big(I_n - \check{x}(t+\Delta,u)\Big)\sigma(u)dW(u)\right|^2\right] \\
&\leq \int_{t_0}^{t}\Big|\check{x}(t+\Delta,u) - \check{x}(t,u)\Big|^2\sigma_+^2 du + \int_{t}^{t+\Delta}\Big|I_n - \check{x}(t+\Delta,u)\Big|^2\sigma_+^2 du \\
&\leq \sigma_+^2(t-t_0)c_R^2\Delta^2 + \sigma_+^2\int_{t}^{t+\Delta} c_R^2\Delta^2 du \leq \mathrm{const}\,\Delta^2,
\end{aligned}$$

where in the second to the last inequality we have used Itô isometry and that $I_n = \check{x}(u,u)$ and therefore, $\big|\check{x}(t+\Delta,u) - I_n\big| \leq c_R\Delta$ for all $u \in [t,t+\Delta]$ for appropriate c_R, see (3.3.12).

From the general theory, we know that $\int_{t_0}^{t} \sigma(u)dW(u)$, $t \in [t_0, T]$, admits Hölder-continuous sample paths of order $\gamma \in (0, 1/2)$ almost surely, and therefore there is an almost surely Hölder-continuous version of $\int_{t_0}^{t} \check{x}(t,u)\sigma(u)dW(u),\ t \in [t_0, T]$ of order $\gamma \in (0, 1/2)$. □

In the following, when considering the stochastic integral process, defined in Lemma 3.4, we will refer to its continuous version. The next objective is to give a generalization of the solution representation that is presented in [Moh84], stated above as Proposition 3.3.

Theorem 3.5 (General Representation Theorem). *Consider the situation of the Lemma 3.4 and let $(T^{\det}_{t,u} : u \geq t_0, t \geq u)$ denote the solution semi group from $\mathcal{C}(J, \mathbb{R}^n)$ to $\mathcal{C}(J, \mathbb{R}^n)$ of the deterministic version of the homogeneous SRFDE*

$$\begin{cases} dx(t) = L(t)x_t dt + \sigma(t)dW(t) & \text{for } t \in [t_0, T), \\ x_{t_0} = \Upsilon. \end{cases} \tag{3.3.21}$$

Then, for arbitrary finite time horizon $T > t_0$, the unique solution of (3.3.21) *is $\mathbb{P}$-almost surely given by*

$$y(t) := T^{\det}_{t,t_0}\Upsilon(0) + \int_{t_0}^{t} \check{x}(t,u)\sigma(u)dW(u) \qquad \text{for all } t \in [t_0, T],$$

where the stochastic integral term is understood as the continuous version ensured by the previous Lemma 3.4.

Proof. We go over the arguments deliberately in small steps. Due to its definition $(T^{\det}_{t,t_0}\Upsilon(0) : t \geq t_0)$ solves the deterministic version of (3.3.21) in t, which is to say that

$$\begin{aligned} T^{\det}_{t,t_0}\Upsilon(0) &= \Upsilon(0) + \int_{t_0}^{t} \frac{\partial}{\partial s} T^{\det}_{s,t_0}\Upsilon(0)ds \qquad \text{and} \\ \frac{\partial}{\partial t} T^{\det}_{t,t_0}\Upsilon(0) &= L(t)(T^{\det}_{t,t_0}\Upsilon) = \int_{-r}^{0} T^{\det}_{t,t_0}\Upsilon(\theta)\, d_\theta \eta(t,\theta) \qquad \text{for all } t \in [t_0, T]. \end{aligned} \tag{3.3.22}$$

Further, we know that the fundamental solution solves the respective integral equation of the deterministic system in the first argument, which means

$$\check{x}(t,u) = \check{x}(u,u) + \int_{u}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\, d_\theta \eta(s,\theta)\, ds \quad \text{for all } u \in [t_0, T], t \geq u. \tag{3.3.23}$$

And due to the fact that $\check{x}(s+\theta, u) = 0$ for all $s \in [t_0, u)$ and $\theta \in [-r, 0]$, we may exchange the u for t_0 in the lower boundary of the right-hand side integral above. We obtain that

$$\check{x}(t,u) = I_n + \int_{t_0}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\, d_\theta \eta(s,\theta)\, ds \qquad \text{for all } t \geq u.$$

Using (3.3.22) and (3.3.23) to rewrite $(y(t))_{t \in [t_0, T]}$ leads to

$$\begin{aligned} y(t) = \Upsilon(0) &+ \int_{t_0}^{t} \frac{\partial}{\partial s} T^{\det}_{s,t_0}\Upsilon(0)ds \\ &+ \int_{t_0}^{t} \check{x}(u,u)\sigma(u) + \int_{u}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\sigma(u)\, d_\theta \eta(s,\theta)\, ds\, dW(u) \qquad \text{for all } t \in [t_0, T]. \end{aligned}$$

As before, we may replace the u in the lower integral boundary on the right by t_0, because the integrand is zero for all $s \in [t_0, u)$.

$$y(t) = \Upsilon(0) + \int_{t_0}^{t} \int_{-r}^{0} T^{\text{det}}_{s,t_0} \Upsilon(\theta) d_\theta \eta(s,\theta)\, ds + \int_{t_0}^{t} \int_{t_0}^{t} \int_{-r}^{0} \check{x}(s+\theta, t_0)\sigma(u) d_\theta \eta(s,\theta)\, ds\, dW(u)$$
$$+ \int_{t_0}^{t} \check{x}(u,u)\sigma(u) dW(u) \qquad \text{for all } t \in [t_0, T].$$

For the triple-integral term, if we understand $\int_{t_0}^{t} \check{x}(s+\theta, u) dW(u)$, $u \in [t_0, t]$ as a stochastic integral parametrized by (s, θ), we may apply the stochastic Fubini theorem (see [Jac79, Théorème 5.44] (or [DPZ14, Theorem 4.33]) to interchange the order of integration. As an intermediate step $\eta(s, \cdot) \otimes ds$ must formally be split into a difference of two positive, finite measures. We have to check that both appearing stochastic integrals are well-defined, which means predictability, i.e. measurability with respect to the filtration that is generated by the left-continuous and adapted processes, of the integrand as well as $\mathcal{L}^2$-integrability. But concerning the first stochastic integral $\int_{t_0}^{t} \check{x}(s+\theta, u)\sigma(u) dW(u)$ intricate measurability issues do not arise, because the integrand $\check{x}(s+\theta, u)\sigma(u)$ is deterministic and bounded, in particular predictable. And therefore, the stochastic integral $\int_{t_0}^{t} \check{x}(s+\theta, u)\sigma(u) dW(u)$ is predictable in t, see e.g. [Jac79]. And also $\mathcal{L}^2$-integrability is ensured by boundedness of the integrand. Regarding the second stochastic integral

$$\int_{t_0}^{t} \int_{u}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\sigma(u) d_\theta \eta(s,\theta)\, ds\, dW(u),$$

the same reasoning holds true and is not affected by a decomposition of the $d_\theta(\eta(t,\theta))dt$-measure in positive and negative part. We obtain that

$$\int_{t_0}^{t} \int_{t_0}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\sigma(u) d_\theta \eta(s,\theta)\, ds\, dW(u) = \int_{t_0}^{t} \int_{-r}^{0} \int_{t_0}^{t} \check{x}(s+\theta, u)\sigma(u) dW(u)\, d_\theta \eta(s,\theta)\, ds$$
$$\mathbb{P}\text{-almost surely for a dense subset in } t \text{ from } [t_0, T]. \tag{3.3.24}$$

Note further that, simply because $\check{x}$ solves the integrated equation for the homogeneous deterministic system,

$$\int_{t_0}^{t} \int_{u}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\sigma(u) d_\theta \eta(s,\theta)\, ds\, dW(u) \tag{3.3.25}$$
$$= \int_{t_0}^{t} \int_{u}^{t} \int_{-r}^{0} \check{x}(s+\theta, u)\sigma(u) d_\theta \eta(s,\theta)\, ds\, dW(u)$$
$$= \int_{t_0}^{t} \int_{u}^{t} L(s)\big(\check{x}(s+\theta, u) : \theta \in J\big)\sigma(u) ds\, dW(u)$$
$$= \int_{t_0}^{t} (\check{x}(t,u) - \check{x}(u,u))\sigma(u) dW(u)$$
$$= \int_{t_0}^{t} (\check{x}(t,u) - I_n)\sigma(u) dW(u) \qquad \text{for all } t \in [t_0, T]. \tag{3.3.26}$$

That provides continuous paths in t almost surely with respect to $\mathbb{P}$ due to Lemma 3.4 and the choice of the continuous version. Continuing from (3.3.24) we may decline the right-hand side inner integral to an upper boundary of $s+\theta$, which ensures continuity of the stochastic integral term on the right due to construction. Further, with regard to (3.3.26), we know that the term in line (3.3.25) is $\mathbb{P}$-almost surely continuous in t, again due to construction. Therefore, we can understand the two sides of (3.3.24) as two continuous processes in t that match $\mathbb{P}$-almost surely on a dense subset of $[t_0, T]$. So, they must be the same up to indistinguishability, i.e.

$$\begin{aligned}\int_{t_0}^{t}\int_{t_0}^{t}\int_{-r}^{0} &\check{x}(s+\theta,u)\sigma(u)d_\theta\eta(s,\theta)\,ds\,dW(u)\\ &= \int_{t_0}^{t}\int_{-r}^{0}\int_{t_0}^{s+\theta}\check{x}(s+\theta,u)\sigma(u)dW(u)\,d_\theta\eta(s,\theta)\,ds \quad \text{for all } t\in[t_0,T] \ \ \mathbb{P}\text{-almost surely.}\end{aligned}$$

Applying that to the term y we find that $\mathbb{P}$-almost surely

$$\begin{aligned}y(t) &= \Upsilon(0) + \int_{t_0}^{t}\int_{-r}^{0} T^{\det}_{s,t_0}\Upsilon(\theta) + \int_{t_0}^{s+\theta}\check{x}(s+\theta,u)\sigma(u)dW(u)\,d_\theta\eta(s,\theta)\,ds + \int_{t_0}^{t}\sigma(u)dW(u)\\ &= \Upsilon(0) + \int_{t_0}^{t}\int_{-r}^{0} y(s+\theta)d_\theta\eta(s,\theta)\,ds + \int_{t_0}^{t}\sigma(u)dW(u) \quad \text{for all } t\in[t_0,T].\end{aligned}$$

Or, in other words and short-hand differential notation respectively, we find that

$$y(t) = \Upsilon(0) + \int_{t_0}^{t} L(s)y_s ds + \int_{t_0}^{t}\sigma(u)dW(u), \ \text{ or } \quad dy(t) = L(t)y_t dt + \sigma(t)dW(t)$$
$$\text{for all } t\in[t_0,T] \ \mathbb{P}\text{-a.s.} \qquad (3.3.27)$$

By uniqueness of solutions, which is covered in Section 3.2 of this work, this settles the proof. □

Of course, due to the linearity of the inhomogeneous nonautonomous system (3.3.1), we find that the solution of (3.3.1) may now be given explicitly. To put a label to it, we stow that fact in the following corollary.

Corollary 3.6. *Under the assumptions of Lemma 3.4, the solution of* (3.3.1) *is $\mathbb{P}$-almost surely given by*

$$x(t) = T^{\det}_{t,t_0}\Upsilon(0) + \int_{t_0}^{t}\check{x}(t,u)h(u)du + \int_{t_0}^{t}\check{x}(t,u)\sigma(u)dW(u) \quad \textit{for all } t\geq t_0. \qquad (3.3.28)$$

In particular, the solution is a continuous Gaussian process in $\mathbb{R}^n$.

Remark 3.7. *We will refer to the solution formulas of the form* (3.3.14), (3.3.20) *and* (3.3.28) *as variation-of-constants formulas. Their kind has approved as a helpful tool in the study of stochastic retarded functional differential equations. And they will do so in the second part of this work.*

3.4. Concentration of Sample Paths in Autonomous Stable Environment

The case of an autonomous homogeneous linear RFDE subject to additive noise will serve us as a basic example of an application of the Fernique inequality where we generally assume the setting and notations from Theorem 3.5. Because it is of some interest on its own, and some special cases will reappear throughout the work, this example has been devoted a section on its own. Let us suppose that $(x(t))_{t\geq t_0-r}$ is the solution of

$$\begin{cases} dx(t) = Lx_t dt + \sigma(t)dW(t) & \text{for } t \geq t_0, \\ x_t = \Upsilon, \end{cases} \tag{3.4.1}$$

where L, σ and Υ satisfy the assumptions of Proposition 3.3, or Theorem 3.5 respectively. We start reviewing known relations between negative eigenvalue real parts and stability neatly presented in [HVL93, Chapter 7 and Lemma 5.3 in Chapter 6]. We assume that all roots of the *characteristic equation*

$$\det D(\lambda) = 0, \qquad \text{where } D(\lambda) = \lambda I - L\big(e^{\lambda u} : u \in [-r, 0]\big), \tag{3.4.2}$$

have negative real parts. For any $\overline{\lambda} \in \mathbb{R}$ the set $\{\lambda \in \mathbb{C} : D(\lambda) = 0 \text{ and } \Re(\lambda) \geq \overline{\lambda}\}$ of roots of the characteristic equation with real part $\Re(\cdot)$ at least of size $\overline{\lambda}$ is finite. So, the assumption of negative real parts includes boundedness away from zero. And this assumption is sufficient for the system to be asymptotically exponentially stable, which means that any solution of (3.4.1), that corresponds to a feasible initial segment $\Upsilon \in \mathcal{C}(J, \mathbb{R}^n)$, approaches 0 exponentially fast. This includes the fundamental solutions which we keep denoting as $\breve{x} = \big(\breve{x}(t, u) : u \in [t_0, \infty), t \geq u - r\big)$. We have the following estimate concerning the fundamental matrix solution $\breve{x}$ with $\breve{x}(t, u) = \breve{x}(t - u)$ for all $u \in [0, \infty), t \in [u - r, \infty)$, namely, then there exist $K > 0$, $\gamma > 0$ such that

$$|\breve{x}(t - t_0)| \leq K e^{-\gamma(t-t_0)} \qquad \text{for all } t \in [t_0, \infty). \tag{3.4.3}$$

Note that in the autonomous case $m(\cdot)$ from (3.3.3) is a constant. Therefore, the additional assumption $\int_t^{t+r} m(u)du < m_1 < \infty$ in [HVL93, Chapter 6, Lemma 5.2] is trivially satisfied here. It is further worth mentioning that in this situation for constant noise amplifier $\sigma(\cdot) = \sigma$, J. A. D. Appleby, X. Mao and H. Wu [AMW10] derived the exact essential growth rate,

$$\limsup_{t\to\infty} \frac{x_i(t)}{\sqrt{2\log t}} = \overline{\sigma}_i \quad \text{and} \quad \liminf_{t\to\infty} \frac{x_i(t)}{\sqrt{2\log t}} = -\overline{\sigma}_i \quad \text{for all } i \in \{1, \dots, n\} \quad \mathbb{P}\text{-a.s.},$$

where

$$\overline{\sigma}_i = \sqrt{\sum_{k=1}^m \int_{t_0}^\infty (\breve{x}(u)\sigma)_{ik}^2 du} \qquad \text{for all } i \in \{1, \dots, n\}.$$

For later reference we additionally note the one-dimensional case;

$$\limsup_{t\to\infty} \frac{x(t)}{\sqrt{2\log t}} = \sigma \sqrt{\int_{t_0}^{t} \check{x}^2(u)du} \quad \text{and} \quad \liminf_{t\to\infty} \frac{x(t)}{\sqrt{2\log t}} = -\sigma \sqrt{\int_{t_0}^{t} \check{x}^2(u)du} \quad \mathbb{P}\text{-a.s.} \tag{3.4.4}$$

To keep things simple, we consider the case $t_0 = 0$, $n = 1$, the time horizon $T > 0$ is finite, and the diffusion coefficient is bounded over $[0, T]$ by some $\sigma_+ > 0$. Due to linearity, the deviation process $y = (y(t))_{t\in[0,T]}$, defined through $y(t) := x(t) - \mathbb{E}[x(t)]$, $t \in [0, T]$, solves (3.4.1) with $\Upsilon = 0$, and due to Proposition 3.3, it allows for an explicit formulation in terms of the fundamental solution $\check{x}$;

$$y(t) = \int_0^t \check{x}(t-u)\sigma(u)dW(u) \quad \text{for all } t \in [0, T]. \tag{3.4.5}$$

The stochastic-integral process $(y(t))_{t\in[0,T]}$ does not have the martingale property, while the deterministic integrand ensures the process to be Gaussian. The next theorem constitutes the main result of this section, and basically captures what can be learned about the distributional concentration of system (3.4.1) through the Fernique inequality. Sample-path continuity of $(y(t))_{t\in[0,T]}$ is clear, and the applicability of the Fernique inequality will be covered during the proof of the following Theorem. Apart from that, the proof consists of the computation of the parameters that are involved in the Fernique inequality. The lack of concrete knowledge of the constants in (3.4.3) remains unsatisfactory.

For an ease of notation we use that $2 + \sqrt{2} \leq 7/2$.

Theorem 3.8. *For the deviation process $(y(t))_{t\in[-r,T]}$ of system (3.4.1) in dimension $n = 1$, we assume that the roots of the characteristic equation (3.4.2) have all negative real part with the notations from (3.4.3). Let $m(\cdot) = m_L$ denote the variation of L from (3.3.3), then*

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |y(s)| > h\left(\sqrt{\operatorname{var} y(T)} + \frac{7}{2}\left(\frac{\sigma_+ K e^{\gamma r} m_L T}{\sqrt{2\gamma}\, 2p\log p} + \frac{\sigma_+ K\sqrt{T}}{\sqrt{p}\log p}\right)\right)\right\} \leq \frac{5p^2}{2}\exp\left(-\frac{h^2}{2}\right)$$

for all $p \in \mathbb{N}, p \geq 2$ and $h > \sqrt{1 + 4\log p}$.

Proof. By (3.4.3), it is easy to see that

$$\|\Gamma\| = \sup_{t\in[0,T]} \mathbb{E}\left[y(t)^2\right] \leq \operatorname{var} y(T) \leq \frac{\sigma_+^2 K^2}{2\gamma}. \tag{3.4.6}$$

Regarding the Fernique inequality in Theorem 2.1, we introduce the notation

$$Q(p, T) := \left(2 + \sqrt{2}\right) \int_1^\infty \varphi\left(Tp^{-u^2}\right) du, \tag{3.4.7}$$

where we understand $p \in \mathbb{N}$ as arbitrary, but fixed. We note that φ takes the form

$$\varphi(Tp^{-u^2}) = \sqrt{\sup_{\substack{s,t\in[0,T],s<t,\\ |t-s|\leq Tp^{-u^2}}} \mathbb{E}\left[\left(\int_0^t \check{x}(t-v)\sigma(v)dW(v) - \int_0^s \check{x}(s-v)\sigma(v)dW(v)\right)^2\right]}. \tag{3.4.8}$$

Due to independent increments of Brownian motion and Itô isometry it is easy to see that

$$\frac{Q(p,T)}{2+\sqrt{2}} \leq \mathcal{Q}_1(p,T) + \mathcal{Q}_2(p,T), \tag{3.4.9}$$

where

$$\mathcal{Q}_1(p,T) := \int_1^\infty \sqrt{\sup_{\substack{s,t\in[0,T],s<t,\\ |t-s|\leq Tp^{-u^2}}} \int_0^s \sigma^2(v)\Big(\check{x}(t-v)-\check{x}(s-v)\Big)^2 dv}\, du, \tag{3.4.10}$$

$$\mathcal{Q}_2(p,T) := \int_1^\infty \sqrt{\sup_{\substack{s,t\in[0,T],s<t,\\ |t-s|\leq Tp^{-u^2}}} \int_s^t \sigma^2(v)\check{x}^2(t-v)dv}\, du. \tag{3.4.11}$$

Making use of the (weak) differential of the fundamental solution, we may deduce that

$$\int_0^s \Big(\check{x}(t-v)-\check{x}(s-v)\Big)^2\sigma^2(v)dv \leq \sigma_+^2 \int_0^s \left(\int_s^t L\big(\check{x}(u-v+\theta):\theta\in J\big)du\right)^2 dv$$
$$\text{for all } s,t\in[0,T],\ s<t.$$

Then, through the uniform stability, we obtain from (3.4.3) that

$$L(\check{x}(t+\theta:\theta\in J)) \leq m_L\|\check{x}(t+\theta:\theta\in J)\| \leq m_L K e^{\gamma r}e^{-\gamma t} \quad \text{for all } t\in[0,T].$$

Therefore, we can deduce that

$$\begin{aligned}
\int_0^s \Big(\check{x}(t-v)-\check{x}(s-v)\Big)^2\sigma^2(v)dv &\leq \sigma_+^2K^2e^{2\gamma r}m_L^2\int_0^s\int_s^t e^{-2\gamma(u-v)}du\, dv\\
&\leq \sigma_+^2K^2e^{2\gamma r}m_L^2\int_0^s\int_s^t e^{-2\gamma(u-s)}du\, e^{-2\gamma(s-v)}dv\\
&\leq \sigma_+^2K^2e^{2\gamma r}m_L^2(t-s)^2\int_0^s e^{-2\gamma v}dv\\
&\leq \sigma_+^2\frac{(t-s)^2}{2\gamma}K^2e^{2\gamma r}m_L^2 \quad \text{for all } s,t\in[0,T],\ s<t.
\end{aligned}$$

Furthermore,

$$\int_s^t \check{x}^2(t-v)\sigma^2(v)dv \leq \sigma_+^2K^2(t-s) \quad \text{for all } s,t\in[0,T],\ s<t.$$

Thus, with the help of an auxiliary computation that has been postponed to the appendix, see Theorem A.2, first,

$$\int_0^\infty \varphi\left(\exp\left(-x^2\right)\right)dx \leq \frac{\sigma_+Ke^{\gamma r}m_L}{\sqrt{2\gamma}}\int_0^\infty e^{-u^2}du + \sigma_+K\int_0^\infty e^{-\frac{u^2}{2}}du < \infty,$$

which justifies the application of the Fernique inequality. And, second

$$\mathcal{Q}_1 \leq \frac{\sigma_+Ke^{\gamma r}m_L}{\sqrt{2\gamma}}\int_1^\infty Tp^{-u^2}du \leq \frac{\sigma_+Ke^{\gamma r}m_L}{\sqrt{2\gamma}\,2p\log p}T \quad \text{and} \quad \mathcal{Q}_2 \leq \frac{\sigma_+K\sqrt{T}}{\sqrt{p}\log p}. \tag{3.4.12}$$

And the Fernique inequality thus yields the claim. □

Remark 3.9. *(i) Theorem 3.8 is only useful if* h *is at least of order* $\sqrt{\log p}$*, so that the condition* $h > \sqrt{1 + 4\log p}$ *turns out not to have any seriously restrictive meaning.*

(ii) Given some $h > 0$*, and choosing* $p = T$*, the theorem implies that in order to leave the neighborhood* $[-h, h]$ *with probability of order* 1*, the deviation process must be given at least an exponentially large amout of time with respect to* h*. Or in other words, the process remains in* $[-h, h]$ *for an exponentially long time in* x *with high probability. The other way around, we might say that in order to capture the deviation process for a given time* $T > 0$ *with high probability, it is sufficient to choose a neighborhood of order* $\sqrt{\log T}$*. The result resembles the behavior of solutions of classical linear SDEs subject to additive noise that one may e.g. find in [BG06, Theorem 3.1.6 (Stochastic linear stable case)]. It is worth mentioning that the concentration result is comparable to the according ones from large deviation theory, but, as Berglund and Gentz noted, an estimate in the form of theorem 3.8 is more precise revealing knowledge on the leading prefactors, and its validity is not restricted to asymptotic limits.*

(iii) With regard to the previous item of this list, we will regularly ignor integer-value restrictions in this work. If we choose $p = T$*, for big* T *the result of Theorem 3.8 reads*

$$\mathbb{P}\left\{ \sup_{s\in[0,T]} |y(s)| > h\left(\sqrt{\operatorname{var} y(T)} + \mathcal{O}\left(\frac{1}{\log T}\right)\right)\right\} \le \frac{5}{2}T^2 \exp\left(-\frac{h^2}{2}\right)$$

$$\text{for all } h > \sqrt{1 + 4\log T}.$$

This form does not reveal anything that was imperceptible before, but it appears to be convenient to emphasize the substantial role that the variance process takes. And although, by

$$\operatorname{var} y(t) = \int_0^t \check{x}^2(t-u)\sigma^2(u)du \qquad \text{for all } t \in [0, T],$$

we even have an explicit representation and by (3.4.6) *an upper boundary, how exactly* $\operatorname{var} y(t)$ *depends on the parameters* m_L *and* r *of the SRFDE* (3.4.1)*, is pretty much unknown. In that form, it is easier to compare the concentration result and the essential-growth result in* (3.4.4) *for constant diffusion coefficient. While the essential growth provides that the process exceeds the level* $\sqrt{2\log(t)\operatorname{var} y(t)}$ *only finitely many times, the concentration results contributes quickly vanishing tail probabilities.*

In the next chapter, it will be shown that, when restricting to special cases, potentially much more information can be received about the connection between the variance process and the underlying setting of parameters. And the key tool will be the fundamental solution by means of the variation-of-constants formula.

4. Stochastic Delay Differential Equations

This chapter reduces the generality of Chapter 3 to the simplest possible case in which the most basic one-dimensional ODE is equipped with white noise and a linear feedback through a time-lag term with fixed delay length. Formally, we will consider systems of the form

$$\begin{cases} dx(t) = -ax(t)dt + bx(t-r)dt + \sigma dW(t) & \text{for } t \geq 0, \\ x_0 = \Upsilon, \end{cases} \tag{4.0.1}$$

where a, b and $r > 0$ are fixed real constants, $\Upsilon \in \mathcal{C}(J, \mathbb{R})$, and W is a one-dimensional Brownian motion and $\sigma > 0$ is a constant. Systems of the form (4.0.1) are commonly referred to as stochastic delay differential equations (SDDEs) and the literature so far provides plenty of information on their corresponding deterministic counterparts, the DDEs, as well as on SDDEs. Once again, we refer to the book of Hale, [HVL93], for basic properties of the deterministic system, which especially includes the characterization of parameter combinations a, b and r that lead to stable solutions. Moreover, the work [KM92] by U. Küchler and B. Mensch serves a couple interesting facts, that we will come back to several times.
Like in the general case, the fundamental solution $(\breve{x}(t))_{t\in[-r,\infty)}$ can be manifested with the help of a Volterra resolvent as in (3.3.9), and interpreted as the unique solution of the system (4.0.1) with initial segment $\Upsilon(t) = \mathbb{1}_{\{0\}}(t)$, $t \in [-r, 0]$. But, other than in the general case, existence and uniqueness of the fundamental solution can be achieved much easier through a step-wise procedure, which means that, given the initial segment, basic techniques of ODE theory yield the desired results first on $[0, r]$, then on $[r, 2r]$ and so on. The stepwise approach easily provides the following useful facts on the fundamental solution:

- The fundamental solution $\breve{x}$ is continuous on $(0, \infty)$, continuously differentiable on (r, ∞), twice continuously differentiable on $(2r, \infty)$, ...
- it is right-continuous on $[0, \infty)$ and continuously right-differentiable on $[r, \infty)$, twice continuously right-differentiable on $[2r, \infty)$, ...

It is worth mentioning that in the same way existence and uniqueness of strong solutions of (4.0.1) can be settled stepwise using basic SODE techniques. The deterministic version is the simplest case of what we have seen in Example 3.2.

We formulate the characteristic equation in this case by means of the *characteristic mapping*

$$h : \mathbb{C} \to \mathbb{C}, \quad \lambda \mapsto h(\lambda) := a - \lambda - be^{-\lambda r}, \tag{4.0.2}$$

and let $\mathcal{R}$ denote the roots of the characteristic equation, or the characteristic mapping respectively, i.e.

$$\mathcal{R}(a, b, r) := \left\{\lambda \in \mathbb{C} : a - \lambda - be^{-\lambda r} = 0\right\}. \tag{4.0.3}$$

In this situation a simple characterization of stationary solutions is served by the work of Küchler and Mensch, see [KM92, Proposition 2.8, Proposition 2.11]. It provides the equivalence of the following properties

a) All characteristic roots $\mathcal{R}$, have negative real part.

b) There is a stationary solution.

c) The fundamental solution is square-integrable.

As we have mentioned in section 3.4, a) is equivalent to the assumption that $q_0 := \max\{\Re(\lambda) : \lambda \in \mathcal{R}\} < 0$. In that case, the stationary solution is unique and a version of it is given by

$$U(t) := \int_{-\infty}^{t} \check{x}(t-u)dW(u) \quad \text{for all } t \geq -r.$$

Further, it is generally true that

$$\begin{aligned} br \geq -e^{-ar-1} \quad &\Rightarrow \quad \check{x}(t) > 0 \text{ for all } t \geq 0 \quad \text{and} \quad \lim_{t\to\infty} \frac{1}{t} \log \check{x}(t) = q_0, \\ br < -e^{-ar-1} \quad &\Rightarrow \quad \check{x}(t) \text{ oscillates around } 0 \quad \text{and} \quad \limsup_{t\to\infty} \frac{1}{t}\check{x}(t) = q_0, \end{aligned} \tag{4.0.4}$$

see [KM92, Proposition 3.2]. Note that the roots $\mathcal{R}$ of the characteristic equation can be understood as those $\lambda \in \mathbb{C}$ for which the mapping $f^{(\lambda)} : \mathbb{R} \to \mathbb{R}$, $t \mapsto e^{\lambda t}$, satisfies the differental law $\dot{f}^{(\lambda)}(t) = -af^{(\lambda)}(t) + bf^{(\lambda)}(t-r)$ for all t. This means those $t \mapsto e^{\lambda t}$ for which there is an initial segment $\Upsilon \in \mathcal{C}(J, \mathbb{R})$ such that $f^{(\lambda)}$ solves the deterministic version of (4.0.1) with $f_0^{(\lambda)} = \Upsilon$.

In this simple case, it is actually possible to explicitly characterize the combinations of parameters a, b, r that lead to roots of the respective characteristic equation all having negative reals parts, and therefore, that correspond to asymptotically stable systems of the form (4.0.1). We follow the literature in declaring this parameter-combination set as the *stability area* S. One finds a neat presentation of the details for the differential law $dx(t) = -ax(t) + bx(t-r)$ embedded in the multidimensional case and with an outlook to the nonautonomous case in the book of Hale and Lunel, especially in their appendix on *stability of characteristic equations*, [HVL93, Chapter 5.2, Appendix]. Outlining, we remember that whenever the combination of a, b and r is located in S, we know that for any arbitrarily small $C > 0$ there is $K > 0$ such that the corresponding fundamental solution $(\check{x}(t))_{t\in[-r,\infty)}$ satisfies

$$|\check{x}(t)| \leq Ke^{(q_0+C)t} \quad \text{for all } t \geq 0, \tag{4.0.5}$$

where $q_0 = \max\{\Re(\lambda) : \lambda \in \mathcal{R}\} < 0$. For an illustration of the stability area S see Figure 4. And undoubtedly this classical result provides the answers to most of the questions concerning stability issues, but there are still some points left unregarded. If nothing else, the concentration result of Theorem 3.8 adds two points of interest concerning the fundamental solutions.

- First, what can be said about the behavior of fundamental solutions in case of parameter combinations that lie on the boundary?
- Second, in particular interesting from an applicationer's perspective is the question whether there can anything be said about the constant K that appears in (4.0.5) and the relation to $C > 0$.

In the subsequent section, we will actually confine to systems that correspond to combinations where $b > 0$ and provide detailed answers to both those questions and some of the

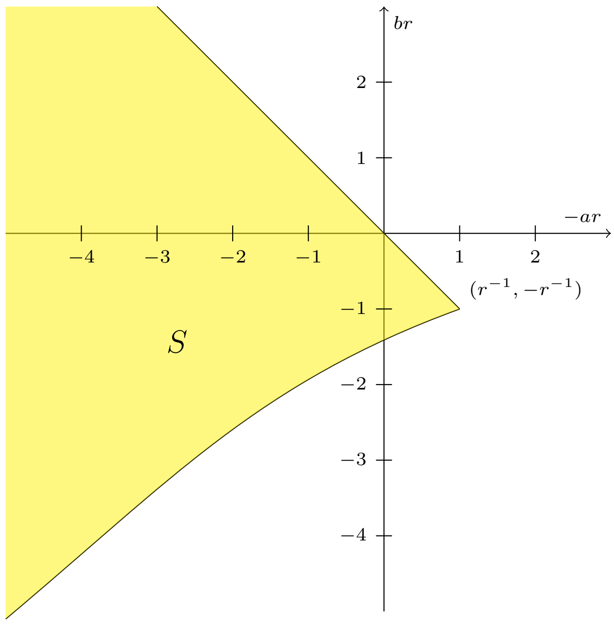

Figure 4: Sketch of the area of stability S. It is common to denote the assumptions in terms of ar and br rather than in a and b and r, see e.g. [HVL93]. Another convenient way is to achieve $r = 1$ by means of time transformation which comes with the advantage to get rid of the symbol r. We favor to always have the time-delay impact explicitly. The lower boundary of S is parametrized by $a = b\cos(\zeta r), -b\sin(\zeta r) = \zeta$ for $\zeta \in (0, \pi/r)$, and the upper boundary is the angle bisector in the second quadrant and a bit of the bisector in the fourth quadrant. The two boundaries meet at (r^{-1}, r^{-1}). It is worth mentioning that in [YB11] one finds beautiful analogue presentations for different delay-feedback mechanisms.

implications.

4.1. Convergence in Critical Regime

In this section we consider delay differential equations with start in some arbitrary $t_0 \in \mathbb{R}$, given by

$$\begin{cases} dx(t) = -ax(t)dt + bx(t-r)dt & \text{for } t \geq t_0, \\ x_{t_0} = \Upsilon, \end{cases} \tag{4.1.1}$$

where $\Upsilon \in \mathcal{C}(J, \mathbb{R})$ and $b > 0$. We will introduce a new approach to the solution properties of systems of the form (4.1.1) that leads to a significantly improved understanding of solutions. In particular we make a noteworthy contribution to the behavior of systems in the *critical regime*, i.e. systems whose parameter combination is nested on the boundary of the stability region S in the second quadrant in Figure 4 which means $a = b > 0$. The outstanding point is that systems in critical regime feature fundamental solutions that converge exponentially fast to $1/(1+ar)$, never exit from the interval $[0, 1]$, and we will provide a minimal convergence rate. Further we will show that this result may be carried over to a class of non-critical systems in case $b > 0$ in a natural way to deduce more concrete estimates on fundamental solutions that correspond to a stable or an unstable regime. The method can in general not be easily generalized to the situation $b < 0$.

To begin with, we remind of a solution representation that can be achieved by an application of the general result (3.3.14) if one understands the influence of the initial segment as an inhomogeinity. The result can for instance be found in [KM92] or again [HVL93, Chapter 1, Theorem 6.1],

$$x(t) = \Upsilon(0)\check{x}(t-t_0) + b\int_{-r}^{0} \check{x}(t-t_0-r-u)\Upsilon(u)du \quad \text{for all } t \geq t_0, \tag{4.1.2}$$

which relies on the fact that the fundamental solution in autonomous case only depends on the difference of the arguments. And let us further remember the classical variation-of-constants formula from ODE theory only for the first segment of length r after t_0. This one is applicable if we interpret the time-delayed term in (4.1.1) as inhomogeneity $f : [t_0, t_0+r) \to \mathbb{R}$, $f(t_0+s) = bx(t_0+s-r)$ for $s \in [0,r)$. The formula tells us that we may write down the explicit solution of $dx(t) = -ax(t)dt + f(t)dt$, $t \in [t_0, t_0+r)$ with start in $x(t_0) = \Upsilon(0)$ as

$$\begin{aligned} x(t_0+s) &= x(t_0)e^{-as} + \int_0^s f(t_0+u)e^{-a(s-u)}du \\ &= x(t_0)e^{-as} + b\int_0^s x(t_0-r+u)e^{-a(s-u)}du \quad \text{for all } s \in [0,r]. \end{aligned} \tag{4.1.3}$$

Of course, this classical variation-of-constants formula can be understood as a special case of (4.1.2) in the case where $b=0$ and therefore, the fundamental solution takes the form e^{-as} for $s \in [0,r)$. The solution representations (4.1.2) and (4.1.3) have been established for an arbitrary starting time t_0. Honestly, we might have spared the effort; we could have started in $t_0 = 0$ to argue in hindsight that it would have been nothing special about starting in 0 compared to starting at any $t_0 \in \mathbb{R}$. But, this way we have evaded all those *woulds*, *mights* and *coulds*. The point is that we have learned all we need about starting solutions of systems (4.1.1) in arbitrary times t_0. In particular, we have the formulas readily prepared for later use. If not surprising, this knowledge is valuable in general and will further play a central role in the proof of the upcoming lemma where we note that the fundamental solution does never leave $[0,1]$. The solution representations will also provide the following seemingly artificial fact: Adding up the fundamental solution at any time t and a-times its integral over the previous segment-length interval $[t-r,t]$ always serves 1, see (4.1.4). And this artificial fact will turn out to be of solid use later on in this section. As the following lemma focuses on properties of the fundamental solution rather than more general solutions, it is convenient to reduce to the case of a start in time 0, because fundamental solutions that start in any other point in time, say t_0, may be regained through a time shift by t_0.

Lemma 4.1. *Let $a = b > 0$ and let $(\check{x}(t))_{t\in[-r,\infty)}$ be the fundamental solution of (4.1.1). Then*

a) $\check{x}(t) \in [0,1]$ for all $t \in [-r,\infty)$ and $\check{x}(t) \in (0,1)$ for all $t \in (0,\infty)$.

b) We can rewrite the fundamental solution as

$$\check{x}(t) = 1 - a\int_{t-r}^{t} \check{x}(u)du \qquad \textit{for all } t \geq 0. \tag{4.1.4}$$

Proof. a) Due to $\check{x}(t) = e^{-at}$ for all $t \in [0,r]$, the claim holds over $[-r,r]$. We define appropriate deterministic stopping times

$$\tau_0 := \inf\{t \geq r : \check{x}(t) = 0\} < \infty \quad \text{and} \quad \tau_1 := \inf\{t \geq r : \check{x}(t) = 1\} < \infty,$$

where $\inf \emptyset = +\infty$, and assume that they are finite. As $\check{x}(r) = e^{-ar}$, we know that $\tau_0 > r$ and $\tau_1 > r$, because the fundamental solution is continuous over $(0, \infty)$. Then, by the classical solution representation (4.1.3), we have that

$$\check{x}(r+s) = \check{x}(r)e^{-as} + a\int_0^s \check{x}(u)e^{-a(s-u)}du \geq \check{x}(r)e^{-as} \quad \text{for all } s \in (0,\infty), s \leq \tau_0 - r,$$

which is a contradiction to $\tau_0 < \infty$. Further, due to the definition of τ_1, clearly

$$\check{x}_{\tau_1}(u) < 1 \quad \text{for all } u \in [-r, 0).$$

Then, with the same representation as above,

$$\begin{aligned}
\check{x}(\tau_1) &= \check{x}(0)e^{-a\tau_1} + a\int_0^{\tau_1} \check{x}(\tau_1 - r + u)e^{-a(\tau_1 - u)}du \\
&< e^{-a\tau_1} + a\int_0^{\tau_1} e^{-a(\tau_1-u)}du \\
&= e^{-a\tau_1} + 1 - e^{-a\tau_1},
\end{aligned}$$

which contradicts $\tau_1 < \infty$.

b) Note that for arbitrary constant $c \in \mathbb{R}$, the constant process $x^{(c)}(t) = c$ for all $t \in [-r, \infty)$ solves (4.1.1) for $\Upsilon = \Upsilon^{(c)}(\cdot) = c$, interpreted as the continuous constant function over J. Let $c = 1$, then an application of the variation-of-constants formula (4.1.2) for $t_0 = 0$ on the known process $x^{(1)}$ reveals the desired knowledge on $(\check{x}(t))_{t \geq -r}$; it reads

$$\begin{aligned}
1 = x^{(1)}(t) &= \Upsilon^{(1)}(0)\check{x}(t) + b\int_{-r}^0 \check{x}(t-r-u)\Upsilon^{(1)}(u)du \\
&= \check{x}(t) + b\int_{-r}^0 \check{x}(t-r-u)du \quad \text{for all } t \geq 0;
\end{aligned}$$

and substituting $v = t - r - u$, we find that

$$1 = \check{x}(t) + b\int_{t-r}^t \check{x}(v)dv \quad \text{for all } t \geq 0,$$

which is the claim. □

Remark 4.2.

- *The lower bound in part a) is covered by* (4.0.4)*, but we will need an analogue of the result in nonautonomous case in the second part of this work. This is why a direct proof is advantageous. Part a) will be generalized to the case* $a \geq b > 0$ *in Lemma 4.16.*

- *An alternate proof of part b) of the above lemma can be obtained in a very simple way* *through the integrated version of differential law:*

$$\check{x}(t) = 1 - a\int_0^t \check{x}(u)du + b\int_0^t \check{x}(u-r)du = 1 - a\int_0^t \check{x}(u)du + b\int_{-r}^{t-r} \check{x}(u)du$$
$$= 1 - a\int_{t-r}^{t} \check{x}(u)du,$$

because $\check{x}$ is zero over $[-r,0)$. Part b) crucially relies on the equality $a=b$, and does not apply in more general settings.

The next lemma shows that a fundamental solution crosses the niveau $(1+ar)^{-1}$ at least once in every interval $(t,t+r)$ for all $t \geq 0$.

Lemma 4.3. *Let $a=b>0$ and let $(\check{x}(t))_{t\in[-r,\infty)}$ be the fundamental solution of* (4.0.1). *Then,*

$$\text{for all } t \geq 0 \text{ there is } t^\star \in (t-r,t): \ \check{x}(t^\star) = \frac{1}{1+ar}. \tag{4.1.5}$$

Proof. First, we show that the claim holds for the interval $(0,r)$. This is due to the two facts that $\check{x}(0) = 1 > (1+ar)^{-1}$ and that $\check{x}(r) = e^{-ar}$. The Taylor expansion of the exponential yields

$$e^{ar} = 1 + ar + \sum_{k=2}^{\infty} \frac{(ar)^k}{k!} > 1 + ar,$$

which is equivalent to say that $e^{-ar} < \frac{1}{1+ar}$. Then, by the mean-value theorem there is $t^\star \in (0,r)$ with $\check{x}(t^\star) = (1+ar)^{-1}$. The rest of the assertion (4.1.5) is shown by contradiction. We suppose that there is $\bar{t} \geq r$ such that

$$\check{x}(v) < \frac{1}{1+ar} \qquad \text{for all } v \in (\bar{t}-r,\bar{t}). \tag{4.1.6}$$

Then rewriting $(\check{x}(t))_{t\in[-r,\infty)}$ as in part (ii) of Lemma 4.1,

$$\check{x}(\bar{t}) = 1 - a\int_{\bar{t}-r}^{\bar{t}} \check{x}(v)dv > 1 - a\int_{\bar{t}-r}^{\bar{t}} \frac{1}{1+ar}dv = 1 - \frac{ar}{1+ar} = \frac{1}{1+ar}.$$

By continuity in $\bar{t}$, this is a contradiction to (4.1.6). Then mainly repeating the above arguments, we assume that there is $\tilde{t} \geq r$ such that

$$\check{x}(v) > \frac{1}{1+ar} \qquad \text{for all } v \in (\tilde{t}-r,\tilde{t}), \tag{4.1.7}$$

implying that

$$\check{x}(\tilde{t}) = 1 - a\int_{\tilde{t}-r}^{\tilde{t}} \check{x}(v)dv < 1 - a\int_{\tilde{t}-r}^{\tilde{t}} \frac{1}{1+ar}dv = 1 - \frac{ar}{1+ar} = \frac{1}{1+ar},$$

which contradicts assumption (4.1.7). And so assertion (4.1.5) is proved. □

The following lemma provides that the time points, where the fundamental solution crosses $(1+ar)^{-1}$, do not have an accumulation point. Let us remark that there is a quick proof based on the fact that difference $\check{x}(t) - (1+ar)^{-1}$ constitutes an *entire* function and therefore cannot have infinitely many zeros over any finite interval. But unwilligly to push open the door to an odyssey upon introducing and veryfying the involved concepts, we bring up a simpler argument that basically relies on the fact that the fundamental solution $\check{x}$ may be represented in one more particular way (see e.g. [KM92]), and two classical results from analysis and algebra. The analytical result is Rolle's Theorem from 1691 (see [KÖ04], [Rol90]) which states that for any continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$ satisfying $f(x_1) = f(x_3)$ for some real $x_1 < x_3$ there must be $x_2 \in (x_1, x_3)$ such that $f'(x_2) = 0$. And the algebraic result is the fundamental theorem of Algebra[1] that implies that a real-valued polynomial of degree $k \in \mathbb{N}$ must not have more than k roots over the reals. We will apply one further representation of the fundamental solution with start in time 0. This one has been seen by [KM92]; compared to the variation-of-constants representation it is bulky, but serves our needs here perfectly.

$$\begin{aligned} \check{x}(t) &= \sum_{k=0}^{K} \frac{b^k}{k!} e^{-a(t-kr)} (t-kr)^k \\ &= e^{-at} \sum_{k=0}^{K} \frac{(be^{ar})^k}{k!} (t-kr)^k \quad \text{for all } t \in [Kr, (K+1)r),\ K \in \mathbb{N} \cup \{0\}. \end{aligned} \tag{4.1.8}$$

Besides the notational burden a formal proof by induction is straightforward.

Lemma 4.4. *Let $a = b > 0$ and let $(\check{x}(t))_{t \in [-r,\infty)}$ be the fundamental solution of (4.1.1) with $t_0 = 0$. Let furthermore $\mathcal{Z}$ be the set of zeros of $\check{x}(t) - (1+ar)^{-1}$, $t \in [0,\infty)$. Then $\mathcal{Z}$ can be written as*

$$\mathcal{Z} = \{t_1^\star, t_2^\star, \ldots\} \quad \text{where } t_1^\star < t_2^\star < t_3^\star < \ldots \quad \text{and} \quad \sum_{i \in \mathbb{N}} t_{i+1}^\star - t_i^\star = \infty. \tag{4.1.9}$$

In other words: There is no accumulation point in $\mathcal{Z}$.

Proof. Rewriting the fundamental solution in the form (4.1.8) reveals $(\check{x}(t))_{t \in [Kr,(K+1)r)}$ to be a ***poly-exponential*** function of degree K over the each interval $[Kr, (K+1)r)$, which means we recognize $(\check{x}_t)_{t \in [-r,\infty)}$ locally, i.e. for $t \in [Kr, (K+1)r)$ for $K \in \mathbb{N}$, as product of a polynomial $P(\cdot)$ of degree K and the exponential function $t \mapsto \exp(-at)$, i.e.

$$\begin{aligned} \check{x}_t &= P(t) \exp(-at) \quad \text{for all } t \in [Kr, (K+1)r), \quad \text{where} \\ P(t) &= \sum_{k=0}^{K} \frac{(be^{ar})^k}{k!} (t-kr)^k \quad \text{for all } t \in [Kr, (K+1)r). \end{aligned} \tag{4.1.10}$$

Let $y : [0,\infty) \to \mathbb{R}$, $y(t) := \check{x}(t) - (1+ar)^{-1}$, denote the difference between fundamental solution and proclaimed limit. The mapping y is continuously differentiable on (r,∞) and has continuous right-hand derivative in r. Let $I \subset [r,\infty)$ denote an arbitrary finite open

[1] The first references of the fundamental theorem of algebra go back to Peter Roth 1608 in *Arithmetica Philosophica*, see [Man06], or Albert Girard's *L'invention nouvelle en l'Algèbre*, [Gir29], in 1629. The first conceptually correct proof was – to the best of our knowledge – published 1746 by Jean d'Alembert even if the proof beared weaknesses. Other noteworthy contributions have been established by Karl Friedrich Gauss in 1815, [Gau15], or Karl Weierstrass in 1891, [Wei82].

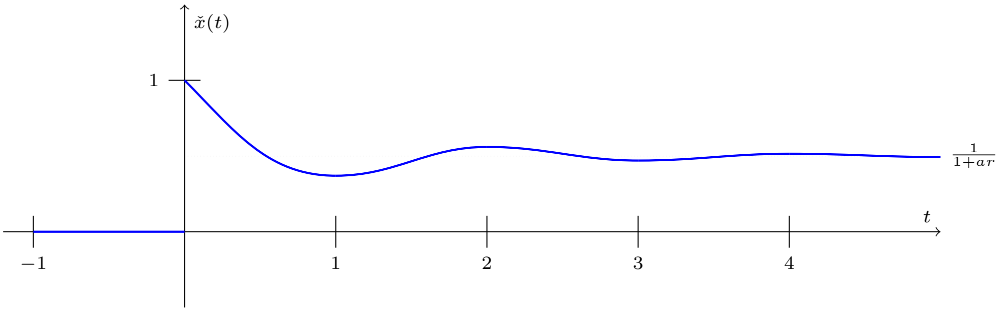

Figure 5: Sketch of the behavior of a fundamental solution for $a = b = r = 1$ to visualize the convergence of fundamental solution in critical regime.

interval and let

$$\mathcal{Z}_{\overline{I}}(y) \text{ denote the set of zeros of } y \text{ on the closure of } I,$$
$$\mathcal{Z}_{I}(\dot{y}) \text{ denote the set of zeros of } \dot{y} \text{ on } I.$$

Then by Rolle's theorem, between any two neighboring zeros of y there must be a zero of $\dot{y}$. Hence,

$$|\mathcal{Z}_{\overline{I}}(y)| \leq 1 + |\mathcal{Z}_{I}(\dot{y})|, \tag{4.1.11}$$

But then, (4.1.10) provides that

$$\begin{aligned} \dot{y}(t) &= \exp(-at)\overline{P}(t) \quad \text{ for all } t \in (Kr, (K+1)r), \text{ where} \\ \overline{P}(t) &= \left(-aP(t) + \frac{dP(t)}{dt}\right) \quad \text{ for all } t \in (Kr, (K+1)r), \end{aligned}$$

revealing $\overline{P}$ as polynomial of order K. So by the Fundamental Theorem of Algebra, $\dot{y}$ can have at most K zeros. And with (4.1.11) we see that $(y(t))_{t\geq 0}$ can have at most $K+1$ zeros over $[Kr, (K+1)r]$. That finishes the proof. □

We are now ready to prove the main result of this section:

Theorem 4.5 (Convergence of Fundamental Solutions in Critical Regime)**.** *Let $a = b > 0$ and $(\check{x}(t))_{t\in[-r,\infty)}$ the fundamental solution of* (4.1.1)*. Then for*

$$\kappa = \frac{|\log(1 - e^{-ar})|}{2r} \tag{4.1.12}$$

the following estimate holds true:

$$\left|\check{x}(t) - \frac{1}{1+ar}\right| \leq e^{-\kappa t} \textit{ for all } t \geq 0. \tag{4.1.13}$$

Proof. Let $t_1^\star < t_2^\star < \ldots$ denote the zeros of $\check{x}(\cdot) - (1+ar)^{-1}$. We fix an arbitrary $k \in \mathbb{N}$ with $t_k^\star \geq r$ as in Lemma 4.3. By continuity, there is some bound on the distance between the fundamental solution and the proclaimed limit $(1+ar)^{-1}$ over the segment prior to $t_k^\star$

given by some constant $C \in (0, \infty)$; formally

$$\left| \breve{x}(t_k^\star - r + s) - \frac{1}{1+ar} \right| \le C \quad \text{for all } s \in [0, r]. \tag{4.1.14}$$

We use the classical-theory representation (4.1.3) with $t_0 = t_k^\star$. Therewith we may deduce a helpful upper-bound estimate for the first interval of length r after time point $t_k^\star$:

$$\breve{x}(t_k^\star + s) = \frac{1}{1+ar} e^{-as} + b e^{-as} \int_0^s e^{au} \breve{x}(t_k^\star - r + u) du \quad \text{for all } s \in [0, r].$$

And therefore, using that $\breve{x}(t_k^\star - r + u) \le (1 + br)^{-1} + C$ yields

$$\begin{aligned} \breve{x}(t_k^\star + s) &\le \frac{1}{1+ar} e^{-as} + e^{-as} \frac{b}{a} \Big(e^{as} - 1 \Big) \left(\frac{1}{1+ar} + C \right) \\ &= \frac{1}{1+ar} e^{-as} + \frac{1 - e^{-as}}{1+ar} + C\big(1 - e^{-as}\big) = \frac{1}{1+ar} + C\big(1 - e^{-as}\big) \quad \text{for all } s \in [0, r]. \end{aligned}$$

So we arrive at the following upper-bound estimate for $\breve{x}(t_k^\star + s)$ for $s \in [0, r]$:

$$\breve{x}(t_k^\star + s) \le \frac{1}{1+ar} + C\big(1 - e^{-ar}\big) \quad \text{for all } s \in [0, r]. \tag{4.1.15}$$

In the same way we may use that $\breve{x}(t_k^\star - r + u) \ge (1 + ar)^{-1} - C$ to obtain

$$\breve{x}(t_k^\star + s) \ge \frac{1}{1+ar} - C\big(1 - e^{-ar}\big) \quad \text{for all } s \in [0, r]. \tag{4.1.16}$$

Combining (4.1.15) and (4.1.16) yields

$$\left| \breve{x}(t_k^\star + s) - \frac{1}{1+ar} \right| \le C(1 - e^{-ar}) \quad \text{for all } s \in [0, r]. \tag{4.1.17}$$

By lemma 4.3 there is $t_{k+1}^\star \in (t_k^\star, t_k^\star + r)$ with $\breve{x}(t_{k+1}^\star) = \frac{1}{1+ar}$, and with the above shown behavior of $\breve{x}$ over $[t_k^\star, t_k^\star + r]$, we may conclude that

$$\left| \breve{x}(t_{k+1}^\star - r + s) - \frac{1}{1+ar} \right| \le C \quad \text{for all } s \in [0, r].$$

We apply the above argument leading to (4.1.17) once more on $t_{k+1}^\star$ instead of $t_k^\star$ to find that

$$\left| \breve{x}(t_{k+1}^\star + s) - \frac{1}{1+ar} \right| \le C\big(1 - e^{-ar}\big) \quad \text{for all } s \in [0, r]$$

which includingly serves an improvement of the estimate (4.1.17) regarding the interval on which it is valid. We observe that

$$\left| \breve{x}(t_k^\star + s) - \frac{1}{1+ar} \right| \le C\big(1 - e^{-ar}\big) \quad \text{for all } s \in [0, t_{k+1}^\star - t_k^\star + r). \tag{4.1.18}$$

Then the picking of those zeros $t_k^\star$, $t_{k+1}^\star$, $t_{k+2}^\star$, $\ldots$ can be iterated and we can deduce

$$\left| \breve{x}(t_k^\star + s) - \frac{1}{1+ar} \right| < C(1 - e^{-ar}) \quad \text{for all } s \in [0, t_{k+1}^\star - t_k^\star + t_{k+2}^\star - t_{k+1}^\star + \ldots).$$

To extend the estimate's validity over the half line starting in $t_k^\star$, it is sufficient that the distance between neighboring zeros of $\check{x}(t) - (1+ar)^{-1}$ is not summable which is granted through Lemma 4.4. And in particular, there is $l \in \mathbb{N}$ such that $t_{k+l}^\star \in (t_k^\star + r, t_k^\star + 2r)$ with $\check{x}(t_{k+l}^\star) - (1+ar)^{-1} = 0$ and

$$\left|\check{x}(t_{k+l}^\star - r + s) - \frac{1}{1+ar}\right| \leq C\big(1 - e^{-ar}\big) \quad \text{for all } s \in [0, r]$$

and with the same reasoning as before, we find that

$$\left|\check{x}(t_{k+l}^\star + s) - \frac{1}{1+ar}\right| \leq C\big(1 - e^{-ar}\big)^2 \quad \text{for all } s \in [0, \infty),$$

which in terms of $t_k^\star$ implies that

$$\left|\check{x}(t_k^\star + 2r + s) - \frac{1}{1+ar}\right| \leq C\big(1 - e^{-ar}\big)^2 \quad \text{for all } s \in [0, \infty).$$

And an iteration of the whole argument yields for arbitrary $n \in \mathbb{N}$:

$$\left|\check{x}(t_k^\star + 2nr + s) - \frac{1}{1+ar}\right| \leq C\big(1 - e^{-ar}\big)^{n+1} \quad \text{for all } s \in [0, \infty).$$

This will play the role of an induction step to deduce the theorem's assertion. Now, we will work out the induction start which means the behavior of $\check{x}(t) - (1+ar)^{-1}$, $t \in [0, r]$, on the first interval of length r. To begin with, we note that

$$\left|\check{x}(t) - \frac{1}{1+ar}\right| = \left|e^{-at} - \frac{1}{1+ar}\right| < 1 - e^{-ar} \quad \text{for all } t \in [0, r].$$

This can be seen by observing $\check{x}(t) - (1+ar)^{-1}$ at the end points of the interval $[0, r]$; the rest follows through monotonicity. First,

$$\check{x}(0) - \frac{1}{1+ar} < 1 - e^{-ar} \quad \Leftrightarrow \quad \frac{1}{1+ar} > e^{-ar} \quad \Leftrightarrow \quad e^{ar} > 1 + ar$$

which is true as we have already seen in the proof of Lemma 4.3; and second,

$$\frac{1}{1+ar} - \check{x}(r) < 1 - e^{-ar} \quad \Leftrightarrow \quad \frac{1}{1+ar} < 1.$$

Further, we know that there is a zero of $\check{x} - (1+ar)^{-1}$ in $(0, r)$ which, by the arguments that we have seen in the first part of the induction step, ensures that the estimate $|\check{x}(t) - (1+ar)^{-1}| \leq 1 - e^{-ar}$ for all $t \in [0, r]$ can be extended to

$$\left|\check{x}(t) - \frac{1}{1+ar}\right| \leq 1 - e^{-ar} \quad \text{for all } t \in [0, \infty). \tag{4.1.19}$$

Therefore, there is $t_m^\star \in (r, 2r)$ with $|\check{x}(t_m^\star + s) - (1+ar)^{-1}| \leq (1 - e^{-ar})$ for all $s \in [-r, 0]$. The usual procedure implies that

$$\left|\check{x}(2r + s) - \frac{1}{1+ar}\right| \leq (1 - e^{-ar})^2 \quad \text{for all } s \in [0, \infty),$$

and iteration due to (4.1.19) yields that

$$\left|\check{x}(2nr+s) - \frac{1}{1+ar}\right| \leq (1-e^{-ar})^{n+1} \quad \text{each for all } s \in [0,\infty) \text{ and } n \in \mathbb{N} \cup \{0\}. \tag{4.1.20}$$

By some simple computations, we see that

$$e^{-\kappa 2r} = 1 - e^{-ar} \quad \Leftrightarrow \quad \kappa = \frac{|\log(1-e^{-ar})|}{2r}.$$

And therefore,

$$e^{-\kappa t} \geq (1-e^{-ar})^{n+1} \quad \text{for all } t \in [2nr, 2(n+1)r) \quad \text{for all } n \in \mathbb{N} \cup \{0\}. \tag{4.1.21}$$

Then, we may conclude from (4.1.20) and (4.1.21) the assertion of the theorem. □

Remark 4.6. *This convergence of fundamental solutions can in principle be achieved in much more general situations, if we assume that the the very solution of the characteristic equation, that features* 0 *real part, is unique and real. The asymptotic limit of convergence has been achieved in [ARS13], [DvGVLW95].*

Due to the central role that fundamental solutions play, there are a couple of informative consequences easily available. The first thing we mention is that the corresponding DDE in critical regime is stable, but not asymptotically stable. This is specified in the next corollary, which provides an exact formula for the limit of the solution initiated at some arbitrary $\Upsilon \in \mathcal{C}(J,\mathbb{R})$, together with its rate of convergence.

Corollary 4.7. *Let* $a = b > 0$ *and assume arbitrary* $\Upsilon \in \mathcal{C}(J,\mathbb{R})$*. Let* $\check{x}$ *denote the fundamental solution of* (4.1.1)*. Then the solution* $(x(t))_{t\in[-r,\infty)}$ *of* (4.1.1) *with initial segment* Υ *converges exponentially fast. If we let* κ *be given by* (4.1.12)

$$x(t) = \frac{1}{1+ar}\left(\Upsilon(0) + b\int_{-r}^{0} \Upsilon(s)\,ds\right) + R(t) \quad \textit{for all } t \geq 0,$$

where the remainder term $R(t)$ *is bounded in absolute value by* $(|\Upsilon(0)| + r\|\Upsilon\|e^{\kappa r})e^{-\kappa t}$ *for all* $t \geq r$*. Actually, we have that*

$$\left|\int_{-r}^{0} \check{x}(t-r-u)\Upsilon(u)\,du - \frac{1}{1+ar}\int_{-r}^{0}\Upsilon(u)\,du\right| \leq r\|\Upsilon\|e^{-\kappa(t-r)} \quad \textit{for all } t \geq r.$$

Proof. Regarding the upper-bound estimate of the corollary, we easily see that

$$\begin{aligned} d(t) &:= \left|\int_{-r}^{0} \check{x}(t-r-u)\Upsilon(u)\,du - \frac{1}{1+ar}\int_{-r}^{0}\Upsilon(u)du\right| \\ &\leq \int_{-r}^{0}\left|\left(\check{x}(t-r-u) - \frac{1}{1+ar}\right)\Upsilon(u)\right|\,du \\ &\leq \int_{-r}^{0}\sup_{v\in J}\left|\check{x}(t-r-v) - \frac{1}{1+ar}\right|\sup_{w\in J}|\Upsilon(w)|\,du \quad \text{for all } t \geq 0. \end{aligned}$$

Then, an application of the convergence result from Theorem 4.5 yields

$$d(t) \leq r\|\Upsilon\|e^{-\kappa(t-r)} \quad \text{for all } t \geq r.$$

Due to the variation-of-constants formula and another application of the Theorem 4.5, we obtain

$$\begin{aligned} x(t) &= \Upsilon(0)\check{x}(t) + \int_{-r}^{0} \check{x}(t-r-u)\Upsilon(u)\,du \\ &= \frac{1}{1+ar}\left(\Upsilon(0) + \int_{-r}^{0} \Upsilon(u)\,du\right) + R(t) \qquad \text{for all } t \geq 0, \end{aligned}$$

where $R(t)$ is of the form that we claimed. □

The two convergence results, Theorem 4.5 and its Corollary 4.7 introduce essentially new insights to the pathwise behavior of solutions of DDEs in a critical regime. Of course, the two results are more or less the two sides of the same coin, and when speaking of the convergence in critical regime, we refer to either the formulation of the theorem or the corollary. But so far, the convergence result only contributes insights when the examined system is in critical regime meaning that its application field might appear too much restricted to be of use. The following section will invalidate this objection by showing that through a simple transformation the convergence result may be used to reveal better insight to pathwise behavior for (4.1.1) whenever $b > 0$.

4.2. Consequences of Convergence

Carrying over the previously presented convergence result will lead us to a unifying picture for DDEs of the form

$$\begin{cases} dx(t) = -ax(t)dt + bx(t-r)dt & \text{for } t \geq 0, \\ x_0 = \Upsilon, \end{cases} \tag{4.2.1}$$

where $\Upsilon \in \mathcal{C}(J,\mathbb{R})$ and $b > 0$. We have already seen that the fundamental solutions can provide a key in deriving results on solutions initiated with an arbitrary segment from $\mathcal{C}(J,\mathbb{R})$. We start by gathering information on fundamental solutions in case $b > 0$ including $a \neq b$. After that we will focus on the consequences for general solutions with arbitrary continuous initial segment. Always regarding the time lag $r > 0$ as arbitrary but fixed constant, we continue to only speak about combinations of a and b rather than taking r into account. The next lemma captures the fact that an exponentially blown up (or shrinked down) fundamental solution is still a fundamental solution, but with respect to a different pair of underlying parameters.

Lemma 4.8. *For fixed $r > 0$ and arbitrary $a_0, b_0, \lambda \in \mathbb{R}$ let $(\check{x}(t))_{t\in[-r,\infty)}$ denote the fundamental solution of (4.2.1) for $a = a_0$, $b = b_0$ (and r), and let $(y(t))_{t\in[-r,\infty)}$ be defined by $y(t) := e^{\lambda t}\check{x}(t)$ for all $t \in [-r,\infty)$. Then $(y(t))_{t\in[-r,\infty)}$ is the fundamental solution of (4.2.1) for $a = a_0 - \lambda =: \tilde{a}$ and $b = b_0 e^{\lambda r} =: \tilde{b}$ (and r).*

Proof. To check that $(y(t))_{t\in[-r,\infty)}$ solves (4.2.1) with $a = \tilde{a}$ and $b = \tilde{b}$, we first observe that the initial-segment condition trivially holds, because

$$y(t) = \exp(\lambda t)\check{x}(t) = 0 \text{ for all } t \in [-r, 0) \quad \text{and} \quad y(0) = \exp(0)\check{x}(0) = 1.$$

And the differential law is easily verified by the product rule,

$$\begin{aligned} d\big(e^{\lambda t}\check{x}(t)\big) &= \lambda e^{\lambda t}\check{x}(t)dt - a_0 e^{\lambda t}\check{x}(t)dt + b_0 e^{\lambda t}\check{x}(t-r)dt \\ &= -(a_0-\lambda)y(t)dt + b_0 e^{\lambda r} y(t-r)dt \qquad \text{for all } t>0. \end{aligned}$$

□

Regarding the critical regime $a_0 = b_0 > 0$, we have achieved detailed knowledge on the fundamental solutions. That raises the question for which combinations $\tilde{a}$ and $\tilde{b}$ there actually is a *real* λ such that – by only an exponential blow up or shrink down – we can go back to the critical-regime world to make use of the improved knowledge about fundamental solutions. In general the answer is the following: There exists such a real λ if the parameter combination $(\tilde{a}, \tilde{b})$ rests in

$$\mathcal{P}_0 := \{(a,b) \in \mathbb{R}^2 : br \geq -e^{-(1+ar)}\}. \tag{4.2.2}$$

In particular, it is always possible, if $b > 0$. The verification is contained in the next lemma.

Lemma 4.9. *Let $r > 0$ and $(a,b) \in \mathcal{P}_0$, defined in* (4.2.2)*, and let $(\check{x}(t))_{t\in[-r,\infty)}$ be the fundamental solution of* (4.2.1)*. Then there is a real λ such that $a - \lambda = be^{\lambda r}$.*

Proof. A simple reformulation yields

$$a - \lambda = be^{\lambda r} \quad \Leftrightarrow \quad ar - \lambda r = bre^{\lambda r - ar}e^{ar} \quad \Leftrightarrow \quad (ar-\lambda r)e^{ar-\lambda r} = bre^{ar}.$$

Therefore, $ar - \lambda r$ is nothing but the inverse of $x \mapsto xe^x$ evaluated at bre^{ar}. This inverse is know as the *Lambert's function* $\mathcal{W}$, and it is a well-known fact[2] that it takes real values over and only over $[-1/e, \infty)$. Therefore,

$$\lambda = \frac{\mathcal{W}(rbe^{ar})}{r} + a \in \mathbb{R} \quad \Leftrightarrow \quad rbe^{ar} \geq -e^{-1} \quad \Leftrightarrow \quad br \geq -e^{-(1+ar)}. \tag{4.2.3}$$

□

The combination of Lemma 4.8 and Lemma 4.9 may be ellegantly reformulated in terms of a correspondence between sets of fundamental solutions, which we state as the corollary below. We abbreviate the word *fundamental solution* for a moment as *f.s.*.

Corollary 4.10. *Let $\mathcal{P}_0$ be given as in* (4.2.2) *and define*

$$\mathbb{F}_0 := \{\check{x} : [-r,\infty) \to \mathbb{R} : \ \check{x} \text{ is the f.s. of (4.2.1) for some } (a,b) \in \mathcal{P}_0\}, \tag{4.2.4}$$

$$\mathbb{F}_c := \{\check{x} : [-r,\infty) \to \mathbb{R} : \ \check{x} \text{ is the f.s. of (4.2.1) for some } (a,b) \in \mathcal{P}_c\}, \tag{4.2.5}$$

$$\text{where } \mathcal{P}_c := \{(a,b) \in \mathbb{R}^2 : a = b\} \cap \mathcal{P}_0. \tag{4.2.6}$$

[2] Due to [SMM06] the Lambert's function was first introduced by Johann Heinrich Lambert (1728-1777) and its magnificent value in physics, in particular electrostatics, statistical mechanics (see [Cai03]), general relativity and quantum chromo dynamics, see e.g. [CGHJ93, AKM05, Cra07], just to name a few, has emerged only in the second half of the 20th century. To some extend its establishment as a *standard function* has come in 1993 through the publication of Scott, see [SBDM93]; up to then it had been kind of unknown to the literature.

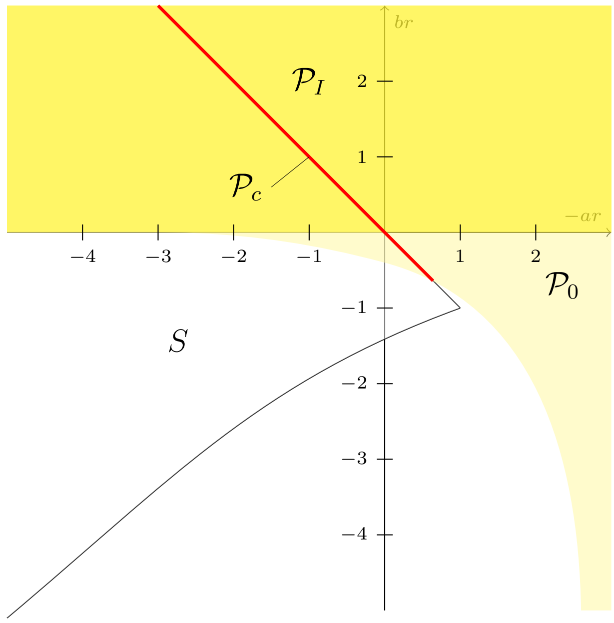

Figure 6: The stability area S has been sketched just for easier orientation and to see the coherence. The area $\mathcal{P}_0$ is represented by the yellow-colored and the pale-yellow-colored area. The solid-yellow-colored area $\mathcal{P}_I$ consists of the first and the fourth quadrant and represents the parameter combinations on which our new method is applicable. It forms a subset of $\mathcal{P}_0$. The set $\mathcal{P}_c$ is the angle bisector through the second and fourth quadrant within $\mathcal{P}_0$, and it is represented through the bold red line segment.

Then, we have that

$$\mathbb{F}_0 = \big\{y : [-r,\infty) \to \mathbb{R} : \exists\ \check{x} \in \mathbb{F}_c,\ \lambda \in \mathbb{R} :\ y(t) = e^{\lambda t}\check{x}(t) \text{ for all } t \in [-r,\infty)\big\}. \quad (4.2.7)$$

The convergence results in critial case are restricted to $b > 0$. And at this point the nonnegativity assumption draws us back to only consider systems of the form (4.2.1) with nonnegative coefficients b, because the multiplication with an exponential function cannot alter the sign when reducing to real exponents. Short-hand references are introduced by the following definition.

Definition 4.11. *We define the set of parameter combinations $\mathcal{P}_I \subset \mathbb{R}^2$ and the set of corresponding fundamental solutions by*

$$\mathcal{P}_I := \{(a,b) \in \mathbb{R}^2 : b > 0\}, \quad (4.2.8)$$

$$\mathbb{F}_I := \{\check{x} : [-r,\infty) \to \mathbb{R} :\ \check{x} \text{ is the f. s. of (4.2.1) for some } (a,b) \in \mathcal{P}_I\}, \quad (4.2.9)$$

$$\mathbb{F}_{I,c} := \{\check{x} : [-r,\infty) \to \mathbb{R} :\ \check{x} \text{ is the f. s. of (4.2.1) for some } (a,b) \in \mathcal{P}_I \cap \mathcal{P}_c\}. \quad (4.2.10)$$

The letter I in the subindex is meant to refer to investigatability by means of the combination of convergence in critical regime and exponential transformation.

Restricted to combinations with $b > 0$, the set $\mathbb{F}_I$ features the same correspondence of fundamental solutions that Corollary 4.10 gives for $\mathbb{F}_0$ in the unrestricted case.

Lemma 4.12. *For arbitrary $a \in \mathbb{R}$ and $b > 0$, there is always a real λ such that $a - \lambda = be^{\lambda r}$; in other words $\mathcal{P}_I \subset \mathcal{P}_0$. Furthermore, the set $\mathbb{F}_I$ is invariant under the exponential blow up or shrink down:*

$$\mathbb{F}_I = \{y : [-r, \infty) \to \mathbb{R} : y(t) = \exp(\lambda t)\check{x}(t)\ \forall\, t \in [-r, \infty) \textit{ for some } \lambda \in \mathbb{R} \textit{ and } \check{x} \in \mathbb{F}_{I,c}\}.$$

Proof. In fact all the statements are trivial with regard to the characterization (4.2.3) in the proof of Lemma 4.9. □

Remark 4.13. *The further analysis of this section confines to the case $a = b > 0$.*

How we actually can learn about the fundamental solution corresponding to a parameter combination $(a, b) \in \mathcal{P}_I$ is clear now - first work out the right exponent λ for a transformation $\check{y}(t) = e^{\lambda t}\check{x}(t),\ t \in [-r, \infty)$ onto a fundamental solution $\check{y}$ corresponding to $\tilde{a} = a - \lambda$ and $\tilde{b} = be^{\lambda r}$ in $\mathcal{P}_I \cap \mathcal{P}_c$, second, apply the convergence result, and third, transform $\check{y}$ back onto $\check{x}$ by $\check{x}(t) = e^{-\lambda t}\check{y}(t)$, $t \in [-r, \infty)$. *What* we actually learn about such a fundamental solution is stated in the next theorem. Furthermore, it is obvious for which combinations of a and b the *right* exponent λ is positive or negative; we also store that point in the theorem for easy later reference.

Theorem 4.14. *Let $(a, b) \in \mathcal{P}_I$ and let λ be the real solution of the characteristic equation $\tilde{a} = a - \lambda = be^{\lambda r} = \tilde{b}$.*

a) The fundamental solution $\check{x}$ with respect to a and b satisfies

$$\left|\check{x}(t) - \frac{e^{-\lambda t}}{1 + \tilde{a}r}\right| \le e^{-(\lambda + \tilde{\kappa})t} \quad \textit{for all } t \ge 0,$$

where $\tilde{\kappa}$ is given by (4.1.12), but with respect to $\tilde{a}$, i.e. $\tilde{\kappa} = \frac{|\log(1 - e^{-\tilde{a}r})|}{2r}$.

b) The solution of the characteristic equation $a - \lambda = be^{\lambda r}$ is negative if $a < b$, and positive if $a > b$.

Proof. To prove part a), one simply multiplies the inquality by $e^{\lambda t}$ to end up with the convergence result in Theorem 4.5 for $\tilde{a} = \tilde{b} > 0$ and $\check{x}(t)e^{\lambda t} = \check{y}(t)$ and $(\check{y}(t))_{t \in [-r, \infty)}$ being the fundamental solution with respect to $\tilde{a}$ and $\tilde{b}$ by Lemma 4.8. Part b) is obvious. □

4.3. Concentration Results for SDDEs

The previous section has provided some surprisingly exact results concerning the behavior of fundamental solutions for SDDEs of the form

$$\begin{cases} dx(t) = -ax(t)dt + bx(t - r)dt + \sigma e^{\mu t}dW(t) & \text{for } t \ge 0, \\ x_0 = \Upsilon, \end{cases} \tag{4.3.1}$$

where $\Upsilon \in \mathcal{C}(J, \mathbb{R})$. One may well ask why it is reasonable to examine the rather artificial diffusion coefficients, and the explanation is that by (4.3.1) we have defined a class of SRFDEs that is closed under the transformation of solution $(x(t))_{t \in [-r, \infty)}$ by multiplication with $e^{ct}, t \in [-r, \infty)$ for some arbitrary $c \in \mathbb{R}$. That means that if $(x(t))_{t \in [-r, \infty)}$ is a solution of (4.3.1) for $a = a_0,\ b = b_0,\ \mu = \mu_0$ and $\Upsilon = \Upsilon_0$, then $y(t) := e^{ct}x(t),\ t \in [-r, \infty)$ solves (4.3.1) for coefficients $a = a_0 - c,\ b = b_0 e^{cr},\ \mu = \mu_0 + c$ and $\Upsilon(u) = e^{cu}\Upsilon_0(u)$ for all $u \in J$, and the new initial segment for $(y(t))_{t\in[-r,\infty)}$ is also in $\mathcal{C}(J,\mathbb{R})$. Then, $(y(t))_{t\in[-r,\infty)}$ is itself a solution to a system of this class. Regarding how we achieved additional information about systems refering to parameter combinations in $\mathcal{P}_I$ in the previous section, that choice is merely natural. The solution representation that we have developed in (3.3.28) may be translated to this simple setting. Remember that $\check{x}(t,u) = \check{x}(t-u)$ due to autonomy of the deterministic version (4.3.1), and so we may rewrite a solution $(x(t))_{t\in[-r,\infty)}$ (4.3.1) in terms of the corresponding fundamental solution $(\check{x}(t))_{t\in[-r,\infty)}$ as

$$x(t) = \Upsilon(0)\check{x}(t) + b\int_{-r}^{0} \check{x}(t-r-u)\Upsilon(u)\,du + \sigma\int_{0}^{t} \check{x}(t-u)e^{\mu u}\,du \quad \text{for } t \geq 0. \tag{4.3.2}$$

The road to concentration inequalities in the spirit that we explained in the introduction passes through the Fernique inequality and best possible quantity estimates for the variance process $\|\Gamma_T\|$ and for the upper $\mathcal{L}^2$-bound on increments $\varphi(\cdot)$, see Section 2.1, for arbitrarily given time horizon $T \in (0,\infty)$. It is not really the behavior of $(x(t))_{t\in[-r,\infty)}$ that we are interested in, but it is the deviation $y(t) = x(t) - \mathbb{E}[x(t)]$, $t \in [-r,\infty)$, from its expectation process which thankfully leads us to a zero-mean Gaussian process. In case of SDDEs this is reflected by the fact that the initial segment of the deviation process is identically zero. And it is a noteworthy point that the previous analysis also provides an accurate description of the expectation process

$$\mathbb{E}[x(t)] = \Upsilon(0)\check{x}(t) + b\int_{-r}^{0} \check{x}(t-r-u)\Upsilon(u)\,du \quad \text{for all } t \geq 0.$$

This means that focussing on the deviation process is far from leaving parts of the analysis behind that one might feel uncomfortable with. And this is why, from now on and regarding concentration results, we will consider the deviation process or, in other words, the solution of

$$\begin{cases} dy(t) = -ay(t)dt + by(t-r)dt + \sigma e^{\mu t}\,dW(t) & \text{for } t \geq 0, \\ y_0 = 0, \end{cases} \tag{4.3.3}$$

where the final 0 is to be understood as the constant-zero mapping on J. That very solution is given by

$$y(t) = \sigma\int_{0}^{t} \check{x}(t-u)e^{\mu u}dW(u) \quad \text{for all } t \geq 0, \tag{4.3.4}$$

which can simply be read off from (4.3.2). If one is interested only in *how* to deduce concentration inequalities making use of the convergence in critical regime, this subsection has in principle come to an end, but as we are so ambitious to ask *what* can be learned, we are far from that. For anyone who fears the worst, namely an endless case analysis, there is not much comfort to spend, only that it is not endless. But one should also see the point that a variety of different settings is described here, some of them are really interesting from an applicationer's point of view, some of them are more technical. Furthermore, we promise to deduce concentration results in numerous cases that are relatively close to the optimum, and in some cases even arbitrarily close to the optimum. To spend some more comfort, we will not work out every case in detail, but merely present one of them detail and defer the rest to the appendix.

The sample case, that we study in detail, is the white-noise case for a critical parameter combination $(a, b) \in \mathcal{P}_I \cap \mathcal{P}_c$ and the reason for choosing this special case is that regarding its first-exit behavior from a given interval, for a large time horizon T, the corresponding solution behaves similar to a properly rescaled Brownian motion regarding first-exit-time distributions. The first part which concerns the concentrational behavior will be covered by the special-case study. The second part, that describes essentially how long the time horizon must be chosen in order to guarantee an exit from a given interval with high probability, will be worked out in the subsequent section. In order to derive an upper-bound estimate for $Q(p,T) = (2+\sqrt{2}) \int_1^\infty \varphi(Tp^{-u^2})\, du$, $p \in \mathbb{N}$, we establish upper bounds for $\mathcal{Q}_1(p,T)$ and $\mathcal{Q}_2(p,T)$, defined in (3.4.10), (3.4.11). We stick to $\kappa > 0$ as the minimal rate with which the fundamental solution $(\check{x}(t))_{t\in[-r,\infty)}$ converges to $(1+ar)^{-1}$ as in (4.1.12), (4.1.13).

Remark 4.15. *Note that one might also be thinking of σ to be time dependent, but bounded in absolute value by $|\sigma(t)| \leq \sigma$ for all $t \in [0,T]$, and the following proof would read all the same. As we have done before in similar situations, we will spare this effort to keep the notations simple.*

Regarding the account on autonomous stable SRFDEs in Section 3.4, we benefitted from the fact that fundamental solution matrices in that case vanish exponentially fast due to solely negative real parts of roots of the characteristic equation (3.4.2), or equivalently the roots of the characteristic mapping (4.0.2). But since we are dealing with $a = b$, it is a trivial fact that $\lambda = 0$ is a solution to the characteristic equation $a - \lambda = be^{\lambda r}$, and therefore, we do not have that very result available here. But, as we are in a simple case, we can compensate this gap rather easily by another result. The following lemma shows that whenever $(a,b) \in \mathcal{P}_I$ and $a \geq b$, it holds true that the corresponding fundamental solution does never leave $(0,1]$. This was shown for $a = b > 0$ in Lemma 4.1.

Lemma 4.16. *Let $(a,b) \in \mathcal{P}_I$ and $a \geq b$. Then the corresponding fundamental solution $(\check{x}(t))_{t\in[-r,\infty)}$ of (4.3.1) remains in $[0,1]$ over $[-r,\infty)$, and in $(0,1)$ over $(0,\infty)$.*

Proof. There are several ways to prove the claim. One way is to use the same arguments as in Lemma 4.1. Again, we assume that

$$\tau_0 := \inf\{t \geq r : \check{x}(t) = 0\} < \infty \quad \text{and} \quad \tau_1 := \inf\{t \geq r : \check{x}(t) = 1\} < \infty.$$

We know that $\check{x}(t) = e^{-at}$ for all $t \in [0,r]$ and an application of the classical solution representation (4.1.3) provides

$$\check{x}(r+s) = \check{x}(r)e^{-as} + b\int_0^s \check{x}(u)e^{-a(s-u)}du \geq \check{x}(r)e^{-as} \quad \text{for all } s \in (0,\infty), s \leq \tau_0 - r,$$

and

$$\begin{aligned}
\check{x}(\tau_1) &= \check{x}(0)e^{-a\tau_1} + b\int_0^{\tau_1} \check{x}(\tau_1 - r + u)e^{-a(\tau_1-u)}du \\
&< e^{-a\tau_1} + b\int_0^{\tau_1} e^{-a(\tau_1-u)}du \\
&= e^{-a\tau_1} + \frac{b}{a}\left(1 - e^{-a\tau_1}\right).
\end{aligned}$$

In the second case the assumption $b/a \leq 1$ serves the contradiction and settles the claim. $\square$

Remark 4.17. *Alternatively, we may recognize that the solution* λ *of the characteristic equation* $a-\lambda = be^{\lambda r}$ *is positive for* $a > b$*, see Theorem 4.14, and that* $\breve{y}(t) = e^{\lambda t}\breve{x}(t)$*,* $t \in [-r, \infty)$*, is the fundamental solution corresponding to the parameter combination* $\tilde{a} = a - \lambda = \tilde{b}$ *for which the result is already known from Lemma 4.1. So we may carry over the result to* $\breve{x} = e^{-\lambda t}\breve{y}(t)$*,* $t \in [-r, \infty)$*. The lower bound can also be taken as a part of* (4.0.4) *due to [KM92].*

As indicated above, we will regard the time horizon $T > 0$ as some arbitrarily fixed number, not too small.

Theorem 4.18 (Concentration inequality in critical regime with white noise)**.** *Let* Γ *and* $Q(T) = Q(p, T)$*, defined in* (2.1.1), (2.1.2) *and* (3.4.7)*, denote the parameters of the Fernique inequality with respect to the solution of* (4.3.1) *with* $a = b > 0$*,* $\mu = 0$*. Let further* κ *be as in* (4.1.12)*. Then*

$$\frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} \leq \frac{\sqrt{T}}{1+ar}\left(1 + \mathcal{R}(a, r, T, \kappa, p)\right), \tag{4.3.5}$$

where $\mathcal{R}(a, r, T, \kappa, p) = \frac{1+ar}{T\kappa} + \frac{(1+ar)^2}{4T\kappa} + \frac{7(1+ar)}{2\sqrt{p}\log(p)} + \frac{a(1+e^{\kappa r})}{\sqrt{2\kappa}}\frac{7(1+ar)\sqrt{T}}{4p\log(p)}$.

Proof. Concerning $\mathcal{Q}_2 = \mathcal{Q}_2(p, T)$, $p \in \mathbb{N}$, defined in (3.4.11), with an application of the Itô isometry and the convergence in critical regime in (4.1.13), we find that

$$\int_s^t \breve{x}^2(t-u)\,du \leq \int_s^t \left(\frac{1}{1+ar} + e^{-\kappa(t-u)}\right)^2 du \quad \text{for all } s, t \in [0, T],\ s < t.$$

The remainder consists of basic computations which reveal that

$$\begin{aligned}
\int_s^t \breve{x}^2(t-u)\,du &\leq \frac{t-s}{(1+ar)^2} + \frac{2}{1+ar}e^{-\kappa t}\int_s^t e^{\kappa u}du + e^{-2\kappa t}\int_s^t e^{2\kappa u}du \\
&= \frac{t-s}{(1+ar)^2} + \frac{2}{1+ar}e^{-\kappa t}e^{\kappa s}\int_0^{t-s} e^{\kappa u}du + e^{-2\kappa t}e^{2\kappa s}\int_0^{t-s} e^{2\kappa u}du \\
&= \frac{t-s}{(1+ar)^2} + \frac{2}{1+ar}e^{-\kappa(t-s)}\frac{e^{\kappa(t-s)}-1}{\kappa} + e^{-2\kappa(t-s)}\frac{e^{2\kappa(t-s)}-1}{2\kappa} \\
&= \frac{t-s}{(1+ar)^2} + \frac{2}{1+ar}\frac{1-e^{-\kappa(t-s)}}{\kappa} + \frac{1-e^{-2\kappa(t-s)}}{2\kappa} \\
&\leq \frac{t-s}{(1+ar)^2} + \frac{2}{1+ar}(t-s) + (t-s) \\
&= \left(\frac{1}{(1+ar)^2} + \frac{2}{1+ar} + 1\right)(t-s) \quad \text{for all } s, t \in [0, T],\ s < t,
\end{aligned}$$

but sometimes simple arguments provide the better estimates. If we use the fact that the fundamental solution does never leave $[0, 1]$ from Lemma 4.16, we directly find that

$$\int_s^t \breve{x}^2(t-u)\,du \leq t-s \quad \text{for all } s, t \in [0, T],\ s < t.$$

Why did we bother with the worthless computation in the first place? Because we will soon realize that $\mathcal{Q}_2$ is the term that serves the leading-order unpleasant term for the concentration inequality. So, maybe we have spared the reader to wonder whether the promising convergence result of the fundamental solutions might have done better than the rough estimate that comes from Lemma 4.16. And as a promise, we will not present that sort of fruitlessness again. For the rest of this special case, the convergence result yields the presented results concerning $\mathcal{Q}_1$ and the optimal one for $\|\Gamma\|$ and for the exponent in the concentration inequality; which is to say that the concentration result worths while.

For the $\mathcal{Q}_1$-term we apply the delay differential law of the fundamental solution to find that

$$\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 du = \int_0^s \left(\int_s^t -a\check{x}(v-u) + a\check{x}(v-u-r) dv \right)^2 du \qquad \text{for all } s,t \in [0,T],\ s<t.$$

Then an application of the convergence in the critical regime leads to

$$\begin{aligned}
&\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 du \\
&= \int_0^s \left(\int_s^t -a \left(\check{x}(v-u) - \frac{1}{1+ar} \right) + a \left(\check{x}(v-u-r) - \frac{1}{1+ar} \right) dv \right)^2 du \\
&\leq \int_0^s a^2 \left(\int_s^t e^{-\kappa(v-u)} + e^{-\kappa(v-u-r)} dv \right)^2 du \\
&= \int_0^s a^2 (1+e^{\kappa r})^2 \left(\int_s^t e^{-\kappa v} dv \right)^2 e^{2\kappa u} du \\
&= a^2(1+e^{\kappa r})^2 \frac{e^{2\kappa s}-1}{2\kappa} e^{-2\kappa s} \left(\int_0^{t-s} e^{-\kappa v} dv \right)^2 \\
&\leq a^2(1+e^{\kappa r})^2 \frac{1-e^{-2\kappa s}}{2\kappa} (t-s)^2 \\
&\leq \frac{a^2(1+e^{\kappa r})^2}{2\kappa} (t-s)^2 \qquad \text{for all } s,t \in [0,T],\ s<t.
\end{aligned}$$

Regarding $\|\Gamma\|$, we again apply the Itô isometry and the fact that the upcoming integral is monotonely increasing and obtain

$$\begin{aligned}
\frac{\|\Gamma\|}{\sigma^2} = \sup_{t\in[0,T]} \mathbb{E}\left[\left(\int_0^t \check{x}(t-u) dW(u) \right)^2 \right] &= \sup_{t\in[0,T]} \int_0^t \check{x}^2(t-u) du \\
&= \sup_{t\in[0,T]} \int_0^t \check{x}^2(u) du = \int_0^T \check{x}^2(u) du = \frac{\operatorname{var} y(T)}{\sigma^2}.
\end{aligned}$$

And then the convergence in the critical regime yields

$$\begin{aligned}
\frac{\|\Gamma\|}{\sigma^2} \leq \int_0^T \left(\frac{1}{1+ar} + e^{-\kappa u} \right)^2 du &= \frac{T}{(1+ar)^2} + \frac{2}{1+ar} \int_0^T e^{-\kappa u} du + \int_0^T e^{-2\kappa u} du \\
&= \frac{T}{(1+ar)^2} + \frac{2}{1+ar} \frac{1-e^{-\kappa T}}{\kappa} + \frac{1-e^{-2\kappa T}}{2\kappa} \\
&\leq \frac{T}{(1+ar)^2} \left(1 + 2\frac{1+ar}{T\kappa} + \frac{(1+ar)^2}{2T\kappa} \right).
\end{aligned}$$

Alltogether, using the estimate worked out in Theorem A.2 in the appendix we find that

$$\frac{\mathcal{Q}_1}{\sigma} \leq \int_1^\infty \sqrt{\sup_{\substack{s,t\in[0,T],\\ |t-s|\leq Tp^{-u^2}}} \varphi(t-s)}\, du \leq \int_1^\infty \sqrt{Tp^{-u^2}}\, du \leq \frac{\sqrt{T}}{\sqrt{p}\log p},$$
$$\frac{\mathcal{Q}_2}{\sigma} \leq \sqrt{\frac{a^2(1+e^{\kappa r})^2}{2\kappa}} \int_1^\infty Tp^{-u^2} du \leq \sqrt{\frac{a^2(1+e^{\kappa r})^2}{2\kappa} \frac{T}{2p\log p}},$$

and the proof is finished, where we applied that $\sqrt{1+x} \leq 1+\frac{x}{2}$ for $x \geq 0$. □

Proposition 4.19. *Consider the situation of Theorem 4.18*

a) We generally have that

$$\frac{\operatorname{var} y(T)}{\sigma^2} \geq \frac{T}{(1+ar)^2}\left(1 - \frac{1}{\kappa T}\log(1+ar) - \frac{2(1+ar)}{\kappa T}\right), \tag{4.3.6}$$
$$\frac{\operatorname{var} y(T)}{\sigma^2} \leq \frac{T}{(1+ar)^2}\left(1 + \frac{1+ar}{\kappa T}\left(2+\frac{1+ar}{2}\right)\right). \tag{4.3.7}$$

In particular,

$$\sqrt{\operatorname{var} y(T)} = \frac{\sigma\sqrt{T}}{1+ar}\left(1+\mathcal{O}(T^{-1})\right) \quad \textit{for big } T. \tag{4.3.8}$$

b) If we additionally assume that

$$T > \max\left\{\frac{5}{2\kappa}(1+ar)^2, 2\log(1+ar) + 4(1+ar)\right\},$$

we have that

$$\frac{\sigma\sqrt{T}}{1+ar}\sqrt{2^{-1}} \leq \sqrt{\operatorname{var} y(T)} \leq \frac{\sigma\sqrt{T}}{1+ar}\sqrt{2}.$$

Proof. We deal with the upper-bound estimates first, and afterwards, we turn on the lower-bound ones. With an application of the convergence of fundamental solutions in critical regime we obtain that

$$\frac{\operatorname{var} y(T)}{\sigma^2} = \int_0^T \check{x}^2(u) du \leq \int_0^T \left(\frac{1}{1+ar} + e^{-\kappa u}\right)^2 du$$
$$\leq \frac{T}{(1+ar)^2}\left(1+\frac{1+ar}{\kappa T}\left(2+\frac{1+ar}{2}\right)\right).$$

Using that $\sqrt{1+\xi} \leq 1+\frac{\xi}{2}$ for all $\xi \in (0,1)$, we find that

$$\sqrt{\frac{T}{(1+ar)^2}\left(1+\frac{1+ar}{\kappa T}\left(2+\frac{1+ar}{2}\right)\right)} \leq \frac{\sqrt{T}}{1+ar}\left(1+\frac{1}{2}\frac{1+ar}{\kappa T}\left(2+\frac{1+ar}{2}\right)\right).$$

Under the assumption $T > \frac{5}{2\kappa}(1+ar)^2$, we have that

$$\frac{1+ar}{\kappa T}\left(2+\frac{1+ar}{2}\right) \leq 1.$$

And as $(1+ar)^{-1} - e^{-\kappa u} \geq 0$ for all $u \geq \frac{1}{\kappa}\log(1+ar)$, we obtain

$$\begin{aligned}\frac{\operatorname{var} y(T)}{\sigma^2} = \int_0^T \check{x}^2(u)du &\geq \int_{\frac{1}{\kappa}\log(1+ar)}^T \left(\frac{1}{1+ar} - e^{-\kappa u}\right)^2 du \\ &\geq \frac{1}{(1+ar)^2}\left(T - \frac{\log(1+ar)}{\kappa}\right) - 2\int_{\frac{1}{\kappa}\log(1+ar)}^T \frac{e^{-\kappa u}}{1+ar} du \\ &\geq \frac{T}{(1+ar)^2}\left(1 - \frac{\log(1+ar)}{\kappa T} - \frac{2(1+ar)}{\kappa T}\right),\end{aligned}$$

which by $\sqrt{1-\xi} \geq 1-\xi$ for $\xi \in (0,1)$ serves the lower boundary for $\sqrt{\operatorname{var} y(T)}$ in (4.3.6).

□

Remark 4.20.

- *We notice the fact that the variance of the one-dimensional distribution is decreasing in r for sufficiently big T which at first glance seems counterintuitive.*

- *As we have mentioned before, an essential-growth rate in this case is provided by the result of [ARS13]. Due to the convergence of the fundamental solution, we obtain that*

$$\limsup_{t\to\infty} \frac{y(t)}{\sqrt{2t\log\log t}} = -\liminf_{t\to\infty} \frac{y(t)}{\sqrt{2t\log\log t}} = \frac{\sigma}{1+ar} \qquad \mathbb{P}\text{-}a.s.$$

- *Regarding Brownian motion, the reflection principle serves an easy way to a concentration inequality with the best-possible exponent. If we consider the rescaled Brownian motion $\frac{\sigma}{1+ar}W(t)$, $t \in [0,T]$, and define*

$$\tilde{\sigma}_T := \frac{\sigma}{1+ar}\sqrt{\operatorname{var} W(T)} = \frac{\sigma\sqrt{T}}{1+ar},$$

we may observe that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \frac{\sigma}{1+ar}|W(s)| > h\tilde{\sigma}_T\right\} \leq 4\exp\left(-\frac{h^2}{2}\right) \qquad \text{for arbitrary } h > 0.$$

For the solution $(y(t))_{t\in[0,\infty)}$ of the SDDE (4.3.1) in critical regime with white noise a reformulation of Theorem 4.18 with $\sigma_T := \sqrt{\operatorname{var} y(T)}$ yields

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |y(s)| > h\sigma_T\right\} \leq \frac{5}{2}p^2 \exp\left(-\frac{h^2}{2}\left(1 + \mathcal{O}\left(1 + T^{-1} + \frac{\sqrt{T}}{p}\right)\right)\right)$$
$$\text{for arbitrary } h > \frac{\sqrt{1+4\log p}}{\sigma_T}, \qquad (4.3.9)$$

when T and p are large compared to the other parameters and T is small compared to p^2, e.g. $p = T^{1/2+\alpha}$ for arbitrary $\alpha > 0$. And as by Proposition 4.19 $\tilde{\sigma}_T = \sigma_T(1 + \mathcal{O}(T^{-1}))$, we might replace σ_T by $\tilde{\sigma}_T$ in (4.3.9) – the error terms merge and do not show up. Therefore, the concentration inequality shows the same exponent as the rescaled Brownian motion up to small correction terms, if T is big. To compensate the undesirable prefactor p^2, it may be drawn as $2\log p$ into the exponent. Therefore, concentration inequality (4.3.9) is useful if we assume that h is at least of order $\sqrt{\log T}$. The classical result for rescaled Brownian motion is not restricted in such way.

- *An amusing fact on the size of $\sqrt{\log T}$. It is formally undoubtedly true that $\sqrt{\log T}$*

converges to $+\infty$ rather slowly when T goes to ∞, but what is slow?

▷ *With regard to the largest number that can technically be displayed by an ordinary calculator, a tiny bit less than 10^{100}, the American mathematician Edward Kasner is rumored to have invented the term googol ($\equiv 10^{100}$) in collaboration with his nine-year old nephew. To get some sort of feeling for the size, let us mention that the overall number of protons in the universe nowadays is estimated between 10^{80} and 10^{89} which is still far from a googol. But if you take the square root of the logarithm of a googol, you end up with barely* 15.2.

▷ *To travel one Planck length in vacuum at the speed of light, one needs an amount of time called the Planck time, and it is about $6 \cdot 10^{-44}$ seconds. And with about $3 \cdot 10^7$ seconds per year we have that the universe is about $2 \cdot 10^{61}$ units of Planck time old. And $\sqrt{\log(2 \cdot 10^{61})} \approx 11{,}9$.*

The remainder of this subsection presents the concentration inequalities for every possible relation of the parameters a, b, μ, λ, and κ. We will distinguish between the different regimes in terms of a and b as before:

- The *critical* regime refers to $a = b > 0$,
- the *instable* regime considers $b > 0$ and $a < b$,
- the *stable* regime refers to $a > b > 0$.

And with respect to the noise parameter μ, we will use the term...

- *increasing noise* when $\mu > 0$,
- *vanishing noise* for $\mu < 0$,
- *white noise* for $\mu = 0$.

All the computational details can be found in the Appendix B. Generally, those are quiet similar to the computations we presented for the white-noise case in the critical regime. We mostly confine to present the best upper-bound estimates for $\frac{\|\Gamma\|}{\sigma^2}$ and $\frac{Q(\cdot)}{\sigma}$ from Section 2.1, that we have achieved, and spare the effort to additionally formulate the corresponding concentration inequality.

Critical regime. The white-noise case has been presented in detail in Theorem 4.18, Proposition 4.19 and discussed in Remark 4.20. In the increasing-noise case we achieve that

$$\begin{aligned}\frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} &\leq \frac{e^{\mu T}}{\sqrt{2\mu}} + \frac{7}{2}\frac{\sqrt{T}e^{\mu T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2(\kappa+\mu)}}\frac{Te^{\mu T}}{2p\log p} \\ &= \frac{e^{\mu T}}{\sqrt{2\mu}}\left(1 + \frac{7\sqrt{2\mu}}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}a(1+e^{\kappa r})\frac{\sqrt{2\mu}}{\sqrt{2(\kappa+\mu)}}\frac{T}{2p\log p}\right).\end{aligned} \tag{4.3.10}$$

In case of vanishing noise we find that

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq \begin{cases} v_0 + \frac{7}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}}\sqrt{1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}}\frac{T}{p\log p} & \text{for } \mu \notin \{-\kappa, -\frac{\kappa}{2}\}, \\ v_1 + \frac{7}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2\kappa e}}\frac{T}{2p\log p} & \text{for } \mu = \kappa, \\ v_2 + \frac{7}{2}\left(\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}}\sqrt{1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}}\frac{T}{2p\log p}\right) & \text{for } \mu = -\frac{\kappa}{2}, \end{cases} \tag{4.3.11}$$

where

$$v_0^2 := \max\left\{\frac{1}{|2\mu|}, \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|} + \frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)\right\}, \tag{4.3.12}$$

$$v_1^2 := \max\left\{\frac{1}{|2\mu|}, \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|} + \frac{1}{2\kappa e}\right\}, \tag{4.3.13}$$

$$v_2^2 := \max\left\{\frac{1}{|2\mu|}, \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{(1+ar)\kappa e} + \frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)\right\}. \tag{4.3.14}$$

Neither in case of increasing noise, nor in case of vanishing noise something truely surprising has occured. If the noise intensifies exponentially with $\mu > 0$, one has to choose x at least of order $\exp(\mu T)$ in order to give Fernique's inequality 2.1 a senseful meaning. In case of vanishing noise, x must have at least a size of order $\sqrt{\log T}$ due to compensate the p^2 of the prefactor and assuming that $p^2/T \geq 1$.

Stable Regime. We let $\check{x}$ denote the fundamental solution with respect to the parameter combination $a > b > 0$. With regard to Theorem 4.14, we let $\lambda > 0$ such that $\tilde{a} = a - \lambda = be^{\lambda r} = \tilde{b}$ implying that

$$\check{x}(t) \leq e^{-\lambda t} \quad \text{for all } t \in [-r, \infty),$$

because $(\check{x}(t)e^{\lambda t})_{t\in[-r,\infty)}$ is the fundamental solution in a critical regime. This estimate improves inequality (4.0.5) from the general case. Further, from Theorem 4.5 for $\tilde{\kappa} = |\log(1-e^{-\tilde{a}r})|/(2r)$ we have that

$$\check{x}(t) \leq \left(\frac{1}{1+\tilde{a}r} + e^{-\tilde{\kappa}t}\right)e^{-\lambda t} \quad \text{for all } t \in [0, \infty).$$

White noise.

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_0 + \frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a+be^{-\lambda r}}{\sqrt{2\lambda}}\frac{T}{2p\log p}.$$

where

$$v_0^2 := \min\left\{\frac{1}{2\lambda}, \frac{1}{(1+\tilde{a}r)^2}\frac{1}{2\lambda} + \frac{2}{1+\tilde{a}r}\frac{1}{\tilde{\kappa}+2\lambda} + \frac{1}{2(\kappa+\lambda)}\right\}.$$

Vanishing noise.

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma}\leq$$
$$\begin{cases} v_2+\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda r}}{\sqrt{2\lambda e}}\frac{T}{2p\log p} & \text{for } \mu=-\lambda,\\ v_3+\sqrt{\frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}}\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda r}}{\sqrt{2|\lambda+\mu|}}\sqrt{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}\frac{T}{2p\log p} & \text{for } \mu=-\lambda-\frac{\tilde\kappa}{2},\\ v_4+\sqrt{\frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{\tilde\kappa}}\frac{\sqrt{T}}{\sqrt{p}\log p}+\sqrt{\frac{(a+be^{\lambda r})^2}{2\tilde\kappa}\left(1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}\right)}\frac{T}{p\log(p)} & \text{for } \mu=-\lambda-\tilde\kappa,\\ v_1+\sqrt{\frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}}\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda r}}{\sqrt{2|\lambda+\mu|}}\sqrt{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}\frac{T}{2p\log p} & \text{else,} \end{cases}$$

where

$$\begin{aligned} v_1^2&=\frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde a r)^2}+\frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde\kappa+2\lambda)}{(2|\mu|)\vee(\tilde\kappa+2\lambda)}\right)}{|\tilde\kappa+2\lambda+2\mu|(1+\tilde a r)}+\frac{1-\frac{|\mu|\wedge(\tilde\kappa+\lambda)}{|\mu|\vee(\tilde\kappa+\lambda)}}{2|\tilde\kappa+\lambda+\mu|},\\ v_2^2&=\max\left\{\frac{1}{2\lambda e},\frac{1}{(1+\tilde a r)^2 2\lambda e}+\frac{2}{1+\tilde a r}\frac{\left(1-\frac{2|\mu|}{\tilde\kappa+2\lambda}\right)}{\tilde\kappa+2\lambda+2\mu}+\frac{\left(1-\frac{|\mu|}{\tilde\kappa+\lambda}\right)}{2(\tilde\kappa+\lambda+\mu)}\right\},\\ v_3^2&=\frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde a r)^2}+\frac{2}{(1+\tilde a r)(\tilde\kappa+2\lambda)e}+\frac{1-\frac{|\mu|\wedge(\tilde\kappa+\lambda)}{|\mu|\vee(\tilde\kappa+\lambda)}}{2(\tilde\kappa+\lambda+\mu)},\\ v_4^2&=\max\left\{\frac{1-\frac{\lambda}{|\mu|}}{2\tilde\kappa},\frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2\tilde\kappa(1+\tilde a r)^2}+\frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde\kappa+2\lambda)}{(2|\mu|)\vee(\tilde\kappa+2\lambda)}\right)}{\tilde\kappa(1+\tilde a r)}+\frac{1}{(\tilde\kappa+2\lambda)e}\right\}. \end{aligned}$$

Increasing noise

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma}\leq v_5 e^{\mu T}\left(1+\frac{7\sqrt{2(\lambda+\mu)}}{2v_0}\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{7(a+be^{\lambda r})}{2v_0}\frac{T}{2p\log p}\right).$$

where

$$v_5^2:=\frac{1}{2(\lambda+\mu)(1+\tilde a r)^2}+\frac{2}{(1+\tilde a r)(\tilde\kappa+2\lambda+2\mu)}+\frac{1}{2(\tilde\kappa+\lambda+\mu)}. \tag{4.3.15}$$

Instable Regime. To simplify notations, we suppose that $-\lambda$ solves the characteristic equation such that $\tilde a := a-\lambda = be^{-\lambda r} =: \tilde b$ and define $\tilde\kappa := |\log(1-e^{-\tilde a r})|/(2r)$ as always. Then,

$$\check x(t)\leq\left(\frac{1}{1+\tilde a r}+e^{\tilde\kappa t}\right)e^{\lambda t},\quad \check x(t)\leq e^{\lambda t}\quad \text{for all } t\in[0,\infty).$$

White Noise.

$$\begin{aligned} \frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma}&=\frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde a r)}\sqrt{1+\mathcal{O}\left(e^{-(\tilde\kappa\wedge(2\lambda))t}\right)}+e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p}+e^{\lambda T}\frac{a+be^{\lambda}r}{\sqrt{2\lambda}}\frac{T}{2p\log p}\\ &\leq\frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde a r)}\left(1+\mathcal{O}\left(e^{-\frac{(\tilde\kappa\wedge(2\lambda))t}{2}}\right)+\sqrt{2\lambda}(1+\tilde a r)\left(\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda}r}{\sqrt{2\lambda}}\frac{T}{2p\log p}\right)\right). \end{aligned}$$

Vanishing noise.

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \le \frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde{a}r)}\left(1+\mathcal{O}\left(\frac{e^{-\min\{\tilde{\kappa}-\nu,2|\mu|\}T}}{\nu}\right)\right.$$
$$\left.+\sqrt{2\lambda}(1+\tilde{a}r)\left(\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda r}}{\sqrt{2(\mu-\lambda)}}\frac{T}{2p\log p}\right)\right).$$

Increasing noise.

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \le$$
$$\begin{cases} \frac{e^{\lambda T}}{\sqrt{2(\lambda-\mu)}(1+\tilde{a}r)}\left(1+\mathcal{O}\left(\frac{1}{\nu}e^{-\frac{\rho_0 T}{2}}\right)\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{-\lambda r}}{\sqrt{2(\mu-\lambda)}}e^{\lambda T}\frac{T}{2p\log p}\right) & \text{for } 0<\mu<\lambda, \\ \frac{\sqrt{T}e^{\lambda T}}{1+\tilde{a}r}\left(1+\sqrt{\frac{1+\tilde{a}r}{T\tilde{\kappa}}}+\frac{1+\tilde{a}r}{\sqrt{2\tilde{\kappa}T}}+\frac{1+\tilde{a}r}{\sqrt{p}\log p}+(1+\tilde{a}r)(a+be^{-\lambda r})\frac{T}{2p\log p}\right) & \text{for } \mu=\lambda, \\ e^{\lambda T}v_0\left(1+\frac{\sqrt{T}}{v_0\sqrt{p}\log p}+\frac{a+be^{-\lambda r}}{v_0\sqrt{2(\mu-\lambda)}}\frac{T}{2p\log p}\right) & \text{for } \mu>\lambda, \end{cases}$$

where

$$\rho_0 := \min\{2(\lambda-\mu),\tilde{\kappa}-\nu\},$$
$$v_0^2 := \frac{1}{2(\mu-\lambda)(1+\tilde{a}r)^2}+\frac{2}{(2\mu-2\lambda+\tilde{\kappa})(1+\tilde{a}r)}+\frac{1}{2(\mu-\lambda+\tilde{\kappa})}.$$

4.4. Small-Ball Probabilities

We let $y = (y(t))_{t\in[-r,\infty)}$ be the deviation process of (4.3.3) in the critical regime under white noise, which means the solution of the SDDE (4.3.3) with initial segment $\Upsilon = 0$. In the course of the previous section we have established a concentration inequality for y with an exponent that resembles the best-possible exponent for an appropriately rescaled Brownian motion up to a logarithmic order term in T. If we assume $h > 0$ to be fixed and regard the concentration probability $P(T) = \mathbb{P}\{\sup_{s\in[0,T]} |y(s)| > h\}$ as a mapping of T, the concentration inequalities of the previous section have provided a reasonable description of the region in terms of T where $P(T)$ is close to zero. Concerning concentration inequalities in general the theory of *large deviations* has approved as an invaluable tool providing exact exponents in rather general situations. In this section we will turn the spotlight on regions where $P(T)$ is close to one. In other words, we bring up the question which size of the time horizon suffices to guarantee a first exit of y from a given tube of radius h prior to T with probability $P(T)$ close to one. In that situation the large-deviation theory is of little use – at least in the classical formulation due to e.g. [DZ92], [Fre12]. In case of Brownian motion a naive start uses that exponential moments $\mathbb{E}\exp(\lambda\tau)$ of first-exit times $\tau = \inf\{t : |W(t)| > h\}$ exist if λ small enough. In that case the Markov inequality yields

$$\mathbb{P}\left\{\sup_{t\in[0,T]} |W(s)| < h\right\} \le \frac{1}{\cos\left(\frac{\pi}{2}\gamma\right)} \exp\left(-\frac{\pi^2}{8h^2}\gamma^2 T\right) \qquad \text{for all } \gamma \in [0,1).$$

Details are deferred to the Appendix A.3.1. This estimate is not optimal, but it is fairly easy to derive through martingale techniques and there is at least some resemblance to the optimal version as we are going to see in a moment, when we have revisited the first-exit-time problem in this case of a Brownian motion. In the literatur, estimates of the form $\mathbb{P}\{\sup_{s\in[0,T]} |y(s)| < h\} = 1 - \mathbb{P}\{\sup_{s\in[0,T]} |y(s)| > h\}$ are well known by the term *small-ball* probabilities, and sometimes for the corresponding theory one finds the term *small deviations* emphasizing the contrast to the large-deviation theory.

The literature provides a remarkable development regarding small-ball probabilites, see e.g. [Li99], [Li03], [BDS01], [LS02]. An introduction as well as brief survey on small-ball probabilities can be found in [LS01] by Wenbo V. Li and Qi-Man Shao; a more recent survey from Hoi H. Nguyen and Van H. Vu can be found in [NV13]. In the case of Brownian motion $(W(t))_{t\in[0,T]}$, asymptotics of small-ball probabilities are well understood meaning that the exact asymptotics of exponent and prefactor are known in this case. For given $h > 0$, see A.3.2, it is generally true that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |W(s)| < h\right\} \le \frac{4}{\pi} \exp\left(-\frac{\pi^2}{8h^2} T\right) \qquad \text{for } h > 0. \tag{4.4.1}$$

The main purpose of this section is an estimate for an analogue of a small-ball probability for SDDEs in critical regime.

The following theorem reveals an estimate that is applicable in a variety of situations, and so it will not always be useful. We will study few special cases in the subsequent corollaries.

Theorem 4.21. *Let $(y(t))_{t\in[-r,\infty)}$ be the deviation process of the SDDE* (4.3.1) *in critical regime with white noise, i.e. $a=b>0$, $\mu=0$, finite time horizon $T>0$, and let κ denote the exponential rate of convergence of the fundamental solution, given in* (4.1.12)*. Let further $\delta_1,\delta_2,\delta_3$ be arbitrary positive constants, and $T_0,T_1>0$ with $T=T_0+T_1$. Then, we denote $\Delta:=\delta_1+\delta_2+\delta_3$ and $\tilde{\Delta}:=\Delta(1+ar)$ and we assume that δ_2 is big enough to satisfy*

$$h_0 := \frac{\delta_2 e^{\kappa T_0}}{\frac{1}{\sqrt{2\kappa}}+\frac{\sqrt{T_1}}{\sqrt{p}\log p}+\frac{a(1+e^{\kappa r})}{\sqrt{2\kappa}}\frac{T_1}{2p\log p}} \geq \sqrt{1+4\log p}, \tag{4.4.2}$$

where $p\in\mathbb{N}$ is some integer, $p\geq 2$. If we denote the fundamental solution of (4.3.1) *by $(\check{x}(t))_{t\in[-r,\infty)}$ and $v(t):=\operatorname{var} y(t)=\int_0^t \check{x}^2(u)du$, $t\geq 0$, Then*

$$\begin{aligned} P_\star &:= \mathbb{P}\left\{\sup_{s\in[0,T]}\frac{|y(s)|}{\sigma}>\delta_1\right\} \\ &\geq 1-\frac{4}{\pi}\exp\left(-\frac{\pi^2}{8\tilde{\Delta}^2}T_1\right)-\frac{5p^2}{2}\exp\left(-\frac{h_0^2}{2}\right)-\exp\left(-\frac{\delta_3^2}{2v(T_0)}\right). \end{aligned}$$

Remark 4.22. *The requirement* (4.4.2) *that originates from the Fernique inequality is a fairly weak assumption. If $\delta_3(T)$ has some minimal size, it is only an assumption on the size of T_1.*

Proof. We use the explicit representation of y through the variation-of-constants formula (4.3.4) and put σ on the left-hand side for a little ease of notation in the subsequent computation. Adding the clever zero $\frac{1}{1+ar}-\frac{1}{1+ar}$ in the integrand and a well-considered decomposition of the dW-integral yields

$$\frac{y(t)}{\sigma}=\int_0^t \check{x}(t-u)dW(u)=\mathcal{J}^{(1)}(t-T_0)+\mathcal{J}^{(2)}(t-T_0)+\mathcal{J}^{(3)}(t) \tag{4.4.3}$$

with the representatives

$$\begin{aligned} \mathcal{J}^{(1)}(t-T_0) &= \int_0^{t-T_0}\frac{1}{1+ar}\,dW(u)=\frac{W(t-T_0)}{1+ar}, \\ \mathcal{J}^{(2)}(t-T_0) &= \int_0^{t-T_0}\check{x}(t-u)-\frac{1}{1+ar}\,dW(u) \text{ and} \\ \mathcal{J}^{(3)}(t) &= \int_{t-T_0}^{t}\check{x}(t-u)\,dW(u) \quad \text{each for all } t\in[T_0,T]. \end{aligned}$$

The capability of the decomposition lies in the improved tractability of the arising terms. The term $\mathcal{J}^{(1)}$ is a rescaled Brownian motion on the interval $[T_0,T_1]$, and the small-ball probability (4.4.1) provides an excellent lower bound for the first-exit-time distribution. Further, $\mathcal{J}^{(2)}$ has an exponentially decaying integrand, and will therefore give a minor contribution compared to $\mathcal{J}^{(1)}$ with high probability when T_1 is sufficiently big. And finally, to have $\mathcal{J}^{(3)}$ relatively small with high probability, it is necessary and sufficient that T_0 is small compared to T_1. Accordingly, we define the stopping times

$$\begin{aligned} \tau_\Delta^{(1)} &:= \inf\left\{t\geq T_0 : \left|\mathcal{J}^{(1)}(t-T_0)\right|>\Delta\right\}, \\ \tau_{\delta_2}^{(2)} &:= \inf\left\{t\geq T_0 : \left|\mathcal{J}^{(2)}(t-T_0)\right|>\delta_2\right\}. \end{aligned}$$

Overlooking the decomposition (4.4.3), we conclude that for y/σ to leave the tube of radius δ_1 prior to T, it is sufficient that the rescaled Brownian motion $\mathcal{J}^{(1)}$ exits from the bigger tube of radius Δ, $\mathcal{J}^{(2)}$ remains relatively tame over the whole time interval and $\mathcal{J}^{(3)}$ behaves nicely in the very moment in which $\mathcal{J}^{(1)}$ exits the Δ-tube. Formally,

$$P_\star \geq \mathbb{P}\left\{\sup_{t\in[T_0,T]} |\mathcal{J}^{(1)}(t-T_0)| > \Delta\right\} - \mathbb{P}\left\{\sup_{t\in[T_0,T]} |\mathcal{J}^{(2)}(t-T_0)| > \delta_2\right\} - \mathbb{P}\left\{|\mathcal{J}^{(3)}(\tau_\Delta^{(1)} \wedge T_1)| > \delta_3\right\}. \tag{4.4.4}$$

The claim follows through the analysis of the involved probabilities for which we define the short-hand notations

$$P_1 := \mathbb{P}\left\{\tau_\Delta^{(1)} < T\right\}, \quad P_2 := \mathbb{P}\left\{\tau_{\delta_2}^{(2)} < T\right\}, \quad P_3 := \mathbb{P}\left\{\left|\mathcal{J}^{(3)}(\tau_\Delta^{(1)} \wedge T_1)\right| > \delta_3\right\}. \tag{4.4.5}$$

In order to derive a lower bound for probability P_1, we reformulate the event in terms of Brownian motion by

$$\left|\int_0^{t-T_0} \frac{1}{1+ar}\, dW(u)\right| > \Delta \quad \Leftrightarrow \quad |W(t-T_0)| > \Delta(1+ar) = \tilde{\Delta}.$$

Then, an application of the small-ball estimate (4.4.1), or A.3.2 respectively, reveals that

$$\mathbb{P}\left\{\sup_{s\in[0,T-T_0]} |W(s)| > \tilde{\Delta}\right\} \geq 1 - \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8\tilde{\Delta}^2}(T-T_0)\right).$$

Regarding the probability P_2, an upper-bound estimate follows from an application of the Fernique inequality. Due to the fact that this is only an instance of a concentration inequality, the applied techniques are naturally similar to the ones of the preceeding chapter. We work out the details to make sure that the result reflects the fact that there is an additional (helpful) term $e^{-\kappa t}$, because $|\check{x}(t) - \frac{1}{1+ar}| < e^{-\kappa t}$ and $t > T_0$, in this case.

$$\int_{s-T_0}^{t-T_0} \left(\check{x}(t-u) - \frac{1}{1+ar}\right)^2 du \leq \int_{s-T_0}^{t-T_0} e^{-2\kappa(t-u)}\, du = e^{-2\kappa T_0} e^{-2\kappa(t-s)} \int_0^{t-s} e^{-2\kappa v}\, dv \quad \text{for all } s,t \in [0,T_1],\ s<t.$$

Here we may apply that $e^{-\kappa u} < 1$ twice; for $u = t-s$ and $u = v$ in the above right most term. Therefore,

$$\int_{s-T_0}^{t-T_0} \left(\check{x}(t-u) - \frac{1}{1+ar}\right)^2 du \leq e^{-2\kappa T_0}(t-s) \quad \text{for all } s,t \in [T_0,T],\ s<t. \tag{4.4.6}$$

And with an application of the delay differential law of the fundamental solution, an additional clever zero, and the convergence of fundamental solutions in critical regime, we obtain

$$\begin{aligned}
&\int_0^{s-T_0} \big(\check{x}(t-u) - \check{x}(s-u)\big)^2 du \\
&= \int_0^{s-T_0} \left(\int_s^t -a\check{x}(v-u) + b\check{x}(v-u-r) dv \right)^2 du \\
&\le a^2 \int_0^{s-T_0} \left(\int_s^t \left|\check{x}(v-u) - \frac{1}{1+ar}\right| + \left|\check{x}(v-u-r) - \frac{1}{1+ar}\right| dv \right)^2 du \\
&\le a^2 \int_0^{s-T_0} \left(\int_s^t e^{-\kappa(v-u)} + e^{-\kappa(v-u-r)} dv \right)^2 du \quad \text{for all } s,t \in [T_0,T],\ s<t.
\end{aligned}$$

Then, the rest of the estimate follows from sheer computations,

$$\begin{aligned}
\int_0^{s-T_0} \big(\check{x}(t-u) - \check{x}(s-u)\big)^2 du &\le a^2\big(1+e^{\kappa r}\big)^2 \int_0^{s-T_0} \left(\int_s^t e^{-\kappa(v-u)} dv \right)^2 du \\
&= a^2\big(1+e^{\kappa r}\big)^2 \int_0^{s-T_0} e^{2\kappa u} du \left(\int_s^t e^{-\kappa v} dv \right)^2 \\
&\le \frac{a^2\big(1+e^{\kappa r}\big)^2}{2\kappa} \left(e^{2\kappa(s-T_0)} - 1 \right) e^{-2\kappa s} \left(\int_s^t e^{-\kappa(v-s)} du \right)^2 \\
&\le \frac{a^2\big(1+e^{\kappa r}\big)^2}{2\kappa} e^{-2\kappa T_0} (t-s)^2 \quad \text{for all } s,t \in [T_0,T],\ s<t.
\end{aligned}$$

Then, for the quantities $\mathcal{Q}_1 = \mathcal{Q}(p,T_1)$ and $\mathcal{Q}_2 = \mathcal{Q}_2(p,T_1)$, defined in (3.4.10) and (3.4.11), we find that

$$\mathcal{Q}_1 \le e^{-\kappa T_0} \frac{\sqrt{T_1}}{\sqrt{p}\log p} \quad \text{and} \quad \mathcal{Q}_2 \le e^{-\kappa T_0} \frac{a(1+e^{\kappa r})}{\sqrt{2\kappa}} \frac{T_1}{2p\log p}.$$

And for the corresponding $\|\Gamma_{\mathcal{J}^{(2)}}\|$-term, where $\Gamma_{\mathcal{J}^{(2)}}$ is defined as Γ in Section (2.1) but with repect to $\mathcal{J}^{(2)}$, another application of the convergence in critical regime serves

$$\begin{aligned}
\|\Gamma_{\mathcal{J}^{(2)}}\| = \sup_{t\in[T_0,T]} \mathbb{E}\left[\left(\mathcal{J}^{(2)}(t-T_0) \right)^2 \right] &= \sup_{t\in[T_0,T]} \int_0^{t-T_0} \left(\check{x}(t-u) - \frac{1}{1+ar} \right)^2 du \\
&\le \sup_{t\in[T_0,T]} e^{-2\kappa T_0} \int_0^{t-T_0} e^{-2\kappa(t-T_0-u)} du.
\end{aligned}$$

Therefore, through a substitution $v = t - T_0 - u$ we obtain

$$\|\Gamma_{\mathcal{J}^{(2)}}\| \le \sup_{t\in[T_0,T]} e^{-2\kappa T_0} \int_0^{t-T_0} e^{-2\kappa v}\, dv = e^{-2\kappa T_0} \int_0^{T-T_0} e^{-2\kappa v}\, dv \le \frac{e^{-2\kappa T_0}}{2\kappa}.$$

Therefore, for the Fernique coefficient $Q_{\mathcal{J}^{(2)}}$, defined as Q in Section 2.1 but with respect to $\mathcal{J}^{(2)}$, we find the following beautiful upper bound

$$\sqrt{\|\Gamma_{\mathcal{J}^{(2)}}\|} + Q_{\mathcal{J}^{(2)}}(T-T_0) \le e^{-\kappa T_0} \left(\frac{1}{\sqrt{2\kappa}} + \frac{\sqrt{T_1}}{\sqrt{p}\log p} + \frac{a(1+e^{\kappa r})}{\sqrt{2\kappa}} \frac{T_1}{2p\log p} \right). \tag{4.4.7}$$

Here we recognize the appearing term from the definition of h_0 in (4.4.2). The corresponding minimality condition on δ_2 implies that

$$h_0\left(\sqrt{\|\Gamma_{\mathcal{J}^{(2)}}\|}+Q_{\mathcal{J}^{(2)}}(T-T_0)\right)\geq\delta_2.$$

We find an upper-bound estimate for P_2 through an application of the Fernique inequality, which provides

$$\begin{aligned}\mathbb{P}\left\{\tau_{\delta_2}^{(2)}<T\right\}&\leq\mathbb{P}\left\{\sup_{t\in[T_0,T]}\left|\mathcal{J}^{(2)}(t-T_0)\right|>h_0\left(\sqrt{\|\Gamma_{\mathcal{J}^{(2)}}\|}+Q_{\mathcal{J}^{(2)}}(T-T_0)\right)\right\}\\&<\frac{5p^2}{2h_0}e^{-\frac{h_0^2}{2}}.\end{aligned}\tag{4.4.8}$$

In order to find an upper bound for $P_3=\mathbb{P}\left\{\left|\mathcal{J}^{(3)}(\tau_\Delta^{(1)}\wedge T_1)\right|>\delta_3\right\}$, we start with an ease of notations and denote $\tau=\tau_\Delta^{(1)}\wedge T_1$ for the rest of the proof. We remember that the underlying probability space features the completed filtration $\mathcal{F}=(\mathcal{F}_t)_{t\geq 0}$, that is generated by the Brownian motion, in particular $W(t)$ is measurable with respect to $\mathcal{F}_t$ for each t. Rewriting the term $\mathcal{J}^{(1)}(t-T_0)=(1+ar)^{-1}W((t-T_0)\wedge 0)$ emphasizes the first essential observation in this part of the proof which is that $\{\tau\leq t\}\in\mathcal{F}_{(t-T_0)\wedge T_1}$ for all $t\geq T_0$, because $\mathcal{J}^{(1)}$ is nothing but the rescaled Brownian motion time-shifted by T_0; Informally speaking, at time $t=T_0$, the process $\mathcal{J}^{(1)}$ starts in $(1+ar)^{-1}W(0)$ and then traces the path of the rescaled Brownian motion with the time lag of T_0. Hence,

$$\{\tau\leq t\}\in\mathcal{F}_{(t-T_0)\wedge T_1}\subset\mathcal{F}_{t-T_0}\quad\text{and therefore,}\quad\{\tau=t\}\in\mathcal{F}_{t-T_0}\quad\text{for all } t\in[0,T].$$

The second essential observation is, informally speaking, that $\mathcal{J}^{(3)}(t)$ evaluated at some arbitrary $t\geq T_0$ can only *see* a time length of T_0 into the past. That means all that $\mathcal{J}^{(3)}(t)$ may observe from the path $(W(u))_{u\in[0,t-T_0]}$ is the very end point, namely $W(t-T_0)$. And that one can not have any meaning to $\mathcal{J}^{(3)}(t)$. To make this idea become a rigorous argument, that works for the stopping time τ instead of t, it is convenient to introduce the notion of

$$W^{(\tau)}(t):=W(\tau+t)-W(\tau)\quad\text{for all } t\in[0,\infty),$$

the Brownian motion restartet at τ. Let us for a moment consider the integrand of $\mathcal{J}^{(3)}$ as a mapping of two arguments: $h(t,u):=\check{x}(t-u)$ for $t,u\in[0,\infty), t-u>-r$. Fix $\hat{t}$ and consider $u\mapsto h(\hat{t},u)$. By the integration-by-parts formula, we deduce that

$$\begin{aligned}h(\hat{t},t)W(t)=h(\hat{t},t-T_0)W(t-T_0)+\int_{t-T_0}^{t}h(\hat{t},u)dW(u)+\int_{t-T_0}^{t}W(u)h(\hat{t},du)\\+\frac{1}{2}2\int_{t-T_0}^{t}(dh(\hat{t},u))(dW(u)),\end{aligned}$$

where the last term is zero. Therefore, an application of the integration-by-parts formula (always understand $\dot{\check{x}}(0)$ as the right-hand derivative in 0) and substituting $s = t - u$ yields

$$\begin{aligned}\mathcal{J}^{(3)}(t) = \int_{t-T_0}^{t} \check{x}(t-u)dW(u) &= \check{x}(0)W(t) - \check{x}(T_0)W(t-T_0) - \int_{t-T_0}^{t} W(u)\frac{d}{du}(\check{x}(t-u))du \\ &= W(t) - \check{x}(T_0)W(t-T_0) + \int_{t-T_0}^{t} W(u)\dot{\check{x}}(t-u)du \\ &= W(t) - \check{x}(T_0)W(t-T_0) + \int_{0}^{T_0} W(t-s)\dot{\check{x}}(s)ds.\end{aligned}$$

Introduction of a *smart zero* $\tau - \tau$ is feasible even pathwise as τ is pathwise bounded by T_1 by definition. We observe that

$$\mathcal{J}^{(3)}(t) = W(\tau + (t-\tau)) - \check{x}(T_0)W(\tau + (t-\tau-T_0)) + \int_0^{T_0} W(\tau + (t-\tau-s))\dot{\check{x}}(s)ds \qquad \text{for all } t \geq T_0.$$

Restating this observation in terms of $W^{(\tau)}$ and remembering that $\check{x}(0) = 1$ reveals

$$\begin{aligned}\mathcal{J}^{(3)}(t) &= \big(W^{(\tau)}(t-\tau) + W(\tau)\big) - \check{x}(T_0)\big(W^{(\tau)}(t-T_0-\tau) + W(\tau)\big) \\ &\quad + \int_0^{T_0} \big(W^{(\tau)}(t-\tau-s) + W(\tau)\big)\dot{\check{x}}(s)ds \\ &= W(\tau)\underbrace{\left(\check{x}(0) - \check{x}(T_0) + \int_0^{T_0} \dot{\check{x}}(s)ds\right)}_{=0} + W^{(\tau)}(t-\tau) - \check{x}(T_0)W^{(\tau)}(t-T_0-\tau) \\ &\quad + \int_0^{T_0} W^{(\tau)}(t-\tau-s)\dot{\check{x}}(s)ds \\ &= W^{(\tau)}(t-\tau) - \check{x}(T_0)W^{(\tau)}(t-T_0-\tau) + \int_0^{T_0} W^{(\tau)}(t-\tau-s)\dot{\check{x}}(s)ds \\ &= \int_{\tau-T_0}^{\tau} \check{x}(\tau-u)dW^{(\tau)}(u) \qquad \text{for all } t \geq T_0.\end{aligned}$$

The sheer stopping-time property of τ suffices to settle the two essential points in the study of $\mathcal{J}^{(3)}(\tau)$, both of them contained in the *new-start property* of Brownian motion:

- The random variable τ and the process $W^{(\tau)}$ are actually independent,
- $W^{(\tau)}$ is a Brownian motion starting in zero.

Making use of that Brownian-motion new-start property and estimating the Gaussian integral provides

$$\begin{aligned}&\mathbb{P}\left\{\left|\int_{\tau-T_0}^{\tau} \check{x}(\tau-u)dW^{(\tau)}(u)\right| > \delta_3\right\} \\ &\quad = \int_{[T_0,T]} \mathbb{P}\left\{\left|\int_{t-T_0}^{t} \check{x}(t-u)\,dW(u)\right| > \delta_3 \,\middle|\, \tau = t\right\}\mathbb{P}\tau^{-1}(dt) \\ &\quad = \mathbb{P}\left\{\left|\int_{t-T_0}^{t} \check{x}(t-u)\,dW(u)\right| > \delta_3\right\} \leq \exp\left(-\frac{\delta_3^2}{\int_0^{T_0} \check{x}^2(u)du}\right).\end{aligned}$$

□

The main result of this section, Theorem 4.21, does not suggest particular choices of the involved parameters – and all that is evident so far is that there are parameter combinations that are useful in the sense that the probability for an exit up to time T can be achieved arbitrarily close to one. But the achieved results so far lack to prove that the special decomposition does more than only leading to additional terms that require concentration estimates each on their own. And that is the duty of the following corollaries, where we will show by means of deliberate choices that the result is capable of providing close-to-optimal estimates. But first, let us point out why it might be a fruitful attempt to compare the typical SDDE solution's first-exit-time behavior to Brownian motion. The first faint hint was given in Proposition 4.19, where we have seen that the variance process $(\operatorname{var} y(t))_{t\in[0,\infty)}$ behaves like

$$\operatorname{var} y(t) = \frac{T}{(1+ar)^2}\Big(1+\mathcal{O}(T^{-1})\Big) \quad \text{for big } T.$$

From that point was rather keen to propose the question in how far there might be further analogues to phenomena of a rescaled Brownian motion $\overline{W}(t) := W(t)/(1+ar), t \in [0,\infty)$. The apparently dissimilar stochastic differential law does not strengthened that suspicion. What we take as a second brief hint are the concentration inequalities for $(y(t))_{t\in[-r,\infty)}$ from the previous section, and which are actually surprisingly similar - at least to some extent - to the one we know from Brownian motion. And so the goal of this section is to study in how far the typical first-exit time of solution paths is similar to the first-exit-time behavior of Brownian motion in terms of small-ball probabilities. Regarding concentration inequalities it is convenient to study first-exit time distribution from a tube with diameter of a multiple of the standard deviation of the examined process. We carry this general idea over to the small-ball probabilities and observe in the case of a rescaled Brownian motion that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |\overline{W}(s)| > h\sqrt{\operatorname{var}\overline{W}(T)}\right\} \geq 1-\frac{4}{\pi}\exp\left(-\frac{\pi^2}{8h^2}\right). \tag{4.4.9}$$

Of course, this is only a trivial reformulation which relates times horizon and boundary. Together with the originally stated version of the small-ball probabilities for Brownian motion in (4.4.1), it covers the cases with radii hT^0 and $hT^{\frac{1}{2}}$. This motivates the slightly more general setting, where the boundary scales with T^α for some $\alpha \in [0,1/2]$.

In the case of rescaled Brownian motion $(W(t)/(1+ar))_{t\in[0,T]}$, result (4.4.1) implies for the first-exit-time distribution from a symmetric interval $[-hT^\alpha, hT^\alpha]$:

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |\overline{W}(s)| > \frac{h}{1+ar}T^\alpha\right\} \geq 1-\frac{4}{\pi}\exp\left(-\frac{\pi^2}{8h^2}T^{1-2\alpha}\right) \quad \text{for every } \alpha\in\mathbb{R}. \tag{4.4.10}$$

We restrict to the case $\alpha \in [0,\frac{1}{2}]$, because it covers the aspects that are mainly interesting for our purpose. The only additional assumption in the following corollary is the relatively weak requirement that the time horizon T is supposed to be big enough, and we will be rather explicit concerning the necessary size of T. With regard to the dependence between time horizon and boundary width for some $\alpha \in \mathbb{R}$, we consider time dependent quantities $\delta_1(T), \delta_2(T), \delta_3(T)$ that describe the tube width, and we maintain to write $\Delta_T = \delta_1(T) + \delta_2(T)+\delta_3(T)$. The subsequent corollaries are based on particular choices for those quantities depending on α. Besides the fact that the boundary parameters $\delta_1, \delta_2, \delta_3$ now depend on time, it is convenient to regard $h_0 = h_0(T)$ from (4.4.2) as time dependent. Then a reformulation of the main theorem reads

$$\begin{aligned} P_\star &:= \mathbb{P}\left\{ \sup_{s\in[0,T]} \frac{|y(s)|}{\sigma} > \frac{\delta_1(T)}{1+ar} \right\} \\ &\geq 1 - \frac{4}{\pi} \exp\left(-\frac{\pi^2}{8\Delta_T^2} T_1 \right) - \frac{5p^2}{2} \exp\left(-\frac{h_0^2}{2} \right) - \exp\left(-\frac{\delta_3^2(T)}{2v(T_0)} \right), \end{aligned}$$

and for easier comparing, we put a label to each of the bounds:

$$\overline{P}_1 := \frac{4}{\pi} \exp\left(-\frac{\pi^2}{8\Delta_T^2} T_1 \right), \quad \overline{P}_2 := \frac{5p^2}{2} \exp\left(-\frac{h_0^2}{2} \right), \quad \overline{P}_3 := \exp\left(-\frac{\delta_3^2(T)}{2v(T_0)} \right).$$

For all three corollaries, we choose $T_0 = \log T_1$, and for an ease of notations, we let $p = T_1 \in \mathbb{N}$ implicitly ignoring the integer-value restriction. The following constants will simplify the study of the relation between the different probabilities for given $\delta_1(T), \delta_2(T), \delta_3(T)$;

$$\begin{aligned} \hat{C}_0 &:= \frac{\pi^2}{8(\delta_1(T)+\delta_2(T)+\delta_3(T))^2}, \\ \hat{C}_1 &:= \sqrt{\frac{1}{2\kappa}} + \frac{\sqrt{T_1}}{\sqrt{p}\log p} + \frac{a(1+e^{\kappa r})}{\sqrt{2\kappa}} \frac{T_1}{2p\log p} = \frac{1}{\sqrt{2\kappa}} + \mathcal{O}\left(\frac{1}{\log T_1} \right), \\ \hat{C}_2 &:= \frac{5\pi p^2}{8} \in \mathcal{O}(T_1^2). \end{aligned}$$

Then, we may rewrite

$$\overline{P}_1 = \frac{4}{\pi} \exp\big(-\hat{C}_0 T_1 \big) \quad \text{and} \quad \overline{P}_2 = \frac{4}{\pi} \hat{C}_2 \exp\left(-\frac{\delta_2^2(T) e^{2\kappa T_0}}{2\hat{C}_1^2} \right).$$

Corollary 4.23. *In case $\alpha \in (1/4, 1/2)$, for arbitrary $h > 0$ we let $\delta_1(T) = hT_1^\alpha$, $\delta_2(T) = h\varepsilon_2 T_1^\alpha$, and $\delta_3(T) = h\varepsilon_3 T_1^\alpha$, where ε_2 and ε_3 are arbitrarily small positive constants. Assume $T = T_0 + T_1$ with $T_0 = \log T_1$ to be big enough such that the following properties hold;*

$$\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} + \frac{\log \hat{C}_2}{T_1^{1-2\alpha}} \leq \frac{h^2\varepsilon_2^2 T_1^{4\alpha-1} T_1^{2\kappa}}{2\hat{C}_1^2}, \tag{4.4.11}$$

$$\frac{T_1^{4\alpha-1}}{\log T_1} \geq \frac{\pi^2}{4h^4(1+\varepsilon_2+\varepsilon_3)^2\varepsilon_3^2} - \frac{2\log\frac{4}{\pi}}{h^2\varepsilon_3^2 T_1^{1-2\alpha}}. \tag{4.4.12}$$

Then the following concentration inequality holds true,

$$\mathbb{P}\left\{ \sup_{s\in[0,T]} \frac{|y(s)|}{\sigma} > \frac{h}{1+ar} T^\alpha \right\} \geq 1 - \frac{12}{\pi} \exp\left(-\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} T_1^{1-2\alpha} \right).$$

Proof. Note that condition (4.4.11) can be equivalently written as

$$\begin{aligned} &\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} + \frac{\log \hat{C}_2}{T_1^{1-2\alpha}} \leq \frac{h^2\varepsilon_2^2 T_1^{4\alpha-1} T_1^{2\kappa}}{2\hat{C}_1^2} \\ \Leftrightarrow \quad &\frac{\pi^2 T_1^{1-2\alpha}}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} \leq \frac{h^2\varepsilon_2^2 T_1^{2\alpha} e^{2\kappa \log T_1}}{2\hat{C}_1^2} - \log \hat{C}_2 \\ \Leftrightarrow \quad &\hat{C}_0 T_1 \leq \frac{\delta_2^2(T) e^{2\kappa T_0}}{2\hat{C}_1^2} - \log \hat{C}_2 \quad \Leftrightarrow \quad \overline{P}_1 \geq \overline{P}_2. \end{aligned}$$

And condition (4.4.12) implies

$$\begin{aligned}
&\frac{T_1^{4\alpha-1}}{\log T_1} \geq \frac{\pi^2}{4h^4(1+\varepsilon_2+\varepsilon_3)^2\varepsilon_3^2} - \frac{2\log\frac{4}{\pi}}{h^2\varepsilon_3^2 T_1^{1-2\alpha}} \\
\Leftrightarrow\quad &\frac{h^2\varepsilon_3^2 T^{4\alpha-1}}{2\log T_1} \geq \frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} - \frac{\log\frac{4}{\pi}}{T_1^{1-2\alpha}} \\
\Rightarrow\quad &\frac{h^2\varepsilon_3^2 T^{2\alpha}}{2v(T_0)} \geq \frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} T_1^{1-2\alpha} - \log\frac{4}{\pi} \quad\Leftrightarrow\quad \overline{P}_3 \leq \overline{P}_1,
\end{aligned}$$

where in the last step we have used that $v(T_0) \leq T_0 = \log T_1$. □

Corollary 4.24. *Let $\alpha \in (0, 1/4)$ and $\delta_1(T) = h\varepsilon_1 T_1^\alpha$, $\delta_2(T) = h\varepsilon_2 T_1^\alpha$, $\delta_3(T) = hT_1^\alpha$ for arbitrarily small constants $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$. We assume $T = T_0 + T_1$ with $T_0 = \log T_1$ big enough to satisfy*

$$\begin{aligned}
\log T_1 > \max\left\{\frac{5}{2\kappa}(1+ar)^2, 2\log(1+ar) + 4(1+ar)\right\}, \\
\frac{\pi^2 T_1^{1-4\alpha}}{8h^4(\varepsilon_1+\varepsilon_2+1)^2} - \frac{\log\frac{4}{\pi}}{T_1^{2\alpha}h^2} \geq \frac{(1+ar)^2}{\log T_1}, \\
\frac{\varepsilon_2^2 T_1^{2\kappa}}{2\hat{C}_1} - T^{-2\alpha}\log\frac{5T_1^2}{2} \geq \frac{(1+ar)^2}{\log T_1}.
\end{aligned}$$

Then,

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \frac{|y(s)|}{\sigma} > \frac{h}{1+ar} T^\alpha\right\} \geq 1 - 3\exp\left(-\frac{h^2 T_1^{2\alpha}}{2\log T_1}\right).$$

Proof. From the first condition on T_1 Proposition 4.19 is applicable and provides that

$$v(T_0) \geq \frac{\log T_1}{2(1+ar)}.$$

A straightforward reformulation of the second and third condition lead to $\overline{P}_1 \leq \overline{P}_3$ and $\overline{P}_2 \leq \overline{P}_3$ just as in the proof of the previous corollary. □

Corollary 4.25. *In case $\alpha = 1/4$ let $\delta_1(T) = hT_1^\alpha$, $\delta_2(T) = h\varepsilon_2 T^\alpha$ and $\delta_3(T) = h\varepsilon_3 T_1^\alpha \log T_1$ for arbitrary small $\varepsilon_2, \varepsilon_3 > 0$. Here we consider $T = T_0 + T_1$ with $T_0 = \log T_1$ to satisfy*

$$\begin{aligned}
\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3\log T_1)^2} \leq \frac{T_1^{2\kappa}\varepsilon_2^2 h^2}{2\hat{C}_1^2} - \frac{\log\hat{C}_2}{\sqrt{T_1}}, \\
\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3)^2} - \log\frac{4}{\pi} \leq \frac{\varepsilon_3^2 h^2\log T_1}{2}.
\end{aligned}$$

Then,

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \frac{|y(s)|}{\sigma} > \frac{h}{1+ar} T^\alpha\right\} \geq 1 - \frac{12}{\pi}\exp\left(-\frac{\pi^2}{8h^2(1+\varepsilon_2+\varepsilon_3\log T_1)^2} T_1^{1-2\alpha}\right).$$

Proof. As before, reformulating the conditions on T_1 yields $\overline{P}_1 \geq \overline{P}_2$ and $\overline{P}_1 \geq \overline{P}_3$. □

Remark 4.26. *a) It is worth mentioning that in the three above corollaries the respective conditions on T_1 are satisfied if only T is big enough, where we preferred to make the necessary size of T_1 rather explicit.*

b) Comparing the result of Corollary 4.23 and the reformulated Brownian motion's small-ball probabilites in (4.4.10) *shows that the main theorem actually provides useful results cherishing the decomposition method that we applied during the proof. In fact, the corresponding exponent from the rescaled Brownian motion case can be achieved up to arbitrary small correction in terms of the prefactor of* $T^{1-2\alpha}$*, which means that*

$$T_1^{1-2\alpha} = (T - \log T_1)^{1-2\alpha} = T^{1-2\alpha}\left(1 + \mathcal{O}\left(\frac{\log T_1}{T}\right)\right) \qquad \text{when } T \text{ is big.}$$

c) To some extend Corollary 4.24 shows the limit of the main theorem. When the boundary is chosen relatively small compared to the time horizon, we can no longer achieve an exponent that resembles the one of rescaled Brownian motion.

The implications of the main theorem have so far primarily aimed for best-possible exponents in concentration results while few attention has been payed to the prefactors. So far, a factor of 3 appears in the above corollaries which reflects the technique of only using the dominant term $\overline{P}_1$ or $\overline{P}_3$ as an upper bound for the other two occuring probabilities. But, this is no real issue because an additional factor $\log 3$ can easily be compensated in the exponent in all of the settings in corollaries 4.23 to 4.25. Only for $\alpha = 1/2$ this is no longer true since $T^{1-2\alpha} = 1$. And for that reason there is one more corollary to cover the special case $\alpha = 1/2$.

Corollary 4.27. *In case* $\alpha = 1/2$ *we let* $\gamma \in (0, 1/2)$ *and* $\delta_1(T) = h\sqrt{T_1}$, $\delta_2(T) = h\varepsilon_2 T_1^{\gamma}$, $\delta_3(T) = h\varepsilon_3 T_1^{\gamma}$*. Then,*

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \frac{|y(s)|}{\sigma} > \frac{h}{1+ar}\sqrt{T_1}\right\} \geq 1 - \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8h^2}\left(1 + \mathcal{O}\left(T_1^{\gamma-\frac{1}{2}}\right)\right)\right).$$

Proof. The proof is due to a couple of simple estimates:

$$\begin{aligned}
\frac{\overline{P}_2}{\overline{P}_1} &= \frac{5\pi T_1^2}{8}\exp\left(-\frac{h^2\varepsilon_2^2 T_1^{2\gamma+2\kappa}}{2\hat{C}_1^2} + \frac{\pi^2}{8h^2\left(1+\varepsilon_2 T_1^{\gamma-\frac{1}{2}} + \varepsilon_3 T_1^{\gamma-\frac{1}{2}}\right)^2}\right)\\
&= \exp\left(-\frac{h^2\varepsilon_2^2 T_1^{2\gamma+2\kappa}}{2\hat{C}_1^2} + \log\left(\frac{5\pi T_1^2}{8}\right) + \frac{\pi^2}{8h^2\left(1+\varepsilon_2 T_1^{\gamma-\frac{1}{2}} + \varepsilon_3 T_1^{\gamma-\frac{1}{2}}\right)^2}\right),
\end{aligned}$$

$$\begin{aligned}
\frac{\overline{P}_3}{\overline{P}_1} &= \exp\left(-\frac{h^2\varepsilon_3^2 T_1^{2\gamma}}{2v(T_0)} + \log\left(\frac{4}{\pi}\right) + \frac{\pi^2}{8h^2\left(1+\varepsilon_2 T_1^{\gamma-\frac{1}{2}} + \varepsilon_3 T_1^{\gamma-\frac{1}{2}}\right)^2}\right)\\
&\leq \exp\left(-\frac{h^2\varepsilon_3^2 T_1^{2\gamma}}{2\log T_1} + \log\left(\frac{4}{\pi}\right) + \frac{\pi^2}{8h^2\left(1+\varepsilon_2 T_1^{\gamma-\frac{1}{2}} + \varepsilon_3 T_1^{\gamma-\frac{1}{2}}\right)^2}\right).
\end{aligned}$$

Let us for a moment denote $\xi := \varepsilon_2 T_1^{\gamma-\frac{1}{2}} + \varepsilon_3 T_1^{\gamma-\frac{1}{2}}$ and $\zeta = \frac{\overline{P}_2}{\overline{P}_1} + \frac{\overline{P}_3}{\overline{P}_1}$, then, by means of a Taylor expansion, we reformulate the leading term

$$\begin{aligned}
\frac{1}{(1+\xi)^2} &= 1 - 2\xi + \mathcal{O}(\xi^2) \quad \text{for small } \xi, \text{ i.e. big } T,\\
\log(1+\zeta) &= \zeta + \mathcal{O}(\zeta^2) \quad \text{for small } \zeta, \text{ i.e. big } T.
\end{aligned}$$

Combining the estimates we find that

$$\begin{aligned}\overline{P}_1 + \overline{P}_2 + \overline{P}_3 &= \frac{\pi}{4} \exp\left(-\frac{\pi^2}{8h^2(1+\xi)^2}\right)(1+\zeta) \\ &= \frac{\pi}{4} \exp\left(-\frac{\pi^2}{8h^2}\big(1 - 2\xi + \mathcal{O}(\xi^2)\big) + \log(1+\zeta)\right).\end{aligned}$$

And as clearly $\zeta = \mathcal{O}(T^{\gamma - \frac{1}{2}})$, the claim follows. □

Remark 4.28. *Keeping in mind that* $\sqrt{\operatorname{var} y(T_1)} \sim \frac{\sigma}{1+ar}\sqrt{T_1}$ *(see Proposition 4.19), and* $\sqrt{T_1} = \sqrt{T}\left(1 + \mathcal{O}\left(\frac{\log T_1}{T_1}\right)\right)$*, the above corollary beautifully resembles the according small-ball probability of Brownian motion in* (4.4.9).

5. From Uniform Stability to Instability

The variety of RFDEs is far too rich to reasonably wish for a uniform discription of typical stochastically perturbed path behavior in terms of concentration inequalities in general. We will focus on a particular generalization of potential-driven SDEs which are not only subject to an instantaneous potential-induced feedback, but also *feel* the time-delayed feedback coming from a second, possibly different potential. For one thing this constitutes a straightforward generalization of the classical potential-driven SDEs and was considered e.g. in [FI05], and for another thing, it is also the obvious generalization of SDDEs from Chapter 4 to time dependence and non-linearity. There are two applications that serve as paragons for our study. The first one is a variant of the linear SDDE (4.0.1), which we studied for constant coefficients in Section 4. We will equip the differential law with time-dependent coefficients that slowly travel out of the area of stability S which we sketched in Figure 4. The second is the symmetric pitchfork bifurcation. Details are presented in the Example 5.32.

5.1. Setting and the Replacement System

We will generally keep the assumption of a fixed delay length, although it is not strictly necessary. We will consider the topic in hindsight in Subsection 5.3.5 to give some details why and how the derived results can be directly extended to more general delay feedback. In general, fixing the time delay $r > 0$, we end up with systems of the form

$$\begin{cases} dx(t) = \tilde{f}(x(t), t, x(t-r), t-r)dt + \sigma dW(t) & \text{for } t \in [0, \tau), \\ x_0 = \Upsilon, \end{cases} \tag{5.1.1}$$

where we assume that $\Upsilon \in \mathcal{C}(J, \mathbb{R})$, $\tau := \inf\{t \geq 0 : (x(t), t, x(t-r), t-r) \notin \tilde{\mathbb{D}}\}$ for some appropriate domain $\tilde{\mathbb{D}} \subset \mathbb{R} \times [0, \infty) \times \mathbb{R} \times [0, \infty)$. As usual, random perturbation through Brownian motion is scaled by the factor $\sigma > 0$. The systemic influence from the current position and delay-related influence are supposed to add up, which means we assume that the drift term in (5.1.1) can be represented as

$$\tilde{f}(x, t, y, t-r) = f(x, t) + g(y, t) \quad \text{for all } (x, t, y, t-r) \in \tilde{\mathbb{D}}. \tag{5.1.2}$$

For notational purpose, we reduce the time dependence of both coefficients to t. This might seem to be a simplification, but as long as the delay r is deterministic, this is only a matter of definition. We conveniently assume that the coefficients f and g share the same domain. And that common domain $\mathbb{D} \subset \mathbb{R} \times [0, \infty)$ of f and g together with the two mappings f and g are assumed to fulfill a catalog of conditions that we are going to present and motivate below. Specific assumptions that rely on the state of the transition phase will be added in the corresponding subsequent subsections. The reason why we do not phrase all conditions at once, but only reinforce them stepwisely is this: We will consider more general situations for instance in the uniformly stable phase, while our techniques for the transition phase require stricter assumptions. Those crucially depend on the nature of the very point where systemic properties change. Those properties are usually linked to different types of bifurcation points.

- The coefficient functions f and g are supposed to be twice continuously differentiable in their spatial argument and also twice continuously differentiable in their time argument, and their derivatives bounded in $\mathbb{D}$. The set of those functions is denoted by $\mathcal{C}_b^{2,2}$, and we use the common short-hand notations

$$h_x(x,t) = \frac{\partial}{\partial x}h(x,t), \quad h_{xx}(x,t) = \frac{\partial^2}{\partial x \partial x}h(x,t), \quad h_t(x,t) = \frac{\partial}{\partial t}h(x,t) \quad \text{for all } x, t,$$

and in the same way the mixed derivatives are denoted $h_{xt} = h_{tx}$ are at least continuous and bounded.

- Bounded continuous differentiability of $f(\cdot,\cdot), g(\cdot,\cdot)$ ensures the existence of ((x,t)-uniform) constants

$$\begin{aligned} m_f &:= \sup_{(x,\nu t)\in\mathbb{D}} |f_t(x,t)| < \infty, \\ m_g &:= \sup_{(x,\nu t)\in\mathbb{D}} |g_t(x,t)| < \infty, \\ d_g &:= \sup_{(x,\nu t)\in\mathbb{D}} |g_x(x,t)| < \infty, \\ d_f &:= \sup_{(x,\nu t)\in\mathbb{D}} |f_x(x,t)| < \infty. \end{aligned}$$

  We denote $M := m_f + m_g$.

We assume that the system evolves slowly in time and we will control that speed with a small parameter $\nu > 0$. This assumption is natural, since there is no hope for path-concentration behavior if the system changes wildly. Thus, based on (5.1.2), we will consider systems of the form

$$\begin{cases} dx(t) = \Big[f(x(t),\nu t) + g(x(t-r),\nu t)\Big]dt + \sigma dW(t) \quad \text{for } t \in [0, T/\nu], \\ x_0 = \Upsilon. \end{cases} \tag{5.1.3}$$

It is worth mentioning that the random noise term is unaffected by this modification.
The one-dimensional setting generally allows for an interpretation of (5.1.3) as potential-driven differential law. To this end denote $F(x,\nu t) := \int_0^x f(u,\nu t)du$ and $G(x,\nu t) := \int_0^x g(u,\nu t)du$ for all $(x,\nu t) \in \mathbb{D}$. Then, the differential law (5.1.3) can be written as

$$dx(t) = -\nabla_x F(x(t),\nu t)dt - \nabla_x G(x(t-r),\nu t)dt \quad \text{for all } t \in [0, T/\nu].$$

Here $\nabla_x F$ for instance denotes the derivative of F with respect to the first argument. In particular, slow evolution in t visually means that the steepest slope of the potentials $F(x,\nu t)$ and $G(x,\nu t)$ with respect to time scales with $\nu > 0$, i.e.

$$\sup_{(x,\nu t)\in\mathbb{D}} \left|\frac{d}{dt}f(x,\nu t)\right| + \sup_{(x,\nu t)\in\mathbb{D}} \left|\frac{d}{dt}g(x,\nu t)\right| \le \nu M.$$

As a key tool for the description we introduce the *replacement* (ordinary) *SDE*

$$\begin{cases} d\overline{x}(t) = f(\overline{x}(t),\nu t)dt + g(\overline{x}(t),\nu t)dt + \sigma dW(t) \quad \text{for } t \in [0, T/\nu), \\ \overline{x}(0) = \overline{\Upsilon}(0) \in \mathbb{R}. \end{cases} \tag{5.1.4}$$

Through a time change $\nu t = s$, we receive an equivalent *fast-time formulation* of the system

(5.1.4):

$$\begin{cases} \nu\, d\overline{x}(s/\nu) = f(\overline{x}(s/\nu), s)ds + g(\overline{x}(s/\nu), s)ds + \frac{\sigma}{\sqrt{\nu}} d\tilde{W}(s) & \text{for } s \in [0, T), \\ \overline{x}(0) = \overline{\Upsilon}(0). \end{cases} \tag{5.1.5}$$

This is convenient, because accelerated Brownian motion $(W(t/\nu))_{t\geq 0}$ has the same distribution as the rescaled Brownian motion $(\nu^{-\frac{1}{2}} W(t))_{t\geq 0}$. The notation $\tilde{W}$ emphasizes that it is not the original Brownian motion that we consider in that place. Setting $\sigma = 0$, the differential law of (5.1.5) can be regarded as a *slow-fast system* via

$$\begin{cases} \left.\begin{aligned} \nu\, \dot{\overline{x}}(t) &= f(\overline{x}(t), y(t)) + g(\overline{x}(t), y(t)), \\ \dot{y}(t) &= \nu, \end{aligned}\right\} & \text{for } t \geq 0, \\ \text{with initial conditions:} \quad \overline{x}(0) = \overline{\Upsilon}(0), \ y(0) = 0. \end{cases} \tag{5.1.6}$$

The multiple reformulation of the system provides that we can acquire results from slow-fast dynamical systems to learn about the intuitively arranged System 5.1.3. It is the central point of the subsequent sections to specify and verify that. We continue with the definition and notations regarding equilibrium branches and adiabatic solutions. We refer to the collection of (x,t)-tuples from $\mathbb{D}$ that satisfy $f(x,t) + g(x,t) = 0$ as the slow (replacement) manifold

$$\mathcal{M} := \{(x, \nu t) \in \mathbb{D} : f(x, \nu t) + g(x, \nu t) = 0\}.$$

We assume that there is a continuous equilibrium branch $(x^\star(t), \nu t)_{t\in[0,T/\nu]}$ that lies in $\mathcal{M}$ and that the potential curvature $A^\star(t) := f_x(x^\star(t), \nu t) + g_x(x^\star(t), \nu t)$ along that path satisfies:

$$A^\star(t) \begin{cases} < 0 & \text{for all } t < \tilde{T}_2/\nu, \\ = 0 & \text{if and only if } t = \tilde{T}_2/\nu, \\ > 0 & \text{for all } t > \tilde{T}_2/\nu. \end{cases}$$

Here $\tilde{T}_2$ is chosen independently of ν. We will refer to $A^\star(t)$, $t \in [0, T/\nu]$, as the *stability matrix* and it may be interpreted as the *curvature* of the potential $(F+G)(x, \nu t)$ along the equilibrium branch $x^\star$. We identify the continuous equilibrium branch $x^\star = (x^\star(t), \nu t)_{t\in[0,T/\nu]}$ with the process $x^\star = (x^\star(t))_{t\in[0,T/\nu]}$ referring to both of them simply using the declaration $x^\star$.

- The $\mathcal{C}_b^{2,2}$-condition on f and g ensures that intersections of equilibrium branches can only occur in T_2/ν due to the implicit-functions theorem. It provides that the equilibrium branch $x^\star$ is $\mathcal{C}^1$ and

$$\frac{d}{dt} x^\star(t) = \Big(A^\star(t)\Big)^{-1} \cdot \Big(\nu (f_t + g_t)(x^\star(t), \nu t)\Big) \quad \text{for all } t \neq \tilde{T}_2/\nu, \tag{5.1.7}$$

  where $t \neq \tilde{T}_2/\nu$ ensures that the stability matrix satisfies $A^\star(t) \neq 0$. This equilibrium-branch concept does neither depend on time-delayed influence nor on any initial segment, and is therefore a suitable choice as a reference.

- To avoid technical issues, we assume that $\mathbb{D}$ generally has sufficiently nice properties, i.e.
  - ▷ For sufficiently small $\nu > 0$, the equilibrium branch $x^\star$ remains inside $\mathbb{D}$, bounded away from the boundary $\partial\mathbb{D}$.
  - ▷ There are constants $R_-, R_+ > 0$ such that $[-R_-, R_-]\times[0, T/\nu] \subset \mathbb{D} \subset [-R_+, R_+]\times[0, T/\nu]$.
  - ▷ To have first-exit-times be stopping times, we further assume $\mathbb{D}$ to have a sufficiently nice boundary.

We will identify the uniformly stable phase with the slow time interval $[0, T_0/\nu]$. In contrast to $\tilde{T}_2, \tilde{T}_3$ there is no canonical choice for T_0. To motivate the concept of a *ν-adiabatic solution* we proceed with a review of the *existence of ν-adiabatic solutions* and the *slaving principle* through the following remark:

Remark 5.1 (Review). *In the above situation suppose that the stability (scalar) matrix is negative and bounded away from zero over $[0, T_0/\nu]$. Then, if the deterministic counterpart of the differential law of* (5.1.4)*, i.e.*

$$\left\{ d\overline{x}(t) = f(\overline{x}(t), \nu t)dt + g(\overline{x}(t), \nu t)dt \quad \textit{for } t \in [0, T_0/\nu), \right. \tag{5.1.8}$$

evolves slow enough, i.e. ν is small enough, there exists a solution $(x^{\mathrm{ad}}(t))_{t\in[0,T_0/\nu]}$ of (5.1.8) *that remains close, which means at a distance of order ν, to the equilibrium branch. Such a particular solution will be called a ν-adiabatic solution. Furthermore there is an open environment $E_\nu \subset \mathbb{D}$, that contains $x^\star$, and all solution paths that are initiated at some $(x, \nu t) \in E_\nu$ run exponentially fast to the adiabatic solution. This phenomenon is sometimes referred to as the slaving principle. Its establishment goes back to the work of Tikhonov [Tih52] and Gradšteĭn [Gra53]. The existence of an adiabatic manifold in a close neighborhood of the slow manifold follows from Fenichel's geometric approach to perturbation theory [Fen79], see also [BG06] or [Kue15] for details and further references. As indicated, the listed literature provides the result for the fast-time variant* (5.1.6) *of the system.*

As another convention we will assume that the equilibrium branch and the ν-adiabatic solution exist over the whole interval $[-r, T/\nu]$.

Remark 5.2. *In principle, we think of the system to exist from time $-r$ on, while the observation time starts in* 0*, and this way we compensate the interpretational dilemma that arises from the need of an initial segment to have the retarded differential equation well-defined.*

Regarding the adiabatic solution and the instantaneous feedback, the following properties are supposed to hold. The first one is a restriction especially on the bifurcation. The second requires a positive curvature of the instantaneous potential F. The third one is a minimal assumption on the domain.

- x^{ad_ν} is a ν-adiabatic solution throughout $[-r, T/\nu]$. Over $[-r, 0]$ this is the above stated convention. Due to the Fenichel theory it is naturally given over $[0, T_0/\nu]$, but, apart from that, it is an assumption that partly characterizes the form of transitions

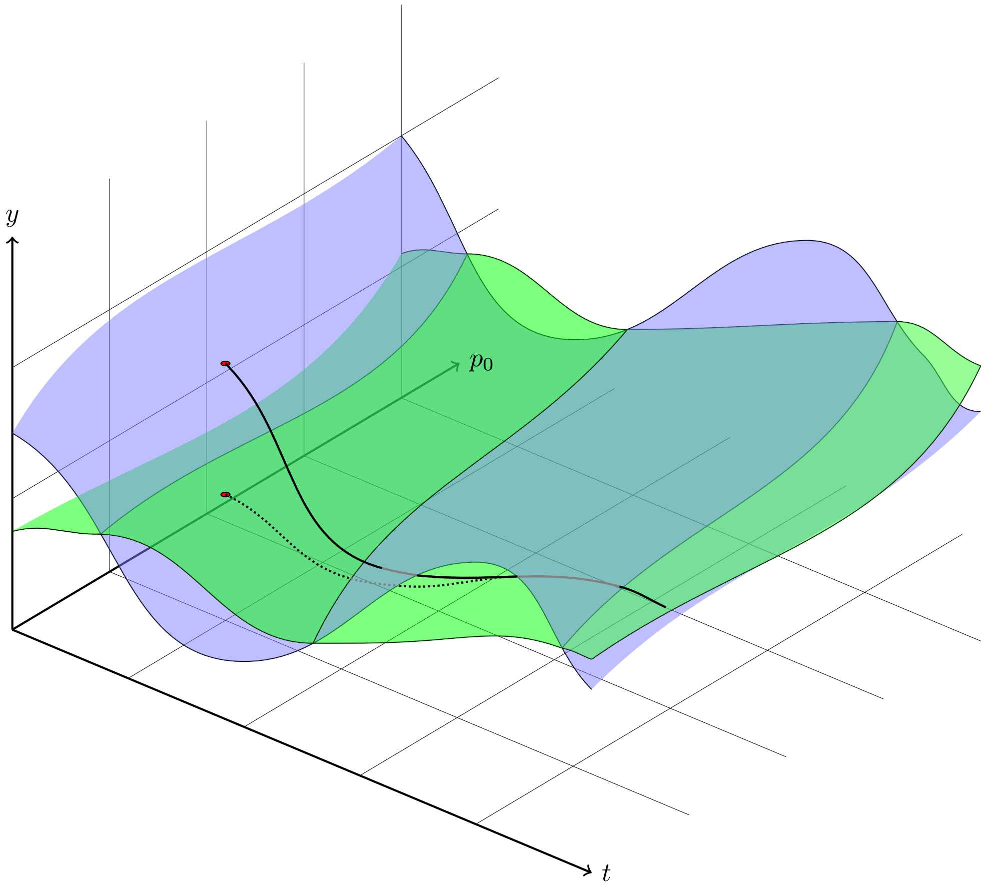

Figure 7: An illustration of the slaving principle, where p_0 symbolically stands for some abstract parameter. The blue surface represents a slow stable manifold, the green one the adiabatic manifold which lies in a close neighborhood of the slow manifold. If the time is rescaled by a small factor ν, then there is an adiabatic manifold within an $\mathcal{O}(\nu)$-neighborhood of the slow manifold. The dotted line represents a particular solution that lies on the adiabatic manifold and which is therefore referred to as adiabatic solution. The full line indicates how an arbitrary solution is *attracted* by a corresponding adiabatic solution.

that we are going to study and describe. This assumption is satisfied for instance if the potential $F + G$ features a symmetric pitchfork bifurcation, while it is not in the asymmetric case.

- The potential $F(x, \nu t)$, that refers to the instantaneous feedback, is stabilizing along the equilibrium branch with a curvature that is bounded away from 0, which means that

$$f_x(x^\star(t), \nu t) < -\tilde{a}_- < 0 \quad \text{for all } t \in [0, T/\nu]. \tag{5.1.9}$$

- We assume that $\nu > 0$ is sufficiently small that, next to the continuous equilibrium branch $x^\star$, also the ν-adiabatic solution x^{ad_ν} remains in $\mathbb{D}$, bounded away from its boundary $\partial\mathbb{D}$. We use the short-hand notations

$$\begin{aligned} a(t) &:= -f_x(x^{\mathrm{ad}_\nu}(t), \nu t), \\ b(t) &:= g_x(x^{\mathrm{ad}_\nu}(t - r), \nu t) \quad \text{for all } t \in [0, T/\nu]. \end{aligned} \tag{5.1.10}$$

We further assume that $a(\cdot) > a_- > 0$ for all $t \in [0, T/\nu]$ and some $a_- > 0$. This assumption is generally satisfied for arbitrary $a_- < \tilde{a}_-$ if ν is small enough, because of the preceding item of this list.

The *initial maximal distance* between the solution $(x(t))_{t\in[0,T/\nu]}$ of (5.1.3) and the adiabatic solution over $[-r, 0]$ is denoted by

$$\sup_{t\in[-r,0]} |x(t) - x^{\mathrm{ad}_\nu}(t)| =: \|\Upsilon_0\| < \infty, \tag{5.1.11}$$

and $\|\Upsilon_0\|$ is assumed at least not to be *too big*. Honestly, this is an assumption for the comfort of notation. In general there is no convenient concept to describe the initial influence of the delayed feedback when a system start is considered in $t = 0$, and there are possibly different interpretations for the delayed feedback on $[0, r]$, when the system originates only in $t = 0$. We are going to specify the respective conditions later on. Regarding the nonlinearity, we assume that there are constants $N_g, N_f > 0$ and *remainders* $\mathcal{R}_f, \mathcal{R}_g : \mathbb{R} \times [0, T] \to \mathbb{R}$ such that

$$\begin{aligned} f(x^{\mathrm{ad}_\nu}(t) + y, \nu t) &= f(x^{\mathrm{ad}_\nu}(t), \nu t) + f_x(x^{\mathrm{ad}_\nu}(t), \nu t)y + \mathcal{R}_f(y, \nu t) \quad \text{with} \\ &|\mathcal{R}_f(y, \nu t)| \le N_f y^2 \quad \text{for all } (x^{\mathrm{ad}_\nu}(t) + y, \nu t) \in \mathbb{D}, \\ g(x^{\mathrm{ad}_\nu}(t-r) + y, \nu t) &= g(x^{\mathrm{ad}_\nu}(t-r), \nu t) + g_x(x^{\mathrm{ad}_\nu}(t-r), \nu t)y + \mathcal{R}_g(y, \nu t) \quad \text{with} \\ &|\mathcal{R}_g(y, \nu t)| \le N_g y^2 \quad \text{for all } (x^{\mathrm{ad}_\nu}(t-r) + y, \nu t) \in \mathbb{D}, \end{aligned}$$

and for convenience, denote $N := N_f + N_g$. We are going to continuously track the role of the nonlinearity although it is neither supposed to be big nor to be small. In order to do so, we will explicitly list the N-terms in the Landau symbols $\mathcal{O}(\cdot)$. We will generally assume that $\nu N < 1$.

The collection of properties, that we have just stated, concerning the existence of an equilibrium branch and a ν-adiabatic solution, the appropriate domain, slow time, and so forth does really only depend on the concept of the replacement system (5.1.4). The whole setting is arranged in a manner that makes established results applicable. That includes the existence of the ν-adiabatic solution as well as the techniques of [BG06] providing concentration inequalities concerning the replacement system. But, roughly speaking the approach bears the flaw that the delay-influenced underlying deterministic version of the system (5.1.3) and the constructed ν-adiabatic solution do not share the same differential, not even up to nonlinearity. The reason why we *did not choose* some appropriate adiabatic solution with respect to the delay-influenced law is the following:

- The *equilibrium branch* is no longer uniquely defined. Proceeding as before, i.e. choosing

  $$x^\star(t) \text{ such that } f(x^\star(t), t) + g(x(t-r)^\star, t-r) = 0,$$

  leads to an equilibrium-branch concept that depends on an initial segment and there is no ultimately convenient choice for that. For example one might take the respective initial segment $\Upsilon \in \mathcal{C}([-r, 0])$, but such a path dependence is highly undesirable for an attempt of a uniform description of sample paths.

Status report. *What we have arranged is a setting that features a well-defined equilibrium branch concept. And based on that we have established the existence of a ν-adiabatic solution in uniform stable environment and, by assumption, over the whole time interval* $[0, T/\nu]$. *There are well-known systems of potential driven SDEs that satisfy these assumptions; or at least that have a decomposition of f and g such that the assumptions are satisfied. Furthermore, the concentration behavior of solutions can be comfortably compared to the related results established in the literature. Altogether, the consideration of an adiabatic solution of the delay-free replacement system provides numerous beautiful properties that make it an ideal reference object. In that regard, the delay-influence term that originally acts on the differential law* (5.1.3) *is understood as a perturbation of the nonlinear replacement system* (5.1.4).

Remark 5.3. *The different phases of transition will be characterized through time points* $\frac{T_0}{\nu}, \frac{T_1}{\nu}, \frac{T_2}{\nu}, \ldots$, *which are related to the potential curvature along a ν-adiabatic solution. But, as the ν-adiabatic solution is not unique, this leads to a blur in the definition of time points. This inaccuracy will be only of order* 1 *in slow time and of order ν in the fast-time formulation, for instance:*

$$\begin{aligned} \tilde{T}_2 &= \inf\{t \in [0, T/\nu] : f(x^\star(t), \nu t) + g(x^\star(t), \nu t) = 0\}, \\ T_2 &:= \inf\{t \in [0, T/\nu] : f(x^{\mathrm{ad}_\nu}(t), \nu t) + g(x^{\mathrm{ad}_\nu}(t-r), \nu t) = 0\}. \end{aligned} \tag{5.1.12}$$

Then, under the given assumptions $\tilde{T}_2 - T_2 = \mathcal{O}(\nu)$. *And therefore, we will neglect it and proceed as if the ν-adiabatic solution as well as the corresponding points in time were uniquely defined.*

5.1.1. Justification for the Approach

The previous subsection has presented a comfortable basic setting serving a ν-adiabatic solution that provides an excellent basis for comparisons with deterministic and stochastically perturbed systems in the delay-free case. This subsection is solely devoted to the presentation of the results that will be achieved in the course of the subsequent sections. Thereby, it also serves as a justification. Roughly, the subsequent study is divided into two parts: The first part describes the uniformly stable phase in Subsection 5.2, while the actual transition phase is studied in the second part in Subsection 5.3. Due to the necessary reinforcements of restrictions, the two parts have been separated. Figure 8 serves an illustration of the different transition phases an typical pathwise behavior. In order to simplify comparisons with the results of [BG06], all time lengths are stated in fast time, although the description as well as the subsequent validation of results will use the slow-time formulation.

- **Initial layer.** In practice, concentration properties can sometimes only be established after the system has *cooled down*, which means after initial conditions have been mostly compensated and the system has approached some kind of invariant state. In that situation the initial time interval that is needed for relaxation is regularly referred to as an *initial layer*. Our study begins in Subsection 5.2 with the system in a uniformly stable regime, especially because it is the only phase of the transition, where we can deal with initial-layer phenomena in principle. We will show that during the uniformly stable phase $[0, T_0/\nu]$, a solution path initiated at a distance up to order 1 from the

ν-adiabatic solution x^{ad_ν} typically approaches x^{ad_ν} up to a size of order ν within a fast time of order $\sqrt{\nu}|\log\nu|$ with high probability if only $\sigma < \frac{\nu}{|\log\nu|}$. In the delay-free case, this cool-down time typically is of order $|\nu\log\nu|$.

- **Uniformly stable phase.** Once the process has entered an environment of order ν around x^{ad_ν}, we will show that

$$\mathbb{P}\left\{\sup_{s\in[0,T_0/\nu]}\frac{|y(s)|}{\sqrt{\zeta(s)}} > h\right\} \leq \frac{T_0}{\nu^2}\exp\left(-(1-\gamma)\frac{h^2}{4\sigma^2}\Big(1+\mathcal{O}(\nu|\log\nu|)\Big)\right),$$

  where $y(t) = x(t) - x^{\mathrm{ad}_\nu}(t)$, $\zeta(t) = \frac{1}{2(a(t)-|b(t)|)}$ with $a(\cdot)$ and $b(\cdot)$ defined in (5.1.10) and $1-\gamma := \min_{t\in[0,T_0/\nu]} 1-\frac{|b(t)|}{a(t)}$. It is worth emphasizing that ζ is defined with regard to the delay-free replacement system, which provides easier comparability. Further, the h must be chosen at least of order ν. The applied approach is closely related to the method from [BG06, Proposition 3.1.5] which we will call Bernstein-based approach, see also A.1. In contrast to the delay-free case, the exponent features an additional factor $\frac{1-\gamma}{2}$, and this one can not be improved further than to $1-\gamma$ with the method we apply.

- **Shallow curvature.** The above introduced γ can naturally be seen as a measure of stability; details are given in Theorem 5.4 below. We consider the fast-time point T_1 as the point where $1-\frac{|b(T_1/\nu)|}{a(T_1/\nu)} = \sqrt{\nu}|\log\nu|$. We will show that a typical solution path remains within a distance of order $\sqrt{\nu}$ up to T_1 in Subsection 5.3.1.

- **Loss of stability, increasing instability.** As we mentioned before, the path behavior through the bifurcation point T_2/ν, defined in (5.1.12), crucially depends on the kind of bifurcation. Considering symmetry conditions similar to the symmetric pitchfork, where the equilibrium branch and the ν-adiabatic solution coincide and form a flat line through the bifurcation point, we will show that a typical solution path remains in an environment of order $\sqrt{\nu}$ at least for a fast time $T_3 - T_2$ of order $\sqrt{\nu}$ after the bifurcation point. In particular, quadratic nonlinearity is assumed to vanish and to be of order at most $\sqrt{\nu}|\log\nu|$ in $[T_1, T_2]$, and at most of order $\sqrt{\nu}$ in $[T_2, T_3]$. This transition is divided into two parts considered in Subsection 5.3.2 and Subsection 5.3.3.

- **Early exit.** Subsection 5.4 deals with the question how long a solution path typically needs to exit from a neighborhood of order 1 from the unstable equilibrium branch. Neglecting nonlinear terms in the differential law, a transformation to a nonautonomous analogue of the critical regime is established. Under further simplifications we will show that typically a fast time $T - T_3$ of order $\sqrt{\nu}|\log\sigma|$ suffices for the path to leave an environment of order 1 around the adiabatic solution.

In Subsection 5.3.4 we consider a further reinforcement of assumptions and confine to the special case of uniformly symmetric potentials F and G rather than assuming that they are only symmetric at the transition time T_2. That implies the absence on quadratic nonlinear influence and provides a significant improvement in the description of the phase between uniformly stable phase and the shallow-curvature phase $[T_0, T_1]$.

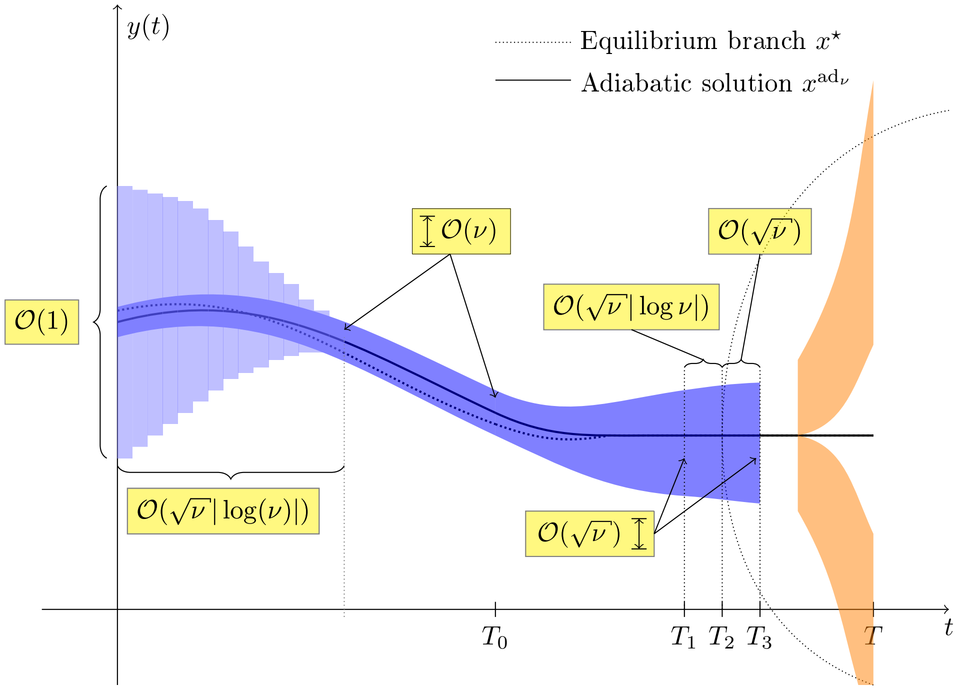

Figure 8: Sketch of the typical path behavior during different transition phases. All time points and durations are labeled in fast-time.

5.2. Uniform Stability

In order to establish concentration phenomena for the system (5.1.3) in the uniformly stable phase $[0, T_0/\nu]$ we start with the consecutive-boxes approach in Subsection 5.2.1 that especially provides a treatment for the solutions that start at some distance from the equilibrium branch, and the adiabatic solution respectively. The resulting typically contractional behavior is represented in the Figure 8 through the light-blue double-sided environment of boxes centered around x^{ad_ν} and the clear blue tube of size ν. Once the initial conditions have been mostly compensated through consecutive boxes, through the Bernstein-based approach we present a technique similar to [BG06] in the subsequent part 5.2.2.

As usual uniform stability is characterized by the property that all eigenvalues of the stability matrix have negative real parts and these are bounded away from zero. In our simple dimension-one case, the *stability scalar (matrix)* at $(x^\star(s), \nu s)$, given by $A^\star(s) = f_x(x^\star(s), \nu s) + g_x(x^\star(s), \nu s)$, has an easy form; the only eigenvalue of $A^\star(s)$ is its very value. Therefore, the uniform stability is characterized by the existence of some $\kappa > 0$, such that

$$A^\star(s) < -\kappa < 0 \quad \text{for all } s \in [-r, T_0/\nu], \tag{5.2.1}$$

and κ must be independent of ν for sufficiently small ν. That prevents us for instance from choosing T_0 such that $A^\star(s) = \mathcal{O}(\nu)$. As we reviewed in some detail in Remark 5.1, through [BG06, Theorem 2.1.8 (Existence of an adiabatic manifold)] or [Kue15, Theorem 3.1.4 (Fenichel's theorem)] the uniform stability assumption (5.2.1) provides the existence of a ν-adiabatic solution $x^{\mathrm{ad}_\nu}(t) = x^\star(t) + \mathcal{O}(\nu)$. In other words, for all $\nu \in (0, \nu_0]$, there is a constant $\delta = \delta(\nu_0) > 0$, that only depends on ν_0, such that

$$\left|x^{\mathrm{ad}_\nu}(t) - x^\star(t)\right| \leq \delta\nu \qquad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.2}$$

As a notational convention we will generally assume that a ν-adiabatic solution satisfies (5.2.2). Uniform stability of the equilibrium branch is the key to path concentration in the delay-free replacement system given that time evolves sufficiently slow. In order to achieve analogue results in the delay-influenced case, we need the additional assumption that

$$|b(t)| < a(t) \qquad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.3}$$

That kind of assumption is unnecessary when only considering the replacement solution. In particular, a big negative $b(\cdot)$ is advantageous for the stability there, but regarding the delay-influenced case, a negative $b(\cdot)$ is no longer necessarily tame and welcome, because if it is sufficiently big, it will trigger oscillations with exponentially increasing amplitude.

Proposition 5.4 (Characterization of (5.2.1)&(5.2.3))**.** *Consider the setting of Subsection 5.1. For sufficiently small $\nu > 0$, the uniform stability condition* (5.2.1) *together with* (5.2.3) *is equivalent to the existence of some $\bar{\gamma} > 0$, independent of ν, such that*

$$1 - \max_{s \in [0, T_0/\nu]} \frac{|b(s)|}{a(s)} > \bar{\gamma}. \tag{5.2.4}$$

Proof. By slow evolution of the equilibrium branch through (5.1.7) we also have slow evolution of the adiabatic solution x^{ad_ν}, i.e. $|x^{\mathrm{ad}_\nu}(t) - x^{\mathrm{ad}_\nu}(t-r)| = \mathcal{O}(\nu)$ for $t \in [r, T_0/\nu]$. Then uniform continuity of f_x and g_x provide that

$$\left.\begin{array}{l} |f_x(x^\star(t), \nu t) - f_x(x^{\mathrm{ad}_\nu}(t), \nu t)| = \mathcal{O}(\nu), \\ |g_x(x^\star(t), \nu t) - g_x(x^{\mathrm{ad}_\nu}(t-r), \nu t)| = \mathcal{O}(\nu) \end{array}\right\} \quad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.5}$$

And hence, the uniform hyperbolicity assumption (5.2.1) and (5.2.3) imply that for sufficiently small ν there is $\bar{\kappa} > 0$ such that

$$\begin{aligned} & f_x(x^{\mathrm{ad}_\nu}(t), \nu t) + |g_x(x^{\mathrm{ad}_\nu}(t-r), \nu t)| < -\bar{\kappa} \\ \Leftrightarrow \quad & |f_x(x^{\mathrm{ad}_\nu}(t), \nu t)| - |g_x(x^{\mathrm{ad}_\nu}(t-r), \nu t)| > \bar{\kappa} \\ \Leftrightarrow \quad & 1 - \frac{|g_x(x^{\mathrm{ad}_\nu}(t-r), \nu t)|}{|f_x(x^{\mathrm{ad}_\nu}(t), \nu t)|} > \frac{\bar{\kappa}}{|f_x(x^{\mathrm{ad}_\nu}(t), \nu t)|} \quad \text{for all } t \in [0, T_0/\nu]. \end{aligned}$$

The final term in the above equivalence relation is in $(0,1)$, because f_x is bounded. Therefore, for sufficiently small ν, there is

$$\bar{\gamma} = 1 - \inf_{t \in [0, T_0/\nu]} \frac{\bar{\kappa}}{|f_x(x^{\mathrm{ad}_\nu}(t), \nu t)|} \in (0,1)$$

such that

$$\frac{|b(t)|}{a(t)} < \bar{\gamma} < 1 \qquad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.6}$$

For the converse it is easy to see that both (5.2.3) and (5.2.4) follow directly from (5.2.6). $\square$

In particular, if ν is sufficiently small, then

- $a(\cdot) > a_- > 0$ uniformly on $[0, T_0/\nu]$ for arbitrary $a_- < \tilde{a}_-$ through (5.1.9) and (5.2.2),

- $\bar{\gamma} < 1$ can be chosen independently of ν.

5.2.1. Consecutive Boxes

The following assumption actually serves the existence of some tube around the ν-adiabatic solution in which we will establish an attraction of paths:

$$\text{There are } \hat{R} > 0, \hat{\gamma} \in (0,1) \text{ with } \gamma\hat{R} + \frac{N\hat{R}^2}{a_-} + \frac{\nu}{a_-}\left(\frac{rM}{\kappa} + 2\delta\right) d_g = \hat{\gamma}\hat{R}. \tag{5.2.7}$$

The appearing $\hat{R}$ can be understood as the radius of that tube. In an intermediate step we will show that solutions, that do not deviate more that such an $\hat{R}$ (satisfying (5.2.7)) from the adiabatic solution, are attracted by the adiabatic solution path. We define the set

$$\mathcal{S} := \{\hat{R} \in (0,\infty) : \text{ There is } \hat{\gamma} \in (0,1) \text{ such that the equality in (5.2.7) holds}\}. \tag{5.2.8}$$

Condition (5.2.7) is a requirement on how small ν, $\hat{R}$, N and δ have to be and it reflects what we can expect of a region that ensures a contractional behavior, namely:

- Condition (5.2.7) is violated, if $\hat{R}$ is too big. This is due to the quadratic influence and the fact that this non-linearity can amplify the effect of large terms (of deviation).

- Condition (5.2.7) is also violated, if $\hat{R}$ is too small, while it is worth mentioning that, to this end, $\hat{R}$ has to be *small* in the sense of ν. The effect is mainly due to the fact, that the system changes with time and that the adiabatic solutions can track the equilibrium branch only at a distance of order ν. Non-linearity only plays a negligible role at this point.

Deterministic Case. Let $x^{\text{det}} = (x^{\text{det}}(t))_{t\in[0,T_0/\nu\wedge\tau_{\mathbb{D}})}$ denote the unique solution of the deterministic counterpart of (5.1.3), where $\tau_{\mathbb{D}} := \inf\{t \geq 0 : (x^{\text{det}}(t), \nu t) \notin \mathbb{D}\}$ as before. In the course of this subsection we show that under reasonable assumptions an initial segment x_0^{det}, which is not situated close to the equilibrium branch, induces a solution path that enters an environment around the equilibrium branch with diameter of order ν before a time of order $|\log\nu|/\sqrt{\nu}$, and does not leave before T_0/ν. We let $y^{\text{det}} = (y_t^{\text{det}})_{t\in[0,T_0/\nu\wedge\tau_{\mathbb{D}}]}$ denote the deviation of x^{det} from the adiabatic solution, i.e.

$$y^{\text{det}}(t) = x^{\text{det}}(t) - x^{\text{ad}_\nu}(t) \quad \text{for all } t \in [0, T_0/\nu \wedge \tau_{\mathbb{D}}]. \tag{5.2.9}$$

We further quantify the initial condition (5.1.11) by assuming that for some $R_0 > 0$, we have that

$$|y^{\text{det}}(t)| \leq R_0 \quad \text{for all } t \in [-r, 0]. \tag{5.2.10}$$

Then, denoting

$$\tau_R(y^{\text{det}}) := \inf\{t \geq 0 : |y^{\text{det}}(t)| > R\} \quad \text{for all } R > 0,$$

and, as long as $t \le \tau_{R_0}(y^{\det}) \wedge \tau_{\mathbb{D}} \wedge (T_0/\nu)$, we have that

$$\begin{aligned}
dy^{\det}(t) &= \Big[f(x^{\det}(t),\nu t) - f(x^{\mathrm{ad}_\nu}(t),\nu t)\Big]dt + \Big[g(x^{\det}(t-r),\nu t) - g(x^{\mathrm{ad}_\nu}(t),\nu t)\Big]dt \\
&= \Big[f_x(x^{\mathrm{ad}_\nu}(t),\nu t)y^{\det}(t) + \mathcal{R}_f\big(x^{\det}(t) - x^{\mathrm{ad}_\nu}(t),\nu t\big)\Big]dt \\
&\quad + \Big[g(x^{\det}(t-r),\nu t) - g(x^{\mathrm{ad}_\nu}(t-r),\nu t)\Big]dt \qquad (5.2.11)\\
&\quad + \Big[g(x^{\mathrm{ad}_\nu}(t-r),\nu t) - g(x^{\mathrm{ad}_\nu}(t),\nu t)\Big]dt \\
&= \Big[f_x(x^{\mathrm{ad}_\nu}(t),\nu t)y^{\det}(t) + \mathcal{R}_f\big(x^{\det}(t) - x^{\mathrm{ad}_\nu}(t),\nu t\big)\Big]dt \\
&\quad + \Big[g_x(x^{\mathrm{ad}_\nu}(t-r),\nu t)y^{\det}(t-r) + \mathcal{R}_g\big(x^{\det}(t-r) - x^{\mathrm{ad}_\nu}(t-r),\nu t\big)\Big]dt \\
&\quad + \Big[g(x^{\mathrm{ad}_\nu}(t-r),\nu t) - g(x^{\mathrm{ad}_\nu}(t),\nu t)\Big]dt. \qquad (5.2.12)
\end{aligned}$$

Let us focus on the process $\Xi = (\Xi(t) : t \in [0, T_0/\nu])$, which is defined by

$$\Xi(t) := g(x^{\mathrm{ad}_\nu}(t-r),\nu t) - g(x^{\mathrm{ad}_\nu}(t),\nu t) \quad \text{for all } t \in [0,T_0/\nu]. \qquad (5.2.13)$$

The quantity Ξ represents the mistake that is caused by using the adiabatic solution with respect to the replacement system ($r = 0$) as the reference for a delay-influenced solution. As stated in (5.2.2), the ν-adiabatic solution $(x^{\mathrm{ad}_\nu}(t))_{t\in[0,T_0/\nu]}$ remains close to the equilibrium branch $(x^\star(t))_{t\in[0,\infty)}$. We use that to deduce an estimate for Ξ. First, we note that

$$|x^{\mathrm{ad}_\nu}(t) - x^{\mathrm{ad}_\nu}(t-r)| \le |x^\star(t) - x^\star(t-r)| + 2\delta\nu \quad \text{for all } t \in [0,T_0/\nu].$$

Then, the implicit-function theorem provides differentiability for the replacement equilibrium branch and by (5.1.7) we know that

$$\begin{aligned}
x^\star(t) - x^\star(t-r) = \int_{t-r}^t dx^\star(u) &= \int_{t-r}^t \big(A^\star(u)\big)^{-1}\cdot\Big(\nu(f_t+g_t)(x^\star(u),\nu u)\Big)du \\
\Rightarrow \quad |x^\star(t) - x^\star(t-r)| &\le \int_{t-r}^t \frac{1}{\kappa}\cdot\nu M du \le \nu\frac{rM}{\kappa} \quad \text{for all } t \in [0,T_0/\nu].
\end{aligned}$$

Furthermore, as the spatial derivative of g is bounded by d_g, we attain the following upper-bound estimate on Ξ

$$|\Xi(t)| \le \nu c_0 \quad \text{for all } t \in [0,T_0/\nu], \quad \text{where } c_0 := \max\left\{\frac{rMd_g}{\kappa} + 2\delta d_g, 1\right\}. \qquad (5.2.14)$$

The lower boundary 1 was taken for technical reasons; it will avoid that we might divide by zero later. Continuing from (5.2.12), and denoting $\alpha(t,s) = \int_s^t a(u)du$ and $\alpha(t) = \alpha(t,0)$, we receive by the (classical) variation-of-constants formula

$$\begin{aligned}
y^{\det}(t) = y^{\det}(0)e^{-\alpha(t)} &+ \int_0^t e^{-\alpha(t,u)}b(u)y^{\det}(u-r)du \\
+ \int_0^t e^{-\alpha(t,u)}&\Big(\mathcal{R}_f\big(x^{\det}(u) - x^{\mathrm{ad}_\nu}(u),\nu u\big) \\
&+ \mathcal{R}_g\big(x^{\det}(u-r) - x^{\mathrm{ad}_\nu}(u-r),\nu u\big) + \Xi(u)\Big)du \quad \text{for all } t \in [0,T_0/\nu].
\end{aligned}$$

By monotony of the integral and the triangle inequality we therefore have that

$$\begin{aligned}\left|y^{\det}(t)\right| &\leq |y^{\det}(0)|e^{-\alpha(t)} + \int_0^t e^{-\alpha(t,u)}|b(u)||y^{\det}(u-r)|du \\ &\quad + \int_0^t e^{-\alpha(t,u)}\left(N_f\left(y^{\det}(u)\right)^2 + N_g\left(y^{\det}(u-r)\right)^2 + \nu c_0\right)du \quad \text{for all } t\in[0,T_0/\nu].\end{aligned} \tag{5.2.15}$$

Using (5.2.6), we obtain that

$$\begin{aligned}&\int_0^t e^{-\alpha(t,u)}|b(u)||y^{\det}(u-r)|du \\ &\quad = \int_0^t e^{-\alpha(t,u)}a(u)\frac{|b(u)|}{a(u)}|y^{\det}(u-r)|du \\ &\quad < \gamma R_0\int_0^t a(u)e^{-\alpha(t,u)}du = \gamma R_0(1-e^{-\alpha(t)}) \quad \text{for all } t\in\left[0,\tau_{R_0}(y^{\det})\wedge(T_0/\nu)\right),\end{aligned}$$

and also that

$$\begin{aligned}&\int_0^t e^{-\alpha(t,u)}\left(N_f\left(y^{\det}(u)\right)^2 + N_g\left(y^{\det}(u-r)\right)^2 + \nu c_0\right)du \\ &\qquad \leq \frac{1}{a_-}\left(NR_0^2+\nu c_0\right)\left(1-e^{-\alpha(t)}\right) \quad \text{for all } t\in\left[0,\tau_{R_0}(y^{\det})\wedge(T_0/\nu)\right].\end{aligned}$$

So we may deduce that

$$\begin{aligned}\left|y^{\det}(t)\right| &< R_0e^{-\alpha(t)} + \left(\gamma R_0 + \frac{NR_0^2+\nu c_0}{a_-}\right)\left(1-e^{-\alpha(t)}\right) \\ &= \left(\gamma R_0 + \frac{NR_0^2+\nu c_0}{a_-}\right) + \left(R_0 - \gamma R_0 - \frac{NR_0^2+\nu c_0}{a_-}\right)e^{-\alpha(t)} \\ &\qquad\qquad \text{for all } t\in\left[0,\tau_{R_0}(y^{\det})\wedge(T_0/\nu)\right].\end{aligned} \tag{5.2.16}$$

Assuming that $R_0\in\mathcal{S}$, we observe that

$$\gamma R_0 + \frac{NR_0^2+\nu c_0}{a_-} < R_0 \quad \Leftrightarrow \quad R_0-\gamma R_0-\frac{NR_0^2+\nu c_0}{a_-} > 0.$$

Therefore, we know that in (5.2.16), we truly observe a monotone exponential decay in t on the right-hand side. We store the straightforward implications inside the following two corollaries:

Corollary 5.5. *If we assume that $|y^{\det}(t)| < R_0\in\mathcal{S}$ for all $t\in[-r,0]$, we have that*

$$|y^{\det}(t)| < R_0 \quad \textit{for all } t\in\left[0,\tau_{R_0}(y^{\det})\wedge(T_0/\nu)\right]$$

actually providing that $\tau_{R_0}(y^{\det}) > T_0/\nu$. Moreover, for $t>\log(2)/a_-$ we have $e^{-\alpha(t)}<1/2$ and so,

$$\begin{aligned}|y^{\det}(t)| &\leq \left(\gamma R_0 + \frac{NR_0^2+\nu c_0}{a_-}\right) + \frac{1}{2}\left(R_0-\gamma R_0-\frac{NR_0^2+\nu c_0}{a_-}\right) \\ &= \frac{R_0}{2}\left(1+\gamma+\frac{NR_0}{a_-}+\frac{\nu c_0}{a_-}\frac{1}{R_0}\right) \quad \textit{for all } t\in\left[\frac{\log 2}{a_-},\tau_{R_0}(y^{\det})\wedge(T_0/\nu)\right].\end{aligned}$$

Corollary 5.5 manifests the concept of a decay factor by which the initial distance between x^{det} and x^{ad_ν} is multiplied after evolving for at least $\log(2)/a_-$ units of time. This relation between the current distance y^{det} and the decay factor is fixed in the following definition. The below remark gathers its essential properties. The underlying mathematics is so basic that is does not seem necessary to formulate a proposition-proof scheme. Furthermore, a bit of a reminder is included to have all the related facts collected at a glance for the forthcoming computations.

Definition 5.6. *Define the decay-factor function* $q : (0, \infty) \to [0, \infty)$ *by*

$$q(x) := \frac{1}{2}\left(1 + \gamma + \frac{1}{a_-}\left(Nx + \frac{\nu c_0}{x}\right)\right) = \frac{1+\gamma}{2} + \frac{N}{2a_-}x + \frac{\nu c_0}{2a_-}\frac{1}{x} \quad \textit{for all } x \in (0, \infty). \tag{5.2.17}$$

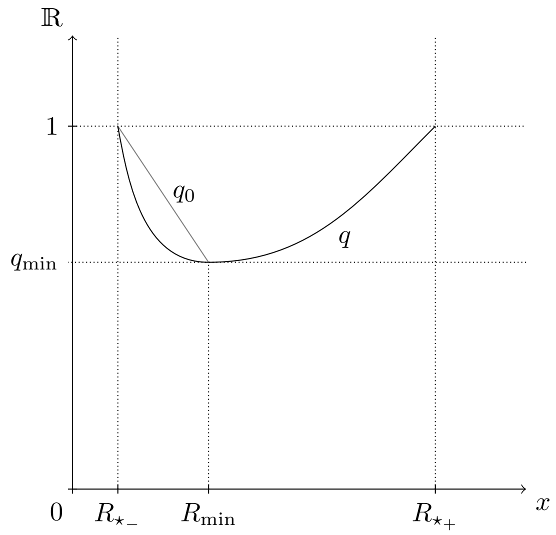

Figure 9: Illustration of the decay-factor function q and related quantities.

Remark 5.7 (Properties of the decay-factor function)**.**

(i) The set $\mathcal{S}$*, defined in* (5.2.8)*, allows the characterization*

$$\mathcal{S} = \{x \in [0, \infty) : q(x) \in (0, 1)\}\,.$$

(ii) Intersections with 1*. Note that* $c_0 = \frac{rMd_g}{\kappa} + 2\delta d_g$*, defined in* (5.2.14)*. Then,* $q(x) = 1 \Leftrightarrow x \in \partial\mathcal{S} = \{R_{\star_-}, R_{\star_+}\}$ *with*

$$\begin{aligned}
R_{\star_-} &= \frac{a_-(1-\gamma)}{2N} - \sqrt{\left(\frac{a_-(1-\gamma)}{2N}\right)^2 - \frac{\nu c_0}{N}} = \frac{a_-(1-\gamma)}{2N} - \frac{a_-(1-\gamma)}{2N}\sqrt{1 - \frac{4\nu c_0 N}{a_-^2(1-\gamma)^2}} \\
&= \frac{\nu c_0}{a_-(1-\gamma)} + \mathcal{O}\left(\nu^2 N\right), \\
R_{\star_+} &= \frac{a_-(1-\gamma)}{2N} + \sqrt{\left(\frac{a_-(1-\gamma)}{2N}\right)^2 - \frac{\nu c_0}{N}} = \frac{a_-(1-\gamma)}{2N} + \frac{a_-(1-\gamma)}{2N}\sqrt{1 - \frac{4\nu c_0 N}{a_-^2(1-\gamma)^2}} \\
&= \frac{a_-(1-\gamma)}{N} + \mathcal{O}(\nu).
\end{aligned}$$

(iii) Derivatives

$$\frac{dq(x)}{dx} = \frac{N}{2a_-} - \frac{\nu c_0}{2a_-}\frac{1}{x^2} \quad \textit{and} \quad \frac{d^2q(x)}{dx^2} = \frac{\nu c_0}{a_-}\frac{1}{x^3} \quad \textit{for all } x \in (0,\infty).$$

In particular, q is strictly convex and strictly decreasing on $(0, R_{\min})$, where $R_{\min} := \sqrt{\frac{\nu c_0}{N}}$. The mapping q has a unique global (restricted to $(0,\infty)$) minumum at $x = R_{\min}$, which comes along with the minimal value

$$q_{\min} := q(R_{\min}) = \frac{1+\gamma}{2} + \frac{N}{2a_-}\frac{\sqrt{\nu c_0}}{\sqrt{N}} + \frac{\nu c_0}{2a_-}\frac{\sqrt{N}}{\sqrt{\nu c_0}} = \frac{1+\gamma}{2} + \frac{1}{a_-}\sqrt{\nu N c_0}.$$

(iv) All points on the secant segment, that connects $(R_{\star-}, 1)$ and $(R_{\min}, q_{\min})$, lie above the graph of q. The points' values of the secant segment are given by the mapping $q_0 : [R_{\star-}, R_{\min}] \to \mathbb{R}$, defined as

$$q_0\big(R_{\star-} + x\big) = q(R_{\star-}) - \frac{q(R_{\star-}) - q(R_{\min})}{R_{\min} - R_{\star-}}x \quad \textit{for all } x \in [0, R_{\min} - R_{\star-}].$$

Convexity yields that $q_0(x) \geq q(x)$ for all $x \in [R_{\star-}, R_{\min}]$. With $q(R_{\star-}) = q_0(R_{\star-})$ that leads to the following inequality:

$$q(R_{\star-}) - q(R_{\star-} + x) \geq q_0(R_{\star-}) - q_0(R_{\star-} + x) = \frac{q(R_{\star-}) - q(R_{\min})}{R_{\min} - R_{\star-}}x \qquad (5.2.18)$$
$$\textit{for all } x \in (0, R_{\min} - R_{\star-}).$$

Corollary 5.8. *Let $\varepsilon_\star > 0$ such that $R_{\star-}(1+\varepsilon_\star) < R_{\min}$, $R_0 \in \big[R_{\star-}(1+\varepsilon_\star), R_{\min}\big]$, and define $R_i := q(R_{i-1})R_{i-1}$ for $i = \{1, \ldots, n\}$. If $n \in \mathbb{N}$ is such that $R_i \geq R_{\star-}(1+\varepsilon_\star)$ for all $i \in \{1, \ldots, n\}$, then $q(R_i) \leq q\big(R_{\star-}(1+\varepsilon_\star)\big)$ for all $i \in \{1, \ldots, n\}$. Furthermore,*

$$R_i \leq R_0 q^i\big(R_{\star-}(1+\varepsilon_\star)\big) \quad \textit{for all } i \in \{0, \ldots, n\}.$$

Lemma 5.9. *Suppose $R_0 \in \big(R_{\star-}(1+\varepsilon_\star), R_{\min}\big)$ for some $\varepsilon_\star > 0$, then*

$$q\big(R_{\star-}(1+\varepsilon_\star)\big) \leq 1 - \sqrt{\nu N}\frac{\sqrt{c_0}}{2a_-}\varepsilon_\star + \mathcal{O}(\nu N \varepsilon_\star).$$

Proof. Setting $x = \varepsilon_\star R_{\star-}$ in (5.2.18), we deduce that

$$\begin{aligned} q\big(R_{\star-}(1+\varepsilon_\star)\big) \leq q_0\big(R_{\star-}(1+\varepsilon_\star)\big) &\leq q(R_{\star-}) - \frac{q(R_{\star-}) - q(R_{\min})}{R_{\min} - R_{\star-}}R_{\star-}\varepsilon_\star \\ &= 1 - \frac{1 - q(R_{\min})}{R_{\min} - R_{\star-}}R_{\star-}\varepsilon_\star. \end{aligned}$$

We note that

$$1 - q(R_{\min}) = 1 - \frac{1+\gamma}{2} - \frac{1}{a_-}\sqrt{\nu N c_0} = \frac{1-\gamma}{2} + \mathcal{O}(\sqrt{\nu N}),$$

and that

$$\begin{aligned}
R_{\min} - R_{\star_-} &= \sqrt{\frac{\nu c_0}{N}} - \frac{\nu c_0}{a_-(1-\gamma)} + \mathcal{O}(\nu^2 N) \\
&= \sqrt{\frac{\nu}{N}} \left(\sqrt{c_0} - \frac{\sqrt{\nu N}\, c_0}{a_-(1-\gamma)} + \mathcal{O}\big(\nu^{3/2} N^{3/2}\big) \right) = \sqrt{\frac{\nu}{N}} \left(\sqrt{c_0} + \mathcal{O}(\sqrt{\nu N}) \right).
\end{aligned}$$

So,

$$\begin{aligned}
\frac{1 - q(R_{\min})}{R_{\min} - R_{\star_-}} &= \frac{\frac{1-\gamma}{2} + \mathcal{O}\big(\sqrt{\nu N}\big)}{\sqrt{\frac{\nu}{N}} \left(\sqrt{c_0} + \mathcal{O}\big(\sqrt{\nu N}\big) \right)} = \frac{\sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2} + \mathcal{O}(N)}{\sqrt{c_0} + \mathcal{O}\big(\sqrt{\nu N}\big)} = \frac{\sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2}}{\sqrt{c_0} + \mathcal{O}\big(\sqrt{\nu N}\big)} + \mathcal{O}(N) \\
&= \sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2} \left(\frac{1}{\sqrt{c_0}} + \mathcal{O}\big(\sqrt{\nu N}\big) \right) + \mathcal{O}(N) = \sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2\sqrt{c_0}} + \mathcal{O}(N).
\end{aligned}$$

Altogether, we have that

$$\begin{aligned}
q\big(R_{\star_-}(1+\varepsilon_\star)\big) &\leq 1 - \left(\sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2\sqrt{c_0}} + \mathcal{O}(N) \right) R_{\star_-} \varepsilon_\star \\
&= 1 - \left(\sqrt{\frac{N}{\nu}} \frac{1-\gamma}{2\sqrt{c_0}} + \mathcal{O}(N) \right) \left(\frac{\nu c_0}{a_-(1-\gamma)} + \mathcal{O}\left(\nu^2 N\right) \right) \varepsilon_\star \\
&= 1 - \left(\sqrt{\nu N} \frac{1-\gamma}{2} \frac{\sqrt{c_0}}{a_-(1-\gamma)} + \mathcal{O}\Big(\nu N + \nu^{3/2} N^{3/2}\Big) \right) \varepsilon_\star \\
&= 1 - \sqrt{\nu N} \frac{\sqrt{c_0}}{2a_-} \varepsilon_\star + \mathcal{O}(\nu N \varepsilon_\star).
\end{aligned}$$

□

Theorem 5.10. *Let $\varepsilon_\star > 0$ and $|y^{\det}(t)| \leq R_0$ for all $t \in [-r, 0]$. Denote $\theta_i = i(r + \log(2)/a_-)$ for $i \in \mathbb{N}$.*

a) For $R_0 \in \big(R_{\star_-}(1+\varepsilon_\star), R_{\min}\big)$ we have that

$$|y^{\det}(t)| \leq R_{\star_-}(1+\varepsilon_\star) \qquad \text{for all } t \in [\theta_{n_\star} - r, T_0/\nu),$$

where $n_\star$ is given by

$$n^\star := \frac{2a_-}{\sqrt{\nu N}\, \varepsilon_\star} \left| \log \frac{R_{\star_-}(1+\varepsilon_\star)}{R_0} \right| \left(1 + \mathcal{O}\big(\sqrt{\nu N}\big) \right).$$

b) For $R_0 \in (R_{\min}, R_{\star_+}(1-\varepsilon_\star))$ we find that

$$|y^{\det}(t)| \leq R_{\min} \qquad \text{for all } t \in [\theta_{m_\star} - r, T_0/\nu),$$

where

$$\begin{aligned}
m_\star := \frac{\log \frac{R_{\min}}{R_0}}{\log q(R_{\star_+}(1-\varepsilon_\star))} &\leq \left| \log \frac{R_{\min}}{R_0} \right| \left(\frac{2}{\varepsilon_\star(1-\gamma)} + \mathcal{O}\left(\frac{\sqrt{\nu N}}{\varepsilon_\star} \right) \right) \\
&= \mathcal{O}\left(\frac{\log(R_0) + |\log \nu|}{\varepsilon_\star} \right).
\end{aligned}$$

Proof. Defining $\hat{R}_0 = R_0$ and $\hat{R}_i = \hat{R}_{i-1} q(R_{\star_-}(1+\varepsilon_\star))$ for all $i \in \mathbb{N}$, by Corollary 5.8 and Corollary 5.5 we know that $R_n \leq \hat{R}_n \vee R_{\star_-}(1+\varepsilon_\star)$ for all $n \in \mathbb{N}$, where $(R_i)_{i\in\mathbb{N}}$ is defined as in Corollary 5.8. In particular, we have the following:

$$R_n \le R_{\star_-}(1+\varepsilon_\star) \quad \text{or} \quad R_n \le R_0 q^n\big(R_{\star_-}(1+\varepsilon_\star)\big) \quad \text{for all } n \in \mathbb{N}.$$

Then, we observe that

$$\begin{aligned} & R_0 q^n\big(R_{\star_-}(1+\varepsilon_\star)\big) \le R_{\star_-}(1+\varepsilon_\star) \\ \Leftrightarrow \quad & n \ge \left|\log \frac{R_{\star_-}(1+\varepsilon_\star)}{R_0}\right| \frac{1}{|\log q(R_{\star_-}(1+\varepsilon_\star))|} =: n_0. \end{aligned}$$

As $1-x \le e^{-x}$ and therefore $\log(1-x) \le -x$ for all $x \in (0,1)$, we have that $\frac{1}{|\log(1-x)|} \le \frac{1}{x}$. That reveals with Lemma 5.9 that

$$\begin{aligned} n_0 &\le \left|\log \frac{R_{\star_-}(1+\varepsilon_\star)}{R_0}\right| \frac{1}{\sqrt{\nu N}\frac{c_0}{2a_-}\varepsilon_\star + \mathcal{O}(\nu N \varepsilon_\star)} \\ &= \left|\log \frac{R_{\star_-}(1+\varepsilon_\star)}{R_0}\right| \frac{2a_-}{c_0\sqrt{\nu N}\,\varepsilon_\star}\Big(1+\mathcal{O}\big(\sqrt{\nu N}\big)\Big), \end{aligned}$$

which through Lemma 5.5 shows the first part. The analogue assertion for $|x^{\det}(t) - x^{\mathrm{ad}_\nu}(t)| \in (R_{\min}, R_{\star_+}(1-\varepsilon_\star))$ is straightforward. The same ideas as before yield a similar decay factor in that case; we find that

$$q(R_{\star_+}(1-\varepsilon_\star)) \le 1 - \frac{q(R_{\star_+}) - q(R_{\min})}{R_{\star_+} - R_{\min}} R_{\star_+}\varepsilon_\star = 1 - \frac{1-\gamma}{2}\varepsilon_\star + \mathcal{O}\big(\varepsilon_\star\sqrt{N\nu}\big). \tag{5.2.19}$$

And the rest of part b) relies on mere computations. □

Remark 5.11.

- *The above condition on $n^\star$ is surely not optimal; in particular, we have used a uniform (lower-bound) rate on which $y^{\det}$ approaches x^{ad_ν}, although the rate is (much) better when the distance between $y^{\det}$ and $x^\star$ is not yet $\mathcal{O}(\nu)$.*
- *The applied upper bound for the decay factor in (5.2.19) used in part b) is unaffected – despite correction terms – from the small parameter ν. This is the reason why the $m_\star$ is small compared to $n_\star$ when ν gets small.*

Summarizing we have shown that if a deviation process $y^{\det}(t) = x^{\det}(t) - x^{\mathrm{ad}_\nu}(t)$, $t \in [0, T_0/\nu]$ for some $t_0 \in [0, T_0/\nu]$ satisfies $\|y^{\det}_{t_0}\| \le R_0$, then the following assertions hold true:

- For any $\varepsilon_\star > 0$, if $R_0 \le R_{\star_-}(1+\varepsilon_\star)$, then $x^{\det}$ remains within a distance of order ν around the equilibrium branch at least up to time T_0/ν.
- For any $\varepsilon_\star > 0$, if $R_0 > R_{\star_-}(1+\varepsilon_\star)$, then $x^{\det}$ enters a neighborhood of order ν around the equilibrium branch within a time of order $|\log\nu|/\sqrt{\nu}$ and does not leave before T_0/ν.

In other words: For every $\varepsilon_\star > 0$ with $R_{\star_-}(1+\varepsilon_\star) < R_{\star_+}$ the set of paths

$$\mathcal{M}_{\varepsilon_\star} = \left\{(x(t))_{t\in[-r,T_0/\nu]} \text{ satisfying (5.1.3)}, \ \|x(t) - x^{\mathrm{ad}_\nu}(t)\|_{[-r,T_0/\nu]} < R_{\star_-}(1+\varepsilon_\star)\right\},$$

constitutes an *invariant manifold* enveloping the ν-adiabatic solution with diameter $R_{\star_-}(1+\varepsilon_\star) > 0$. Further, we have shown that the invariant manifold $\mathcal{M}_{\varepsilon_\star}$ is attracting with *basin of attraction*

$$\mathcal{A} := \left\{\Upsilon \in \mathcal{C}([-r,0],\mathbb{R}) : \|\Upsilon\| < |R_{\star_+}| \text{ and the according solution remains in } \mathbb{D}\right\}.$$

Additive Noise. In this section we extend the consecutive-boxes approach to white noise. The ideas remain mostly the same and so do the calculations. To begin with, we review a concentration inequality of an Ornstein–Uhlenbeck process with additive noise in stable regime, which will turn out to be helpful for our attempt later on. The result can be found beautifully presented in [BG06, Chapter 3.1.1], where it is formulated in greater generality than the adapted form that we present below. We consider the Ornstein–Uhlenbeck process with white noise $(z(t))_{t\in[0,\infty)}$ which is the unique solution of

$$\begin{cases} dz(t) = -a(t)z(t)dt + dW(t) & \text{for } t \geq 0, \\ z(0) = 0, \end{cases} \tag{5.2.20}$$

where $a(t) = \tilde{a}(\nu t) > a_-$ for all $t \in [0,\infty)$ for some continuously differentiable $\tilde{a} : [0,\infty) \to [a_-,\infty)$. The initial value 0 reflects that it is truely all about deviation. The related variance process $v(t) = \operatorname{var} z(t)$, $t \in [0,\infty)$, then satisfies the differential law

$$dv(t) = -2a(t)v(t)dt + 1 \quad \text{for all } t \in [0,\infty), \tag{5.2.21}$$

and its equilibrium branch is given by $t \mapsto 1/2a(t), t \in [0,\infty)$. Then, we obtain the existence of a ν-adiabatic solution $(\zeta(t))_{t\in[0,\infty)}$ that satisfies

$$\zeta(t) := \frac{1}{2a(t)} + \mathcal{O}(\nu) \quad \text{for } t \in [0,T_0] \tag{5.2.22}$$

for any finite $T_0 > 0$. Essentially by [BG06, Theorem 3.1.5], denoting $\alpha(t) = \int_0^t a(u)du$ for all $t \in [0,T_0]$ we have that

$$\mathbb{P}\left\{ \sup_{s\in[0,T_0/\nu]} \frac{|z(t)|}{\sqrt{\zeta(t)}} > \beta \right\} \leq \frac{2eT_0\beta^2\big(1+\mathcal{O}(\nu)\big)}{\nu\alpha(T_0/\nu)} \exp\left(-\frac{\beta^2}{2}\right) \quad \text{for } \beta > 0. \tag{5.2.23}$$

For the above probability to become small, it suffices to choose β of order $|\log\nu|$. A full proof, that contains all the above claims, can be found in the appendix A.1.

In this part, the central role is again taken by the deviation $y = (y(t))_{t\in[0,\infty)}$ of the solution $(x(t))_{t\in[-r,\infty)}$ of RFDE (5.1.3) from the ν-adiabatic solution x^{ad_ν} of the replacement system. Generally speaking, performing the same computational steps as before, we observe that the differential law of $(y(t))_{t\in[-r,T_0/\nu]}$ allows consecutive estimates in every $t_0 \in [0,T_0/\nu]$, which resembles the technique, that has lead to the results of the deterministic case. Let c_0 be given as in (5.2.14). For arbitrary $t_0 \in [0,T_0/\nu]$ the same basic ideas, that have lead to Corollary 5.5, in particular estimate (5.2.15), yield

$$\begin{aligned} |y(t_0+s)| \leq{}& |y(t_0)|e^{-\alpha(t_0+s,t_0)} \\ &+ \int_0^s e^{-\alpha(t_0+s,t_0+u)}|b(t_0+u)||y(t_0-r+u)|du \\ &+ \int_0^s e^{-\alpha(t_0+s,t_0+u)}\left(N_f y^2(t_0+u) + N_g y^2(t_0+u-r) + \nu c_0\right)du \\ &+ \sigma\left|\int_{t_0}^{t_0+s} e^{-\alpha(t_0+s,u)}dW(u)\right| \quad \text{for all } s \in [0,T_0/\nu - t_0]. \end{aligned} \tag{5.2.24}$$

We will continue to assume $|y(t)| \le R_0$ for all $t \in [-r, 0]$ and we will work out sufficient conditions on R_0 that have lead to a contractional behavior just like in the deterministic part. As an analogue of assumption (5.2.7) from the deterministic case, if $|y(t)| \le R_0$ for all $t \in [-r, 0]$, we assume that σ and β are small enough such that

$$\frac{1}{2}\left(R_0 + \gamma\left(R_0 + \sigma\beta\right) + \frac{N(R_0 + \sigma\beta)^2}{a_-} + \frac{\nu c_0}{a_-} + 2\sigma\beta\right) < R_0. \tag{5.2.25}$$

And if condition (5.2.7) is satisfied, then (5.2.25) is only an assumption on how small σ must be.

Definition 5.12. *Let us for arbitrary $t \in [0, \infty)$ denote $\tau_R^{(t)}(y) = \inf\{u \ge t : |y(u)| > R\}$ for arbitrary $R > 0$. Further denote*

$$\begin{aligned}
\theta_i &:= i\left(\frac{\log(2)}{a_-} + r\right) \qquad \text{for all } i \in \{0, 1, 2, \ldots\}, \\
\xi^{(i)}(t) &:= \int_{\theta_i}^{t} e^{-\alpha(t,u)} dW(u) \quad \text{for all } t \in [\theta_i, \theta_{i+1}], i \in \{0, 1, 2, \ldots\}, \\
\xi(t) &:= \int_{0}^{t} e^{-\alpha(t,u)} dW(u) \quad \text{for all } t \in [0, T_0/\nu], \\
\tau_\beta^{[i]}(\xi) &:= \inf\left\{t \ge 0 : t \in [\theta_i, \theta_{i+1}] : \xi_t^{(i)} > \beta\right\} \quad \text{for all } i \in \{0, 1, 2, \ldots\}, \\
\tau_\beta(\xi) &:= \min\{\tau_\beta^{[i]}(\xi) : i \in \{0, 1, 2, \ldots\}\} \quad \text{for } \beta > 0.
\end{aligned}$$

Compared to the determinstic estimate for the deviation process $y^{\det}$ in Corollary 5.5, the additive noise gives rise to additional terms. Consequently, the results of the stochastically perturbed version are based on a slightly modified version of the formulation in Definition 5.6 and Remark 5.7.

Definition 5.13. *We define a decay-factor function $\tilde{q} : (0, \infty) \to \mathbb{R}$ by*

$$\tilde{q}(x) := \frac{1}{2x}\left(x + \gamma\left(x + \sigma\beta\right) + \frac{N(x + \sigma\beta)^2}{a_-} + \frac{\nu c_0}{a_-} + 2\sigma\beta\right).$$

The above defined decay-factor function $\tilde{q}$ allows the following representations which will be helpful for computations; for all $x \in (0, \infty)$ we have that

$$\begin{aligned}
\tilde{q}(x) &= \frac{1}{2}\left(1 + \gamma\left(1 + \frac{\sigma\beta}{x}\right) + \frac{Nx}{a_-} + \frac{2N\sigma\beta}{a_-} + \frac{N\sigma^2\beta^2}{a_- x} + \frac{\nu c_0}{x a_-} + \frac{2\sigma\beta}{x}\right) \\
&= \frac{N}{2a_-}x + \frac{1}{2}\left(1 + \gamma + \frac{2N\sigma\beta}{a_-}\right) + \frac{1}{2}\left(\gamma\sigma\beta + \frac{N\sigma^2\beta^2}{a_-} + \frac{\nu c_0}{a_-} + 2\sigma\beta\right)\frac{1}{x} \\
&= \frac{N}{2a_-}\left(x + \frac{a_-(1+\gamma)}{N} + 2\sigma\beta + \left(\frac{a_-\sigma\beta}{N}(\gamma + 2) + \sigma^2\beta^2 + \frac{\nu c_0}{N}\right)\frac{1}{x}\right).
\end{aligned}$$

Just like in the deterministic case, the decay-factor function $\tilde{q}$ is analytically simple, but a little bulky when it comes to computations. That is why we will state the interesting properties as a lemma this time. Section 5.3 deals with the situation when the stability, manifested as $1 - \gamma$ gets small. To this end, we will no longer drop terms like $\frac{1}{1-\gamma}$ in the Landau symbols, because it will spare us lots of extra computational effort then.

Lemma 5.14. *The decay-factor function $\tilde{q}$ has the following properties.*

a) Intersections with 1. *Using the notation*

$$\tilde{c} := (\gamma+2)\sigma\beta + \frac{N}{a_-}\sigma^2\beta^2 + \frac{\nu c_0}{a_-} \tag{5.2.26}$$

we have that $\tilde{q}(x) = 1 \Leftrightarrow x \in \{\tilde{R}_{\star_-}, \tilde{R}_{\star_+}\}$, where

$$\begin{aligned}
\tilde{R}_{\star_-} &= \frac{\tilde{c}}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right),\\
\tilde{R}_{\star_+} &= \left(\frac{a_-}{N}(1-\gamma) - 2\sigma\beta\right)\left(1 + \mathcal{O}\left(\frac{N\tilde{c}}{(1-\gamma)^2}\right)\right).
\end{aligned}$$

b) Derivatives and $(0,\infty)$-global minimum. For all $x \in (0,\infty)$,

$$\frac{d}{dx}\tilde{q}(x) = \frac{N}{2a_-} - \frac{\tilde{c}}{2x^2} \quad \text{and} \quad \frac{d^2}{dx^2}\tilde{q}(x) = \frac{\tilde{c}}{x^3}.$$

In particular, $\tilde{q}$ is strictly convex on $(0,\infty)$ and there is a unique minimum at

$$\tilde{R}_{\min} = \sqrt{\frac{a_-\tilde{c}}{N}} \quad \text{with} \quad \tilde{q}(\tilde{R}_{\min}) = \frac{1+\gamma}{2} + \sqrt{\frac{N\tilde{c}}{a_-}} + \frac{N}{a_-}\sigma\beta.$$

Proof. a) Let us introduce the notations

$$\lambda_0 = \frac{a_-\sigma\beta}{N}(\gamma+2) + \sigma^2\beta^2 + \frac{\nu c_0}{N} \quad \text{and} \quad \lambda_1 = \frac{a_-}{N}(1-\gamma) - 2\sigma\beta.$$

Then, we find that

$$\begin{aligned}
&\tilde{q}(x) = 1\\
\Leftrightarrow\quad & \frac{N}{2a_-}\left(x + \frac{a_-(1+\gamma)}{N} - \frac{2a_-}{N} + 2\sigma\beta + \left(\frac{a_-\sigma\beta}{N}(\gamma+2) + \sigma^2\beta^2 + \frac{\nu c_0}{N}\right)\frac{1}{x}\right) = 0\\
\Leftrightarrow\quad & x - \frac{a_-}{N}(1-\gamma) + 2\sigma\beta + \left(\frac{a_-\sigma\beta}{N}(\gamma+2) + \sigma^2\beta^2 + \frac{\nu c_0}{N}\right)\frac{1}{x} = 0\\
\Leftrightarrow\quad & x - \lambda_1 + \frac{\lambda_0}{x} = 0.
\end{aligned}$$

It is obvious that 0 is no solution of the equation. Therefore, $\tilde{q}(x) = 1 \Leftrightarrow x \in \{\tilde{R}_{\star_-}, \tilde{R}_{\star_+}\}$, where, using $\sqrt{1+z} = 1 + \frac{z}{2} + \mathcal{O}(z^2)$ for z near zero, we have that

$$\begin{aligned}
\tilde{R}_{\star_+} &= \frac{\lambda_1}{2} + \sqrt{\left(\frac{\lambda_1}{2}\right)^2 - \lambda_0} = \frac{\lambda_1}{2} + \frac{\lambda_1}{2}\sqrt{1 - \frac{4\lambda_0}{\lambda_1^2}} = \frac{\lambda_1}{2} + \frac{\lambda_1}{2}\left(1 - \frac{4\lambda_0}{2\lambda_1^2} + \mathcal{O}\left(\frac{\lambda_0^2}{\lambda_1^4}\right)\right)\\
&= \lambda_1\left(1 + \mathcal{O}\left(\frac{\lambda_0}{\lambda_1^2}\right)\right),\\
\tilde{R}_{\star_-} &= \frac{\lambda_1}{2} - \frac{\lambda_1}{2}\left(1 - \frac{4\lambda_0}{2\lambda_1^2} + \mathcal{O}\left(\frac{\lambda_0^2}{\lambda_1^4}\right)\right) = \frac{\lambda_0}{\lambda_1}\left(1 + \mathcal{O}\left(\frac{\lambda_0}{\lambda_1^2}\right)\right).
\end{aligned}$$

Note that with $\frac{1}{1+z} = 1 - z + \mathcal{O}(z^2)$ for small z,

$$\begin{aligned}
\frac{1}{\lambda_1} &= \frac{1}{\frac{a_-}{N}(1-\gamma)\left(1 - \frac{2\sigma\beta N}{a_-(1-\gamma)}\right)} = \frac{N}{a_-(1-\gamma)}\left(1 - \frac{2\sigma\beta N}{a_-(1-\gamma)} + \mathcal{O}\left(\frac{\sigma^2\beta^2N^2}{(1-\gamma)^2}\right)\right) \\
&= \frac{N}{a_-(1-\gamma)}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right) = \mathcal{O}\left(\frac{N}{1-\gamma}\right), \\
\frac{1}{\lambda_1^2} &= \frac{N^2}{a_-^2(1-\gamma)^2}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right)^2 = \frac{N^2}{a_-^2(1-\gamma)^2}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right).
\end{aligned}$$

Then, for the leading term of $\tilde{R}_{\star_+}$, we find that

$$\begin{aligned}
\frac{\lambda_0}{\lambda_1} &= \left(\frac{\gamma+2}{1-\gamma}\sigma\beta + \frac{N\sigma^2\beta^2}{a_-(1-\gamma)} + \frac{\nu c_0}{a_-(1-\gamma)}\right)\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right) \\
&= \frac{\tilde{c}}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right), \\
\frac{\lambda_0}{\lambda_1^2} &= \frac{N\tilde{c}}{a_-(1-\gamma)^2}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right).
\end{aligned}$$

and then, it is easy to see that $\mathcal{O}\left(\lambda_0/\lambda_1^2\right) = \mathcal{O}\left(\frac{N\tilde{c}}{(1-\gamma)^2}\right)$. Therefore,

$$\begin{aligned}
\tilde{R}_{\star_-} &= \frac{\lambda_0}{\lambda_1}\left(1 + \mathcal{O}\left(\frac{\lambda_0}{\lambda_1^2}\right)\right) = \frac{\tilde{c}}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma}\right)\right)\left(1 + \mathcal{O}\left(\frac{N\tilde{c}}{(1-\gamma)^2}\right)\right) \\
&= \frac{\tilde{c}}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right), \\
\tilde{R}_{\star_+} &= \lambda_1\left(1 + \mathcal{O}\left(\frac{\lambda_0}{\lambda_1^2}\right)\right) = \left(\frac{a_-}{N}(1-\gamma) - 2\sigma\beta\right)\left(1 + \mathcal{O}\left(\frac{N\tilde{c}}{(1-\gamma)^2}\right)\right).
\end{aligned}$$

b) For all $x \in (0, \infty)$ we have

$$\begin{aligned}
\frac{d}{dx}\tilde{q}(x) &= \frac{N}{2a_-}\left(1 - \left(\frac{a_-\sigma\beta}{N}(\gamma+2) + \sigma^2\beta^2 + \frac{\nu c_0}{N}\right)\frac{1}{x^2}\right) = \frac{N}{2a_-} - \frac{\tilde{c}}{2x^2}, \\
\frac{d^2}{dx^2}\tilde{q}(x) &= \frac{\tilde{c}}{x^3},
\end{aligned}$$

serving strict convexity of $\tilde{q}$ on $(0, \infty)$, which implies that there is at most one minimum over $(0, \infty)$. The first-order criterion serves the existence of a minimum at $x = \tilde{R}_{\min}$, where

$$\tilde{R}_{\min} := \sqrt{\frac{a_-\tilde{c}}{N}}.$$

To compute the value $\tilde{q}(\tilde{R}_{\min})$, we rewrite $\tilde{q}$ as

$$\tilde{q}(x) = \frac{N}{2a_-}x + \frac{1}{2}\left(1 + \gamma + \frac{2N\sigma\beta}{a_-}\right) + \frac{\tilde{c}}{2x} \quad \text{for all } x \in (0, \infty).$$

Then, it is easy to see that

$$\tilde{q}(\tilde{R}_{\min}) = \frac{1+\gamma}{2} + \sqrt{\frac{N\tilde{c}}{a_-}} + \frac{N}{a_-}\sigma\beta.$$

□

Regarding (5.2.24), the following consecutive estimate on $(y(t))_{t\in[-r,T_0/\nu]}$ is straighforward.

Lemma 5.15. *With the notations from above, the following two assertions hold true.*

a) Assume that for some $i \in \mathbb{N}$

$$\tilde{q}(\tilde{R}_{k-1}) < 1, \quad \textit{for all} \quad k \in \{1, \dots, i\}, \ \textit{where}$$
$$\tilde{R}_{k+1} := R_0 \prod_{i=0}^{k} \tilde{q}(\tilde{R}_i) \quad \textit{for all } k \in \{0, 1, 2, \dots\}.$$

Let further $(y(t))_{t\in[-r,T_0/\nu]}$ be the unique solution of (5.1.3) and assume that $\|y_0\| \le R_0$. Then

$$\begin{cases} |y(t)| \le \tilde{R}_{i-1} + \sigma\beta & \textit{for all } t \in \left[\theta_i - r, \left(\theta_i + \frac{\log 2}{a_-}\right) \wedge \tau_\beta(\xi) \wedge (T_0/\nu)\right], \\ |y(t)| \le \tilde{R}_{i-1}\tilde{q}(\tilde{R}_{i-1}) = \tilde{R}_i & \textit{for all } t \in \left[\theta_i + \frac{\log 2}{a_-}, \theta_{i+1} \wedge \tau_\beta(\xi) \wedge (T_0/\nu)\right]. \end{cases} \tag{5.2.27}$$

b) For arbitrary $i \in \mathbb{N}$, $\varepsilon_\star > 0$ with $\tilde{R}_{\star\_}(1+\varepsilon_\star) < \tilde{R}_{\min}$,

$$\|y_{\theta_i}\| \le \tilde{R}_{\star\_}(1+\varepsilon_\star)$$
$$\Rightarrow \begin{cases} |y(t)| \le \tilde{R}_{\star\_}(1+\varepsilon_\star) + \sigma\beta & \textit{for all } t \in \left[\theta_i, \left(\theta_i + \frac{\log 2}{a_-}\right) \wedge \tau_\beta(\xi) \wedge (T_0/\nu)\right], \\ |y(t)| \le \tilde{R}_{\star\_}(1+\varepsilon_\star) & \textit{for all } t \in \left[\theta_i + \frac{\log 2}{a_-}, (\theta_i + \theta_1) \wedge \tau_\beta(\xi) \wedge (T_0/\nu)\right]. \end{cases}$$

Proof. a) The assumption $\tilde{q}(\tilde{R}_k) < 1$ implies for all $k \in \{0, 1, \dots, n\}$ that

$$\gamma\left(\tilde{R}_k + \sigma\beta\right) + \frac{N(\tilde{R}_k + \sigma\beta)^2}{a_-} + \frac{\nu c_0}{a_-} < \tilde{R}_k - 2\sigma\beta < \tilde{R}_k. \tag{5.2.28}$$

For given $l \in \{0, 1, \dots, i\}$, we assume that the assertion (5.2.27) is true for all $\bar{l} \in \{0, 1, \dots, l-1\}$. Then continuing from (5.2.24) for $t_0 = \theta_l$ we observe that

$$\begin{aligned} |y(\theta_l + s)| &\le \tilde{R}_{l-1} e^{-\alpha(\theta_l+s,\theta_l)} \\ &\quad + \left(1 - e^{-\alpha(\theta_l+s,\theta_l)}\right)\left(\gamma\left(\tilde{R}_{l-1} + \sigma\beta\right) + \frac{N(\tilde{R}_{l-1} + \sigma\beta)^2}{a_-} + \frac{\nu c_0}{a_-}\right) \\ &\quad + \sigma|\xi^{(i)}(\theta_l + s)| \quad \text{for all } \theta_l + s \in [\theta_l, (T_0/\nu)], \ l \in \{0, 1, \dots, n-1\}. \end{aligned}$$

Applying (5.2.28) to the second summand for $k = l - 1$ yields

$$|y(\theta_l + s)| \le \tilde{R}_{l-1} + \sigma\beta \text{ for all } \theta_l + s \in [\theta_l, (T_0/\nu) \wedge \tau_\beta(\xi)].$$

Note that the term

$$\tilde{R}_{l-1} e^{-\alpha(\theta_l+s,\theta_l)} + \left(1 - e^{-\alpha(\theta_l+s,\theta_l)}\right)\left(\gamma\left(\tilde{R}_{l-1} + \sigma\beta\right) + \frac{N(\tilde{R}_{l-1} + \sigma\beta)^2}{a_-} + \frac{\nu c_0}{a_-}\right)$$

starts in $\tilde{R}_{l-1}$ for $s=0$ and converges to $\gamma\left(\tilde{R}_{l-1}+\sigma\beta\right)+\frac{N(\tilde{R}_{l-1}+\sigma\beta)^2}{a_-}+\frac{\nu c_0}{a_-}$ monotonically (decreasing) and exponentially fast. Moreover, since $\exp(-\alpha(\theta_l+s,\theta_l))<\frac{1}{2}$ for all $s>\frac{\log 2}{a_-}$ we have that

$$|y(\theta_l+s)|\le\frac{\tilde{R}_{l-1}}{2}+\frac{1}{2}\left(\gamma\left(\tilde{R}_{l-1}+\sigma\beta\right)+\frac{N(\tilde{R}_{l-1}+\sigma\beta)^2}{a_-}+\frac{\nu c_0}{a_-}\right)+\sigma\beta=\tilde{R}_{l-1}\tilde{q}(\tilde{R}_{l-1})$$
$$\text{for all } s\in[\theta_1-r,\theta_1]\cap[0,\tau_\xi(\beta)\wedge(T_0/\nu))\,.$$

b) Let us for a moment denote $\overline{R}=\tilde{R}_{\star\_}(1+\varepsilon_\star)$, and if we assume that $\|y_{\theta_i}\|\le\overline{R}$, then, we may deduce with the same arguments as above, restarting the segment process in θ_i, that

$$\begin{aligned}|y(\theta_i+s)|&\le\overline{R}e^{-\alpha(\theta_i+s,\theta_i)}+\left(1-e^{-\alpha(\theta_l+s,\theta_l)}\right)\left(\gamma\left(\overline{R}+\sigma\beta\right)+\frac{N(\overline{R}+\sigma\beta)^2}{a_-}+\frac{\nu c_0}{a_-}\right)\\&\quad+\sigma\left|\xi^{(i)}(\theta_i+s)\right|\\&\le\overline{R}+\sigma\beta\qquad\text{for all } s\in[0,(T_0/\nu)\wedge\tau_\beta(\xi)].\end{aligned}$$

And in the same way,

$$|y(\theta_i+s)|\le\frac{\overline{R}}{2}+\frac{1}{2}\left(\gamma\left(\overline{R}+\sigma\beta\right)+\frac{N(\overline{R}+\sigma\beta)^2}{a_-}+\frac{\nu c_0}{a_-}\right)+\sigma\beta\le\overline{R}$$
$$\text{for all } s\in[\theta_1-r,\theta_1]\cap[0,\tau_\beta(\xi)\wedge(T_0/\nu)],$$

where we have used that $\tilde{q}(\overline{R})\le 1$. □

Remark 5.16.

- *The first part of Lemma 5.15 shows that it is in principle possible to apply the decay argument sequentially, where subsequent results provide bounds for the decay speed and the limiting size of $\tilde{R}_n$. The iteration principally works as long as $\tilde{q}(\tilde{R}_n)<1$, and it stucks if $\tilde{R}_n$ is close to $\tilde{R}_{\star\_}$*
- *The second part of the lemma shows that the deviation y remains within a tube of radius $\tilde{R}_{\star\_}(1+\varepsilon_\star)+\sigma\beta$ at least up to the time $\tau_\beta(\xi)\wedge T_0/\nu$, which means, as long as the stochastic perturbation behaves friendly.*

Repeating the ideas from the deterministic case, the following lemma provides a uniform upper bound of the decay factor in case $\tilde{R}_n>\tilde{R}_{\star\_}(1+\varepsilon_\star)$ for some $\varepsilon_\star>0$.

Corollary 5.17. *For arbitrary $x\in(\tilde{R}_{\star\_},\tilde{R}_{\min})$, we have that*

$$\tilde{q}(\tilde{R}_{\star\_}(1+\varepsilon_\star))\le 1-\frac{\tilde{q}(\tilde{R}_{\star\_})-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min}-\tilde{R}_{\star\_}}\varepsilon_\star\tilde{R}_{\star\_}=1-\frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}}\,\varepsilon_\star\left(1+\mathcal{O}\left(\sqrt{N\tilde{c}}+N\sigma\beta\right)\right).$$

where

$$\begin{aligned}\frac{\tilde{q}(\tilde{R}_{\star\_})-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min}-\tilde{R}_{\star\_}}\tilde{R}_{\star\_}&=\left(\frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}}-\frac{N\tilde{c}}{a_-(1-\gamma)}-\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{a_-^{3/2}(1-\gamma)}\right)\\&\quad\cdot\left(1+\mathcal{O}\left(\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{(1-\gamma)^2}+\frac{\sqrt{N\tilde{c}}}{1-\gamma}+\frac{\sigma\beta N}{1-\gamma}+\frac{N\tilde{c}}{(1-\gamma)^2}\right)\right).\end{aligned}$$

Proof. The first inequality is straightforwardly following the arguments of the previous case, and the rest is mere cumbersome computation. First, we observe that

$$\begin{aligned}\tilde{R}_{\min} - \tilde{R}_{\star_-} &= \sqrt{\frac{a_-\tilde{c}}{N}}\left(1 - \sqrt{\frac{N\tilde{c}}{a_-}}\frac{1}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right)\right)\\ &= \sqrt{\frac{a_-\tilde{c}}{N}}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{(1-\gamma)^2} + \frac{N^{3/2}\tilde{c}^{3/2}}{(1-\gamma)^3} + \frac{\sqrt{N\tilde{c}}}{1-\gamma}\right)\right).\end{aligned}$$

That leads to

$$\begin{aligned}\frac{1-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min} - \tilde{R}_{\star_-}} = &\left(\frac{1-\gamma}{2} - \sqrt{\frac{N\tilde{c}}{a_-}} - \frac{N}{a_-}\sigma\beta\right)\sqrt{\frac{N}{a_-\tilde{c}}}\\ &\cdot\left(1 + \mathcal{O}\left(\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{(1-\gamma)^2} + \frac{N^{3/2}\tilde{c}^{3/2}}{(1-\gamma)^3} + \frac{\sqrt{N\tilde{c}}}{1-\gamma}\right)\right),\end{aligned}$$

and then we end up with

$$\begin{aligned}&\frac{1-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min} - \tilde{R}_{\star_-}}\tilde{R}_{\star_-}\\ &= \left(\frac{1-\gamma}{2} - \sqrt{\frac{N\tilde{c}}{a_-}} - \frac{N}{a_-}\sigma\beta\right)\sqrt{\frac{N}{a_-\tilde{c}}}\frac{\tilde{c}}{1-\gamma}\\ &\quad\cdot\left(1 + \mathcal{O}\left(\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{(1-\gamma)^2} + \frac{N^{3/2}\tilde{c}^{3/2}}{(1-\gamma)^3} + \frac{\sqrt{N\tilde{c}}}{1-\gamma}\right)\right)\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right)\\ &= \left(\frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}} - \frac{N\tilde{c}}{a_-(1-\gamma)} - \frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{a_-^{3/2}(1-\gamma)}\right)\\ &\quad\cdot\left(1 + \mathcal{O}\left(\frac{\sigma\beta N^{3/2}\sqrt{\tilde{c}}}{(1-\gamma)^2} + \frac{\sqrt{N\tilde{c}}}{1-\gamma} + \frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right).\end{aligned}$$

which is the claim. □

The characterizing property of the uniformly stable phase is the property that $|b(\cdot)|/a(\cdot) < \gamma$, where $\gamma < 1$ is bounded away from 1, see (5.2.6). Therefore, the implications of Corollary 5.17 can be further simplified to

$$\frac{1-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min} - \tilde{R}_{\star_-}}\tilde{R}_{\star_-} = \frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}}\left(1 - \mathcal{O}\left(\sqrt{N\tilde{c}}\right)\right).$$

Theorem 5.18. *Assume that $\|y_0\| \le R_0$ and that $\sigma < \frac{\nu}{|\log\nu|}$ and $\beta = \mathcal{O}(|\log\nu|)$. Assume further that ν is small enough such that there is $\delta > 0$ of order 1 such that for ζ from (5.2.22) and appropriate $\tilde{a}_-, \tilde{a}_+ > 0$ we have*

$$\frac{1}{2\tilde{a}_+} \le \frac{1-\delta\nu}{2a(t)} \le \zeta(t) \le \frac{1+\delta}{2a(t)} \le \frac{1}{2\tilde{a}_-} \qquad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.29}$$

a) Let $R_0 \in (\tilde{R}_{\star-}, \tilde{R}_{\min})$. *Let further* $\varepsilon_\star > 0$ *such that* $\tilde{R}_{\star-}(1+\varepsilon_\star) < \tilde{R}_{\min}$. *If*

$$n_\star \geq \left|\log\left(\frac{\tilde{R}_{\star-}(1+\varepsilon_\star)}{R_0}\right)\right| \frac{2}{\varepsilon_\star}\sqrt{\frac{a_-}{N\tilde{c}}}\left(1+\mathcal{O}\left(\sqrt{N\tilde{c}}\right)\right),$$

then

$$|y(t)| \leq \tilde{R}_{\star-}(1+\varepsilon_\star) + \sigma\beta \in \mathcal{O}(\nu) \qquad \textit{for all } t \in [\theta_{n_\star}, \tau_\xi(\beta) \wedge (T_0/\nu)),$$

and,

$$\mathbb{P}\left\{\tau_\beta(\xi) < T_0/\nu\right\} \leq \frac{4e\beta^2 e T_0(1+\mathcal{O}(\nu))}{\theta_1 \nu} \exp\left(-\beta^2 \tilde{a}_-\right) \qquad \textit{for } \beta > 0, \tag{5.2.30}$$

where integer-value restrictions have been ignored.

b) If $\|y_0\| \leq R_0 \in (\tilde{R}_{\min}, \tilde{R}_{\star+})$, *we obtain that*

$$\left(t \geq \theta_{m_\star} \textit{ with } m_\star = \frac{1}{\log \tilde{q}(R_0)} \log \frac{\tilde{R}_{\min}}{R_0}\right) \quad \Rightarrow \quad |y(t)| \leq \tilde{R}_{\min},$$

where again, integer-value restrictions have been ignored.

Proof. Due to the definition $\tilde{R}_i = \tilde{R}_{i-1}\tilde{q}(\tilde{R}_{i-1})$, $i\ in \{0,1,\ldots\}$, we remember that $|y(t)| \leq \tilde{R}_i$ for $t \in [\theta_i - r, \theta_i]$ for all $i \in \mathbb{N}$ satisfying $\tilde{R}_i \geq \tilde{R}_{\star-}$ by Lemma 5.15. And we also know that $\tilde{q}(\tilde{R}_i) \leq \tilde{q}(\tilde{R}_{\star-}(1+\varepsilon_\star))$. Therefore,

$$\begin{aligned} &R_0\left(\tilde{q}(\tilde{R}_{\star-}(1+\varepsilon_\star))\right)^{n_\star} < \tilde{R}_{\star-}(1+\varepsilon_\star) \\ &\qquad \Rightarrow \quad |y(t)| \leq \tilde{R}_{\star-}(1+\varepsilon_\star) + \sigma\beta \qquad \text{for all } t \in [\theta_{n_\star}, \tau_\beta(\xi) \wedge (T_0/\nu)). \end{aligned}$$

The left-hand condition is equivalent to

$$n_\star > \left|\log \frac{\tilde{R}_{\star-}(1+\varepsilon_\star)}{R_0}\right| \frac{1}{|\log \tilde{q}(\tilde{R}_{\star-}(1+\varepsilon_\star))|}. \tag{5.2.31}$$

Moreover, with Corollary 5.17 and $\frac{1}{|\log(1-x)|} \leq \frac{1}{x}$ for all $x \in (0,1)$, we know that

$$\begin{aligned} |\log \tilde{q}(\tilde{R}_{\star-}(1+\varepsilon_\star))| &\geq 1 - \tilde{q}(\tilde{R}_{\star-}(1+\varepsilon_\star)) \\ &\geq \frac{1-\tilde{q}(\tilde{R}_{\min})}{\tilde{R}_{\min} - \tilde{R}_{\star-}} \tilde{R}_{\star-}\varepsilon_\star = \frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}}\,\varepsilon_\star\left(1+\mathcal{O}\left(\sqrt{N\tilde{c}}\right)\right). \end{aligned}$$

Therefore, condition (5.2.31) is satisfied if

$$\begin{aligned} n_\star &> \left|\log \frac{\tilde{R}_{\star-}(1+\varepsilon_\star)}{R_0}\right| \frac{1}{\frac{1}{2}\sqrt{\frac{N\tilde{c}}{a_-}}\,\varepsilon_\star\left(1+\mathcal{O}\left(\sqrt{N\tilde{c}}\right)\right)} \\ &= \left|\log \frac{\tilde{R}_{\star-}(1+\varepsilon_\star)}{R_0}\right| \frac{2}{\varepsilon_\star}\sqrt{\frac{a_-}{N\tilde{c}}}\left(1+\mathcal{O}\left(\sqrt{N\tilde{c}}\right)\right). \end{aligned}$$

Furthermore, with regard to (5.2.23), note that there are $\frac{T_0}{\theta_1 \nu}$ steps of size θ_1 and $\alpha(\theta_i, \theta_{i+1}) \geq \tilde{a}_- \theta_1$. Therefore,

$$\begin{aligned}
\mathbb{P}\left\{\tau_\beta(\xi) < T_0/\nu\right\} &\leq \sum_{i=0}^{n-1} \mathbb{P}\left\{\sup_{t \in [\theta_i, \theta_{i+1}]} \frac{|\xi^{(i)}(t)|}{\sqrt{\zeta(t)}} > \beta\sqrt{2\tilde{a}_-}\right\} \\
&\leq \frac{T_0}{\theta_1 \nu} \frac{4\tilde{a}_- e \theta_1 \beta^2 (1 + \mathcal{O}(\nu))}{\tilde{a}_- \theta_1} \exp\left(-\beta^2 \tilde{a}_-\right) \\
&= \frac{4e\beta^2 e T_0 (1 + \mathcal{O}(\nu))}{\theta_1 \nu} \exp\left(-\beta^2 \tilde{a}_-\right) \qquad \text{for } \beta > 0.
\end{aligned}$$

Part b) is much easier, because consecutive iterations $\tilde{R}_k = \tilde{R}_{k-1}\tilde{q}(\tilde{R}_{k-1})$, $k \in \mathbb{N} \cup \{0\}$, yield improving factors of decay $\tilde{q}(R_0) \geq \tilde{q}(\tilde{R}_1) \ldots$ as long as $\tilde{R}_k \geq \tilde{R}_{\min}$. □

Remark 5.19.
- *The results presented in the previous theorem beautifully capture the behavior of a delay-influenced solution of* (5.1.3) *when it is not initiated close to the equilibrium branch. The corollary allows an initial maximal distance of order* 1 *between the initial segment and the* ν*-adiabatic solution. A (slow) time of order* $|\log(\nu)|/\sqrt{\nu}$ *suffices for the solution to approach the* ν*-adiabatic solution, and therefore also the equilibrium branch* $x^\star$*, up to a distance of order* ν*. Furthermore, it actually suffices to choose* β *of order* $|\log \nu|$.

5.2.2. Bernstein-based Approach

The goal of this subsection is to establish concentration of solution paths of (5.1.3) initiated in a close neighborhood to the ν-adiabatic solution $x^{\mathrm{ad}_\nu} = (x^{\mathrm{ad}}(t))_{t\in[-r,T_0/\nu]}$ of the replacement system. As an innovation of this section, we will study concentration inequalities in the formulation that was consistently used in [BG06] and focuses on excursions of the process related to its standard deviation; an instance of which we have seen in (5.2.23). We continue to study the deviation $y = (y(t))_{t\in[0,T_0/\nu]}$ of the delay-influenced solution $x = (x(t))_{t\in[0,T_0/\nu)}$ from a ν-adiabatic solution x^{ad_ν} of the corresponding replacement system in a uniformly stable regime that we introduced at the beginning of this section. That especially includes that assumption (5.2.6) holds true. A particularly modified linearization $(y^{\mathrm{lin}}(t))_{t\in[0,T_0/\nu]}$ of the replacement system (5.1.4) is given as the unique solution of

$$\begin{cases} dy^{\mathrm{lin}}(t) = -a(t)y^{\mathrm{lin}}(t) + |b(t)|y^{\mathrm{lin}}(t) + \sigma dW(t) & \text{for all } t \in [0, T_0/\nu], \\ y^{\mathrm{lin}}(0) = y(0). \end{cases}$$

In analogy to the discussion at the beginning of the previous Subsection 5.2.1, the according rescaled variance process $(v(t))_{t\in[0,T_0/\nu]}$, defined by $v(t) = \sigma^{-2} \operatorname{var} y^{\mathrm{lin}}(t)$ for $t \in [0, T_0/\nu]$, fulfills the differential law $dv(t) = -2\big(a(t)-|b(t)|\big)v(t)dt+1$ for all $t \in [0, T_0/\nu)$, and that one admits the equilibrium branch $t \mapsto \frac{1}{2}(a(t) - |b(t)|)^{-1}$ together with a ν-adiabatic solution $(v^{\mathrm{ad}}(t))_{t\in[0,T_0/\nu]}$ with $v^{\mathrm{ad}}(t) = \frac{1}{2}(a(t) - |b(t)|)^{-1} + \mathcal{O}(\nu)$ for all $t \in [0, T_0/\nu]$ for small ν. Again, the results can be directly deduced from the appendix A.1. In other words there are $\nu_0, \delta_0 > 0$ such that

$$\frac{1-\delta_0\nu}{2(a(t)-|b(t)|)} \le v^{\mathrm{ad}}(t) \le \frac{1+\delta_0\nu}{2(a(t)-|b(t)|)} \qquad \text{for all } \nu < \nu_0,\ t \in [0, T_0/\nu] \tag{5.2.32}$$

which is due to the fact that the $a(\cdot) - |b(\cdot)| \ge a_-(1-\gamma) \in \mathcal{O}(1)$. In case of the process $(y^{\mathrm{lin}}(t))_{t\ge 0}$ due to the Appendix A.1 the following concentration inequality holds:

$$\mathbb{P}\left\{ \sup_{t\in[0,T_0/\nu]} \frac{|y^{\mathrm{lin}}(t)|}{\sqrt{v^{\mathrm{ad}}(t)}} > h \right\} \le \frac{2eT_0h^2\big(1+\mathcal{O}(\nu)\big)}{\sigma^2\nu\gamma(T_0/\nu)} \exp\left(-\frac{h^2}{2\sigma^2}\right) \qquad \text{for } h > 0,$$

where $\gamma(T/\nu) = \int_0^{T_0/\nu} a(u) - |b(u)|du$. The goal of this section is to establish a concentration inequality for $(y(t))_{t\in[0,T_0/\nu]}$ that shows resemblance to the one above for y^{lin}. The subsequent extension of the techniques from [BG06] to time-delayed perturbations comes with a price: In order to study the probability for events of the form

$$\left\{ \sup_{s\in[0,T_0/\nu]} \frac{|y(s)|}{\sqrt{\zeta(s)}} > h \right\}, \tag{5.2.33}$$

we are compelled to assume $(\zeta(t))_{t\in[-r,T_0/\nu]}$ nondecreasing. Therefore, for the course of this subsection, we let

$$\zeta(t) := \sup_{s\in[0,t]} v^{\mathrm{ad}}(s) \qquad \text{for all } t \in [0, T_0/\nu]. \tag{5.2.34}$$

With the notation $a_+ = \sup_{t\in[0,T_0/\nu]} a(t)$ we derive useful upper and lower bounds in the following lemma.

Lemma 5.20. *Assume that ν is small enough such that the inequality in* (5.2.32) *holds true. For $(\zeta(t))_{t\in[0,T_0/\nu]}$, defined as in* (5.2.34)*, we have that*

$$\frac{1}{2a_+}(1-\delta_0\nu) \le \zeta(s) \le \frac{1}{2a_-}\frac{1}{1-\gamma}(1+\delta_0\nu) \quad \textit{for all } s\in[0,T_0/\nu].$$

Proof. The lower bound is easily derived, and for the upper bound we find that

$$v^{\mathrm{ad}}(s) \le \frac{1+\delta_0\nu}{2(a(s)-|b(s)|)} = \frac{1+\delta_0\nu}{2a(s)\frac{a(s)-|b(s)|}{a(s)}} \le \frac{1}{2a_-\left(1-\frac{|b(s)|}{a(s)}\right)}(1+\delta_0\nu).$$

Together with the basic assumption (5.2.6) that shows the claim. □

For $\nu>0$ small enough, some arbitrarily small $t_\Delta>0$ and $h>0$ we will see that

$$\mathbb{P}\left\{\sup_{s\in[0,T_0/\nu]}\frac{|y(s)|}{\sqrt{\zeta(s)}}>h\right\} \le \frac{T_0}{\nu t_\Delta}\exp\left(-(1-\gamma)\frac{h^2}{2\sigma^2}\left(1+\mathcal{O}\left(\frac{\|y_0\|\vee\nu}{h}+Nh+t_\Delta\right)\right)\right),$$

which is only useful, if

$$\max\left\{\sigma,\|y_0\|,\nu\right\} < h < \frac{1}{N}. \tag{5.2.35}$$

We make this restriction an assumption in order to omit error terms that are not of leading order. From the preceding part, inequality (5.2.24) yields for starting point $t_0=0$

$$\begin{aligned}|y(t)| \le \|y_0\|e^{-\alpha(t)} &+ \int_0^t e^{-\alpha(t,u)}|b(u)||y(u-r)|du \\ &+ \int_0^s e^{-\alpha(t,u)}\left(N_f y^2(u)+N_g y^2(u-r)+\nu c_0\right)du \\ &+ \sigma\left|\int_0^t e^{-\alpha(t,u)}dW(u)\right| \quad \text{for all } s\in[0,T_0/\nu\wedge\tau_{\mathbb{D}}(x)],\end{aligned}$$

where $\tau_{\mathbb{D}}(x) := \inf\{t\ge 0 : (x(t),\nu t)\notin\mathbb{D}\}$ denote the first time that x leaves $\mathbb{D}$, and c_0 is still the same we defined in (5.2.14). In order to emphasize the martingale part of the stochastic-integral term, let us introduce the notation

$$\mathcal{M}(t) := \int_0^t e^{\alpha(u)}dW(u) \quad \text{for all } t\in[0,T_0/\nu].$$

We define the family of auxiliary events

$$E_t := \left\{\omega\in\Omega : \sup_{s\in[0,t]}\left|\frac{\sigma e^{-\alpha(s)}\mathcal{M}(s)}{\sqrt{\zeta(s)}}\right| + \frac{1}{\sqrt{\zeta(s)}}\|y_0\|e^{-\alpha(s)}\right. \tag{5.2.36}$$

$$\left. + \frac{\int_0^s e^{-\alpha(s,u)}\left(\nu c_0+|b(u)|h\sqrt{\zeta(u)}+Nh^2\zeta(u)\right)du}{\sqrt{\zeta(s)}} > h\right\} \text{ for all } t\in[0,T_0/\nu]. \tag{5.2.37}$$

And for given $t\in[0,T_0/\nu]$, we denote the event that the deviation y has left the tube of radius h at least once over $[0,t]$ by

$$A_t := \left\{ \sup_{s\in[0,t]} \frac{|y(s)|}{\sqrt{\zeta(s)}} > h \right\} \quad \text{for all } t \in [0,T_0/\nu] \quad \text{and } \tau_h = \inf\left\{ t \geq 0 : \frac{|y(t)|}{\sqrt{\zeta(t)}} > h \right\}, \tag{5.2.38}$$

where both definitions are reasonably defined after the following technical assumption concerning a suitable size of the domain:

Assumption 5.21. *We let h be small enough that*

$$\left\{(z,\nu t) : t \in [0,T_0/\nu] \text{ and } |z - x^{\mathrm{ad}_\nu}(t)| \leq h\sqrt{\zeta(t)} \right\} \subseteq \mathbb{D},$$

which informally ensures that we always see y leaving the $h\sqrt{\zeta(\cdot)}$-tube before x leaves $\mathbb{D}$.

Lemma 5.22. *Under the Assumption 5.21 we find the following properties of the involved families of sets defined in* (5.2.37) *and* (5.2.38).

a) We have that $A_t \subseteq E_t$ for all $t \geq 0$, or in other words $\omega \notin E_t \Rightarrow \omega \notin A_t$.

b) For an arbitrary integer $n \in \mathbb{N}$, we let $0 = t_0 < t_1 < \ldots < t_n = T_0/\nu$ be a partition of the time interval $[0,T_0/\nu]$. Then

$$\mathbb{P}\big(E_{T_0/\nu}\big) = \sum_{i=0}^{n-1} \mathbb{P}\big(E_{t_{i+1}} \cap E_{t_i}^c\big).$$

Proof. a) Pathwise interpretation of the stochastic-integral term is justified through the integrand's finite variation and serves with estimate (5.2.24)

$$\begin{aligned}
h &= \left| \frac{y(\tau_h)}{\sqrt{\zeta(\tau_h)}} \right| \\
&\leq \frac{1}{\sqrt{\zeta(\tau_h)}} \Bigg(\quad e^{-\alpha(\tau_h)}\|y_0\| \\
&\qquad + \int_0^{\tau_h} e^{-\alpha(\tau_h,u)} \Big(\nu c_0 + |b(u)||y(u-r)| + N_f y^2(u) + N_g y^2(u-r)\Big) du \\
&\qquad + \sigma e^{\alpha(\tau_h)} \big|\mathcal{M}(\tau_h)\big| \quad \Bigg) \\
&\leq \frac{1}{\sqrt{\zeta(\tau_h)}} \Bigg(\quad e^{-\alpha(\tau_h)}\|y_0\| \\
&\qquad + \int_0^{\tau_h} e^{-\alpha(\tau_h,u)} \Big(\nu c_0 + |b(u)| h\sqrt{\zeta(u)} + N_f h^2\zeta(u) + N_g h^2\zeta(u)\Big) du \\
&\qquad + \sigma e^{-\alpha(\tau_h)} \big|\mathcal{M}(\tau_h)\big| \quad \Bigg) \qquad \text{for all } \omega \in \{\tau_h < T_0/\nu\}.
\end{aligned}$$

That directly yields the first claim.
b) As the family $(E_{t_i})_{i\in\{0,1,\ldots,n\}}$ is obviously increasing, the events $E_{t_{i+1}} \cap E_{t_i}^c$, $i \in \{0,1,\ldots,n-1\}$, are pairwise disjoint and together constitute $E_{T_0/\nu}$. That shows the second part. □

Then, for arbitrary $i \in I := \{0, \dots, N-1\}$, we find that

$$\mathbb{P}\Big\{E_{t_i}^c \cap E_{t_{i+1}}\Big\}$$
$$\leq \mathbb{P}\Bigg\{ \sup_{s \in [t_i, t_{i+1}]} \frac{\sigma e^{-\alpha(s)}|\mathcal{M}(s)| + \|y_0\| e^{-\alpha(s)} + \int_0^s e^{-\alpha(s,u)}\Big(\nu c_0 + |b(u)| h\sqrt{\zeta(u)} + Nh^2\zeta(u)\Big)du}{\sqrt{\zeta(s)}} > h\Bigg\}. \tag{5.2.39}$$

We use the following short-hand notations

$$F(s) := \nu c_0 \int_0^s e^{-\alpha(s,u)}du + e^{-\alpha(s)}\|y_0\|, \quad G(s) := \int_0^s e^{-\alpha(s,u)}|b(u)|\sqrt{\zeta(u)}\,du,$$
$$H(s) := N\int_0^s e^{-\alpha(s,u)}\zeta(u)du, \qquad V(s) := \int_0^s e^{-2\alpha(s,u)}du \quad \text{for all } s \in [0, T_0/\nu].$$

It is easy to see that

(i) The mapping $F : [0, T_0/\nu] \to \mathbb{R}$ is the unique solution of the ODE

$$\begin{cases} \dot F(t) = -a(t) + \nu c_0 & \text{for } t \in [0, T_0/\nu), \\ F(0) = \|y_0\|. \end{cases}$$

(ii) The mapping $G : [0, T_0/\nu] \to \mathbb{R}$ is the unique solution of

$$\begin{cases} \dot G(t) = -a(t)G(t) + |b(t)|\sqrt{\zeta(t)} & \text{for } t \in [0, T_0/\nu), \\ G(0) = 0. \end{cases}$$

(iii) The mapping $H : [0, T_0/\nu] \to \mathbb{R}$ is the unique solution of

$$\begin{cases} \dot H(t) = -a(t)H(t) + N\zeta(t) & \text{for } t \in [0, T_0/\nu), \\ H(0) = 0. \end{cases}$$

(iv) The mapping $V : [0, T_0/\nu] \to \mathbb{R}$ is the unique solution of

$$\begin{cases} \dot V(t) = -2a(t) + 1 & \text{for } t \in [0, T_0/\nu), \\ V(0) = 0. \end{cases}$$

Lemma 5.23. *With the notations from above, we have that*

a)
$$F(s) \text{ “follows” } \frac{\nu c_0}{2a(s)} \quad \text{and} \quad F(s) \leq \|y_0\| \vee \frac{\nu c_0}{2a_-} \text{ for all } s \in [0, T_0/\nu].$$

b) There is $\nu_1 > 0$ such that there is a constant $\delta_1 > 0$ independent of ν that fulfills all three *below estimates for all $\nu \le \nu_1$:*

$$G(s) \text{ ``follows'' } \frac{|b(s)|}{a(s)}\sqrt{\zeta(s)} \quad \text{and} \quad G(s) \le \frac{|b(s)|}{a(s)}\sqrt{\zeta(s)} + \delta_1 \nu \text{ for all } s \in [0, T_0/\nu], \tag{5.2.40}$$

$$H(s) \text{ ``follows'' } N\frac{\zeta(s)}{a(s)} \quad \text{and} \quad H(s) \le N\frac{\zeta(s)}{a(s)} + \delta_1 \nu \text{ for all } s \in [0, T_0/\nu], \tag{5.2.41}$$

$$V(s) \text{ ``follows'' } \frac{1}{2a(s)} \quad \text{and} \quad V(s) \le \frac{1+\delta_1\nu}{2a(s)} \text{ for all } s \in [0, T_0/\nu]. \tag{5.2.42}$$

Proof. This is due to Fenichel's theory, which ensures the existence of ν-adiabatic solutions that *follow* the respective slow manifolds in a distance of order $\mathcal{O}(\nu)$. The assertions of the lemma follow from the fact that solution paths can not intersect in case of ordinary differential equations. In (5.2.42) we have additionally used that $a(\cdot) \le a_+$ over $[0, T_0/\nu]$. □

Without loss of generality, from now on, we assume that $\nu_0 = \nu_1$ and $\delta_0 = \delta_1$, where ν_0, δ_0 are defined in (5.2.32).

Lemma 5.24. *Let $t_\Delta = \sup_{i\in I} |t_{i+1} - t_i|$ denote the maximum step width of the partition $0 = t_0 < \ldots < t_n = T_0/\nu$. Let further*

$$\mu_1 := \Big((d_f + d_g)\delta + N\delta^2\nu\Big) \sup_{(z,u)\in\mathbb{D}} |f_{xx}(z,u)| + \sup_{(z,u)\in\mathbb{D}} |f_{xt}(z,u)|,$$

$$\mu_2 := \mu_1 + \Big((d_f + d_g)\delta + N\delta^2\nu\Big) \sup_{(z,u)\in\mathbb{D}} |g_{xx}(z,u)| + \sup_{(z,u)\in\mathbb{D}} |g_{xt}(z,u)|.$$

Then

$$\inf_{i\in\{0,\ldots,n-1\}} \inf_{s\in[t_i,t_{i+1})} \frac{a(t_{i+1})}{a(s)} \ge 1 - \frac{\nu\mu_1}{a_-} t_\Delta, \tag{5.2.43}$$

$$\inf_{i\in\{0,\ldots,n-1\}} \inf_{s\in[t_i,t_{i+1})} \frac{a(s) - |b(s)|}{a(t_i) - |b(t_i)|} \ge 1 - \frac{\nu\mu_2}{a_-(1-\gamma)} t_\Delta, \tag{5.2.44}$$

$$\inf_{i\in\{1,\ldots,n\}} \exp\big(-2\alpha(t_i, t_{i+1})\big) \ge 1 - 2a_+ t_\Delta. \tag{5.2.45}$$

Proof. The ν-adiabatic solution x^{ad_ν} is a solution of the replacement system and therefore, for each $t \in (0, T_0/\nu)$ we find that

$$\begin{aligned}
\frac{d}{dt} x^{\mathrm{ad}_\nu}(t) &= f(x^{\mathrm{ad}_\nu}(t), \nu t) + g(x^{\mathrm{ad}_\nu}(t), \nu t) \\
&= f(x^\star(t) + x^{\mathrm{ad}_\nu}(t) - x^\star(t), \nu t) + g(x^\star(t) + x^{\mathrm{ad}_\nu}(t) - x^\star(t), \nu t) \\
&= f_x(x^\star(t), \nu t)(x^{\mathrm{ad}_\nu}(t) - x^\star(t)) + g_x(x^\star(t), \nu t)(x^{\mathrm{ad}_\nu}(t) - x^\star(t)) \\
&\quad + \mathcal{R}_f(x^{\mathrm{ad}_\nu}(t) - x^\star(t), \nu t) + \mathcal{R}_g(x^{\mathrm{ad}_\nu}(t) - x^\star(t), \nu t).
\end{aligned}$$

Due to assumption (5.2.2), we obtain that

$$\left|\frac{d}{dt} x^{\mathrm{ad}_\nu}(t)\right| \le d_f\delta\nu + d_g\delta\nu + N\delta^2\nu^2 \quad \text{for all } t \in (0, T_0/\nu). \tag{5.2.46}$$

For the differential of $a(\cdot)$ we observe that for all $u \in (0, T_0/\nu)$,

$$\frac{d}{du} a(u) = \frac{d}{du} f_x(x^{\mathrm{ad}_\nu}(u), \nu u) = f_{xx}(x^{\mathrm{ad}_\nu}(u), \nu u) \frac{dx^{\mathrm{ad}_\nu}(u)}{du} + \nu f_{xt}(x^{\mathrm{ad}_\nu}(u), \nu u). \tag{5.2.47}$$

Altogether, we have that

$$\frac{a(t_{i+1})}{a(s)} \geq 1 - \frac{1}{a_-}\int_s^{t_{i+1}} \frac{d}{du}a(u)du \geq 1 - \frac{\nu\mu_1}{a_-}t_\Delta \quad \text{for all } s \in [t_i, t_{i+1}], i \in I.$$

The second inequality can be seen as follows. First, observe that

$$a(t_i) - |b(t_i)| = a(t_i)\left(1 - \frac{|b(t_i)|}{a(t_i)}\right) \geq a_-(1-\gamma) \quad \text{for all } i \in I.$$

And, as $(-1)(|b(t_i)| - |b(s)|) \leq |b(t_i) - b(s)|$, we have that

$$\begin{aligned}\frac{a(s) - |b(s)|}{a(t_i) - |b(t_i)|} &= 1 - \frac{a(t_i) - |b(t_i)| - a(s) + |b(s)|}{a(t_i) - |b(t_i)|} \\ &\geq 1 - \frac{a(t_i) - a(s) + |b(t_i) - b(s)|}{a_-(1-\gamma)} \quad \text{for all } s \in [t_i, t_{i+1}], i \in I. \end{aligned} \tag{5.2.48}$$

Then, for all $s \in [t_i, t_{i+1}], i \in I$ we have that

$$\begin{aligned} b(t_i) - b(s) &= \int_s^{t_i} \frac{db(u)}{du}du \\ &= \int_s^{t_i} g_{xx}(x^{\mathrm{ad}_\nu}(u), \nu u)\frac{dx^{\mathrm{ad}_\nu}(u)}{du}du + \int_s^{t_i} g_{xt}(x^{\mathrm{ad}_\nu}(u), \nu u)\nu du. \end{aligned}$$

Then, an application of (5.2.46) provides for all $s \in [t_i, t_{i+1}], i \in I$ that

$$|b(t_i) - b(s)| \leq t_\Delta \nu \left(\sup_{(z,u)\in\mathbb{D}} |g_{xx}(z,u)| \left((d_f + d_g)\delta + N\delta^2\nu\right) + \sup_{(z,u)\in\mathbb{D}} |g_{xt}(z,u)| \right). \tag{5.2.49}$$

And analogously with the help of (5.2.47), we find for all $s \in [t_i, t_{i+1}], i \in I$ that

$$\begin{aligned} a(t_i) - a(s) &= \int_s^{t_i} \frac{da(u)}{du}du \\ &= \int_s^{t_i} f_{xx}(x^{\mathrm{ad}_\nu}(u), \nu u)\frac{dx^{\mathrm{ad}_\nu}(u)}{du}du + \int_s^{t_i} f_{xt}(x^{\mathrm{ad}_\nu}(u), \nu u)\nu du. \end{aligned}$$

And therefore, using (5.2.46) we find that

$$|a(t_i) - a(s)| \leq t_\Delta \nu \left(\sup_{(z,u)\in\mathbb{D}} |f_{xx}(z,u)| \left((d_f + d_g)\delta + N\delta^2\nu\right) + \sup_{(z,u)\in\mathbb{D}} |f_{xt}(z,u)| \right) \quad \text{for all } s \in [t_i, t_{i+1}], i \in I. \tag{5.2.50}$$

From estimates (5.2.49) and (5.2.50) we reveive for all $s \in [t_i, t_{i+1}], i \in I$ that

$$\begin{aligned} &a(t_i) - a(s) + |b(t_i) - b(s)| \\ &\leq t_\Delta \nu \left(\sup_{(z,u)\in\mathbb{D}} |f_{xx}(z,u)| \left((d_f + d_g)\delta + N\delta^2\nu\right) + \sup_{(z,u)\in\mathbb{D}} |f_{xt}(z,u)| \right) \\ &\quad + t_\Delta \nu \left(\sup_{(z,u)\in\mathbb{D}} |g_{xx}(z,u)| \left((d_f + d_g)\delta + N\delta^2\nu\right) + \sup_{(z,u)\in\mathbb{D}} |g_{xt}(z,u)| \right), \end{aligned}$$

which is the claim when plugged into (5.2.48).

For the third inequality, we use that $e^{-z} \geq 1 - z$ for arbitrary $z \in \mathbb{R}$ and that

$$\alpha(t_{i+1}, t_i) = \int_{t_i}^{t_{i+1}} a(u)du \leq a_+ t_\Delta \quad \text{for all } i \in I.$$

$\square$

Theorem 5.25. *Let t_Δ be given as in Lemma 5.24 and assume that ν is small enough such that the estimate on v^{ad} in (5.2.32) and the inequalities of part b) of Lemma 5.23 hold true. Then, for $h > 0$, we have that*

$$\mathbb{P}\left\{ \sup_{s\in[0,T_0/\nu]} \frac{|y(s)|}{\sqrt{\zeta(s)}} > h \right\} \leq \frac{T_0}{\nu t_\Delta} \exp\left(-(1-\gamma)\frac{h^2}{2\sigma^2}\Big(1 + \mathscr{R}(\nu, \|y_0\|, h, t_\Delta)\Big)\right),$$

where $\mathscr{R}(\nu, \|y_0\|, h, t_\Delta) = \mathcal{O}\left(\frac{\|y_0\| \vee \nu}{h} + Nh + t_\Delta\right)$.

Proof. Rearranging terms in (5.2.39) yields

$$\begin{aligned}
&\mathbb{P}\Big\{E_{t_i} \cap E_{t_{i+1}}\Big\} \\
&\quad \leq \mathbb{P}\Bigg\{ \sup_{s\in[t_i,t_{i+1})} \sigma|\mathcal{M}(s)| > \Bigg(h - \sup_{s\in[t_i,t_{i+1})} \frac{1}{\sqrt{\zeta(s)}}\Big(F(s) + G(s)h + H(s)h^2\Big)\Bigg) \\
&\qquad\qquad \cdot \inf_{u\in[t_i,t_{i+1})} e^{\alpha(u)}\sqrt{\zeta(u)}\Bigg\}.
\end{aligned}$$

So, we can apply the concentration result for stochastic integrals with deterministic integrands from [BG06, Lemma B.1.3 (Bernstein-type inequality)]. That leads to

$$\begin{aligned}
&\mathbb{P}\Big(E_{t_i} \cap E^c_{t_{i+1}}\Big) \\
&\leq \exp\left(- \frac{\left(\left(h - \sup_{s\in[t_i,t_{i+1})} \frac{1}{\sqrt{\zeta_s}}\Big(F(s) + G(s)h + H(s)h^2\Big)\right) \inf_{u\in[t_i,t_{i+1})} e^{\alpha(u)}\sqrt{\zeta(u)}\right)^2}{2\sigma^2 \int_0^{t_{i+1}} e^{2\alpha(u)}du} \right) \\
&\leq \exp\left(- \frac{\left(\left(h - \sup_{s\in[t_i,t_{i+1})} \frac{1}{\sqrt{\zeta(s)}}\Big(F(s) + G(s)h + H(s)h^2\Big)\right) \inf_{u\in[t_i,t_{i+1})} \sqrt{\zeta(u)}\right)^2}{2\sigma^2 \int_0^{t_{i+1}} e^{-2\alpha(t_{i+1},u)}du \sup_{u\in[t_i,t_{i+1}]} e^{2\alpha(t_{i+1},u)}} \right) \\
&\leq \exp\left(- \frac{\left(\left(h - \sup_{s\in[t_i,t_{i+1})} \frac{1}{\sqrt{\zeta(s)}}\Big(F(s) + G(s)h + H(s)h^2\Big)\right) \sqrt{\zeta(t_i)}\right)^2}{2\sigma^2 \int_0^{t_{i+1}} e^{-2\alpha(t_{i+1},u)}du \; e^{2\alpha(t_{i+1},t_i)}} \right) \\
&= \exp\left(-\frac{1}{2\sigma^2 V(t_{i+1})}\left(h - \sup_{s\in[t_i,t_{i+1}]} \zeta(s)^{-\frac{1}{2}}\big(F(s) + G(s)h + H(s)h^2\big)\right)^2 \frac{\zeta(t_i)}{e^{2\alpha(t_{i+1},t_i)}}\right).
\end{aligned}$$

We introduce the auxiliary notation

$$q_i := \frac{1}{2\sigma^2 V(t_{i+1})}\left(h - \sup_{s\in[t_i,t_{i+1}]} \zeta(s)^{-\frac{1}{2}}\big(F(s)+G(s)h+H(s)h^2\big)\right)^2 \frac{\zeta(t_i)}{e^{2\alpha(t_{i+1},t_i)}} \qquad \text{for all } i\in I.$$

And first, for all $i \in I$, we obtain for the terms that are not contained in the outer squared parantheses

$$\begin{aligned} q_i^{(1)} := \frac{1}{2\sigma^2 V(t_{i+1})}\frac{\zeta(t_i)}{e^{2\alpha(t_{i+1},t_i)}} &\geq \frac{1}{2\sigma^2}\frac{a(t_{i+1})}{a(t_i)-|b(t_i)|}\frac{1-\delta_1\nu}{1+\delta_1\nu}(1-2a_+t_\Delta) \\ &= \frac{1}{2\sigma^2}\frac{a(t_{i+1})}{a(t_i)-|b(t_i)|}\Big(1+\mathcal{O}(\nu+t_\Delta)\Big), \end{aligned} \tag{5.2.51}$$

where we use first that $\zeta(t_i) \geq \frac{1}{2(a(t_i)-|b(t_i)|)}(1-\delta_1\nu)$ from (5.2.32), (5.2.34), second that $V(s) \leq \frac{1+\delta_1\nu}{2a(s)}$ from (5.2.42), and third the estimate (5.2.45). And for the terms that appear inside the squared parantheses of q_i applying the estimates of Lemma 5.23, we achieve for all $i \in I$ that

$$\begin{aligned} &q_i^{(2)} := \left(h - \sup_{s\in[t_i,t_{i+1}]}\frac{F(s)+G(s)h+H(s)h^2}{\sqrt{\zeta(s)}}\right)^2 \qquad (5.2.52) \\ &\geq \left(h - \sup_{s\in[t_i,t_{i+1})}\left\{\frac{|b(s)|}{a(s)}h + \frac{1}{\sqrt{\zeta(s)}}\left(\|y_0\|\vee\frac{\nu c_0}{2a_-}\right) + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}\right\}\right)^2 \\ &= h^2\left(1 - \sup_{s\in[t_i,t_{i+1})}\left\{\frac{|b(s)|}{a(s)} + \frac{1}{h}\left(\frac{1}{\sqrt{\zeta(s)}}\left(\|y_0\|\vee\frac{\nu c_0}{2a_-}\right) + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}\right)\right\}\right)^2 \\ &= h^2 \inf_{s\in[t_i,t_{i+1})}\left(1-\frac{|b(s)|}{a(s)}\right)^2\left(1 - \frac{\frac{1}{\sqrt{\zeta(s)}}\left(\|y_0\|\vee\frac{\nu c_0}{2a_-}\right) + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}}{h\left(1-\frac{|b(s)|}{a(s)}\right)}\right)^2. \end{aligned}$$

Using that $1-\sup_{s\in[t_i,t_{i+1})}|b(s)|/a(s) \geq 1-\gamma$ on the one hand, and that

$$1 - \sup_{s\in[t_i,t_{i+1})}\frac{|b(s)|}{a(s)} = \inf_{s\in[t_i,t_{i+1})}\frac{a(s)-|b(s)|}{a(s)}$$

on the other hand, yields

$$\left(1 - \sup_{s\in[t_i,t_{i+1})}\frac{|b(s)|}{a(s)}\right)^2 \geq (1-\gamma)\inf_{s\in[t_i,t_{i+1})}\frac{a(s)-|b(s)|}{a(s)} \qquad \text{for all } i\in I.$$

And also we apply $1-\sup_{s\in[t_i,t_{i+1})}|b(s)|/a(s) \geq 1-\gamma$ in the denominator in the parantheses on the right-hand side to find that

$$q_i^{(2)} \geq h^2(1-\gamma)\inf_{s\in[t_i,t_{i+1})}\frac{a(s)-|b(s)|}{a(s)}\left(1 - \frac{\frac{\|y_0\|\vee\frac{\nu c_0}{2a_-}}{\sqrt{\zeta(s)}} + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}}{h(1-\gamma)}\right)^2 \qquad \text{for all } i\in I.$$

Recombining $q_i^{(1)}$ and $q_i^{(2)}$ allows for every $i \in I$ to write

$$q_i \geq \frac{h^2(1-\gamma)}{2\sigma^2} \inf_{s\in[t_i,t_{i+1})} \frac{a(t_{i+1})}{a(s)} \frac{a(s)-|b(s)|}{a(t_i)-|b(t_i)|} \Big(1+\mathcal{O}(\nu+t_\Delta)\Big) \\ \cdot \left(1 - \frac{\frac{\|y_0\|\vee\frac{\nu c_0}{2a_-}}{\sqrt{\zeta(s)}} + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}}{h(1-\gamma)}\right)^2.$$

Then, the auxiliaries from Lemma 5.24 are applicable and both additional factors, that can be derived from the infimum, each of them $1+\mathcal{O}(\nu t_\Delta)$, get absorbed in the Laudau symbol. The former infimum may then be taken to act on the minuend in the parantheses, on which it becomes a supremum. We obtain that

$$q_i \geq \frac{h^2(1-\gamma)}{2\sigma^2}\Big(1+\mathcal{O}(\nu+t_\Delta)\Big)\left(1 - \sup_{s\in[t_i,t_{i+1})} \frac{\frac{\|y_0\|\vee\frac{\nu c_0}{2a_-}}{\sqrt{\zeta(s)}} + N\frac{\sqrt{\zeta(s)}}{a(s)}h^2 + \nu\delta_1\frac{h+h^2}{\sqrt{\zeta(s)}}}{h(1-\gamma)}\right)^2 \\ \text{for all } i \in I.$$

Then, due to Lemma 5.2.6, $\zeta(\cdot)$ and ζ^{-1} are bounded above by something of order 1. Therefore, we have that

$$\begin{aligned} q_i &\geq \frac{h^2(1-\gamma)}{2\sigma^2}\Big(1+\mathcal{O}(\nu+t_\Delta)\Big)\left(1+\mathcal{O}\left(\frac{\|y_0\|\vee\nu}{h} + Nh + \frac{\nu}{h}\right)\right)^2 \\ &= \frac{h^2(1-\gamma)}{2\sigma^2}\left(1+\mathcal{O}\left(\frac{\|y_0\|\vee\nu}{h} + Nh + t_\Delta + \nu\right)\right) \qquad \text{for all } i \in I. \end{aligned}$$

Applying Lemma 5.24 yields

$$\mathbb{P}\big(E_{t_i}\cap E^c_{t_{i+1},}\big) \leq \exp\left(-(1-\gamma)\frac{h^2}{2\sigma^2}\left(1+\mathcal{O}\left(\frac{\|y_0\|\vee\nu}{h} + Nh + t_\Delta\right)\right)\right) \\ \text{for all } i \in \{1,\dots,n\}.$$

And so, consequently, we have shown that

$$\mathbb{P}\big(E_{T_0/\nu}\big) \leq \frac{T_0}{\nu t_\Delta}\exp\left(-(1-\gamma)\frac{h^2}{2\sigma^2}\left(1+\mathcal{O}\left(\frac{\|y_0\|\vee\nu}{h} + Nh + t_\Delta\right)\right)\right).$$

□

So far, the result features several degrees of freedom. And of course, there is no ultimately convenient way to deminish generality in order to enhance clearity. The below remark suggest a relatively concrete instance of relations between parameters.

Remark 5.26. *a) Remember that assumption* (5.2.35) *demands* $\max\{\sigma, \|y_0\|, \nu\} < h < 1$. *Further, the usefulness of the result of the theorem depends on small terms in the Landau symbol. One way to achieve that is the following: For arbitrary* $\alpha, \beta \in (0,1)$, *we let* $\nu = t_\Delta = \sigma^\alpha$ *implying that* $\|y_0\| = \mathcal{O}(\sigma^\alpha)$. *Further, we let* $h = \tilde{h}\sigma^{\alpha\beta}$, *where* $\tilde{h}$ *denotes some neither small nor big constant. Then, the result of Theorem 5.25 reads*

$$\mathbb{P}(E_{T_0/\nu}) \leq \exp\left(-(1-\gamma)\frac{\tilde{h}^2}{2\sigma^{2(1-\alpha\beta)}}\left(1+\mathcal{O}\Big(\sigma^{\alpha(1-\beta)} + N\sigma^{\alpha\beta}\Big)\right) + 2\alpha|\log\sigma| + \log T_0\right).$$

b) *Reaching for the smallest possible* h *for which Theorem 5.25 provides useful results, we observe the most unpleasant term* $\frac{\|y_0\| \vee \nu}{h}$ *allows for a careful choice of* h *of order* ν *such that the term* $\mathscr{R}(\nu, \|y_0\|, h, t_\Delta) = \mathcal{O}\big(\frac{\|y_0\| \vee \nu}{h} + Nh + t_\Delta\big)$ *remains strictly smaller than* $1/2$ *for instance. Then,* σ *of order* $\frac{\nu}{|\log \nu|}$ *is sufficient to have the overall early-escape probability small.*

5.3. Transition - Stability Comes to an End

During the phase when the stability matrix $A^\star(t)$ is not any longer negative and uniformly bounded away from 0, the *slaving principle* in the form of Remark 5.1 does no longer apply. That means that all former ν-adiabatic solution paths might possibly leave the environment of order ν along the equilibrium branch. This phenomenon occurs e.g. for the asymmetric pitchfork bifurcation. We have deliberately excluded such behavior from our study by assuming that there actually is a ν-adiabatic solution path $(x^{ad_\nu}(t))_{t\in[0,T]}$ close to the equilibrium branch $x^\star$.
This subsection will successively study the transition phase from a uniformly stable environment to instability. In the first part 5.3.1 of this section, under suitable conditions on the relation between the adiabatic solution and the equilibrium branch, we will show that the consecutive-boxes approach presented in Subsection 5.2 provides useful results for the case when stability gets small, but is not lost. Formally, we will consider the system up to a time T_1/ν, defined through

$$T_1/\nu := \inf\left\{ t \geq 0 : \frac{|b(t)|}{a(t)} = 1 - \sqrt{\nu}|\log \nu| \right\}.$$

As we argued in Remark 5.3, the definition of T_1/ν depends on x^{ad_ν}, which is not unique in general, but we neglect the inaccuracy and continue to speak of the time points $T_0, T_1, \ldots$ as well as of the ν-adiabatic solution as if they were unique. Subsection 5.3.2 shows that, as long as the bifurcation point is reached in a reasonable time, the perturbed solution remains relatively close to the adiabatic solution. The following Subsection 5.3.3 provides that a typical solution will remain close to the adiabatic solution even for a short period after the point of instability is surpassed. In Subsection 5.3.4 we study the effect of uniformly symmetric potentials, which provides significant improvement concerning the description of solutions between T_0/ν and a fast time of order $\sqrt{\nu}|\log \nu|$ later than the actual transition point T_2/ν. Finally, Section 5.4 treats the question how long a typical solution path needs to depart from the adiabatic solution.

5.3.1. Shallow Curvature

This first step of the transition-phase description applies the techniques of Theorem 5.18 b) and provides the following insight: The deviation $y = x - x^{\mathrm{ad}_\nu}$ typically remains within a distance at most of order $\sqrt{\nu}$ from the equilibrium branch as long as $\frac{|b(s)|}{a(s)} \leq 1 - \sqrt{\nu}|\log \nu|$. Accordingly, we redefine the stability indicator γ as $\gamma := 1 - \sqrt{\nu}|\log \nu|$ for this subsection. Here, we profit from our foresight when we formulated Lemma 5.14. In particular, in (5.2.26) we defined

$$\tilde{c} := (\gamma + 2)\sigma\beta + \frac{N}{a_-}\sigma^2\beta^2 + \frac{\nu c_0}{a_-},$$

and concluded that

$$\tilde{R}_{\star_-} = \frac{\tilde{c}}{1-\gamma}\left(1 + \mathcal{O}\left(\frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2}\right)\right), \tag{5.3.1}$$

$$\tilde{R}_{\star_+} = \left(\frac{a_-}{N}(1-\gamma) - 2\sigma\beta\right)\left(1 + \mathcal{O}\left(\frac{N\tilde{c}}{(1-\gamma)^2}\right)\right). \tag{5.3.2}$$

We keep assuming that $\sigma < \frac{\nu}{|\log \nu|}$ as before, and additionally that ν is sufficiently small to ensure $\tilde{R}_{\star_-} < \tilde{R}_{\star_+}$. This assumption ensures that the decay-factor function is actually smaller than 1 at least somewhere including that $\tilde{R}_{\star_-}$ and $\tilde{R}_{\star_+}$ are well-defined, especially real. Formally, it ensures that the corresponding radicant is positive, see proof of Lemma 5.14. Since this is no longer true in general for $\gamma \geq 1 - \sqrt{\nu}$, we may understand the choice $\gamma = 1 - \sqrt{\nu}|\log \nu|$ as an application of the consecutive-boxes approach at its limits. Altogether, Lemma 5.14 and Lemma 5.15 are applicable for $\gamma = 1 - \sqrt{\nu}|\log \nu|$, and provide that:

$$\text{There is } t_0 \in [0, T_1/\nu] : \|y_{t_0}\| \in (\tilde{R}_{\star_-}, \tilde{R}_{\star_+})$$
$$\Rightarrow \begin{cases} |y(t)| \leq \|y_{t_0}\| + \sigma\beta & \text{for all } t \in [t_0, T_1/\nu \wedge \tau_\beta(\xi)], \\ |y(t)| \leq \|y_{t_0}\| & \text{for all } t \in [t_0 + a_-^{-1}\log(2), T_1/\nu \wedge \tau_\beta(\xi)]. \end{cases}$$

If we plug $\gamma = 1 - \sqrt{\nu}|\log \nu|$, $\sigma < \frac{\nu}{|\log \nu|}$ in (5.3.1) we receive that

$$\begin{aligned} \tilde{c} &= (3 - \sqrt{\nu}|\log \nu|)\sigma\beta + \frac{N}{a_-}\sigma^2\beta^2 + \frac{\nu c_0}{a_-} \\ &\leq 3\frac{\nu}{|\log \nu|}\beta + \frac{N}{a_-}\frac{\nu^2}{|\log \nu|^2}\beta^2 + \nu\frac{c_0}{a_-}. \end{aligned}$$

And therefore,

$$\frac{\tilde{c}}{1-\gamma} < 3\frac{\sqrt{\nu}}{|\log \nu|^2}\beta + \frac{N}{a_-}\frac{\nu^{3/2}\beta^2}{|\log \nu|^3} + \frac{\sqrt{\nu}c_0}{a_-|\log \nu|}.$$

Furthermore,

$$\begin{aligned} \frac{\sigma\beta N}{1-\gamma} + \frac{N\tilde{c}}{(1-\gamma)^2} &\leq \frac{\nu\beta}{\sqrt{\nu}|\log \nu|^2} + \frac{1}{\nu|\log \nu|^2}\left(3\frac{\nu}{|\log \nu|}\beta + \frac{N}{a_-}\frac{\nu^2}{|\log \nu|^2}\beta^2 + \nu\frac{c_0}{a_-}\right) \\ &= \frac{\sqrt{\nu}\beta}{|\log \nu|^2} + 3\frac{\beta}{|\log \nu|^3} + \frac{N}{a_-}\frac{\nu}{|\log \nu|^4}\beta^2 + \frac{c_0}{a_-|\log \nu|^2}. \end{aligned}$$

With regard to the definition of $\tilde{R}_{\star_-}$ in (5.3.1), the last estimate justifies the second item in the below assumption. The third item is an assumption on a minimal transition speed is natural.

Assumption 5.27. *We add the following brief assumptions:*

- *Suppose that ν is small enough such that $\tilde{R}_{\star_-} < \tilde{R}_{\star_+}$.*
- *We assume that ν is sufficiently small that $\tilde{R}_{\star_-} \leq \frac{2\tilde{c}}{1-\gamma}$.*
- *We assume that for sufficiently small ν the system needs a time at most of order $\frac{T_1 - T_0}{\nu} \in \mathcal{O}(1/\nu)$ to transform from uniform stability to the shallow-curvature phase that we associate with $\gamma = 1 - \sqrt{\nu}|\log \nu|$.*

In Theorem 5.18 we have noted that, in order to have $\mathbb{P}\{\tau_\beta(\xi) < T_1/\nu\}$ small, it suffices to choose β of order $|\log \nu|$. Therefore, $\tilde{R}_{\star_-} \in \mathcal{O}(\sqrt{\nu})$ in the shallow-curvature phase. It is further natural to assume that $\|y_{T_0/\nu}\| = \mathcal{O}(\nu)$, because Theorem 5.18 provides that the deviation process enters a neighborhood of order ν of the equilibrium branch with high probability. Furthermore, by either Theorem 5.18, Lemma 5.15 b) and (5.2.30) or

Theorem 5.25 and Remark 5.26, we have that solution paths remain with high probability in a neighborhood of size at most of order ν until time T_0/ν if we assume that $\sigma < \frac{\nu}{|\log \nu|}$. Summarizing, due to the groundwork we prepared in Subsection 5.2.1, we conclude the following main result of this section.

Theorem 5.28. *Suppose that Assumptions 5.27 hold true. Assume further that $\sigma < \frac{\nu}{|\log \nu|}$, that $\|y_{T_0/\nu}\| \in \mathcal{O}(\nu)$, and that $\frac{|b(t)|}{a(t)} \leq 1 - \sqrt{\nu}|\log \nu|$ for all $t \in [T_0/\nu, T_1/\nu]$. Assume further that* (5.2.29) *holds true for all $t \in [T_0/\nu, T_1/\nu]$. Then for sufficiently small ν,*

$$\mathbb{P}\left\{\sup_{s\in[T_0/\nu,T_1/\nu]} |y(s)| > \frac{2\tilde{c}}{1-\gamma}\right\} \leq \frac{4e\beta^2 e(T_1 - T_0)(1+\mathcal{O}(\nu))}{\theta_1 \nu} \exp\left(-\beta^2 \tilde{a}_-\right) \qquad \textit{for } \beta > 0,$$

which is useful if β is of order $|\log \nu|$ and therefore $\tilde{c}$ is of order $\sqrt{\nu}$.

Theorem 5.18 a) provides that the deviation process enters a neighborhood of order ν of the equilibrium branch with high probability, and solution paths remain with high probability in a neighborhood of size at most of order ν until time T_0/ν. Therefore, the assumption $\|y_{T_0/\nu}\| \leq \mathcal{O}(\nu)$ is absolutely natural.

The result of Theorem 5.27 will be tremendously improved under the following two additional assumptions in Subsection 5.3.4:

- Equilibrium branch and adiabatic solution are identical zero,
- The potentials F and G are uniformly symmetric over $[T_0/\nu, T_1/\nu]$.
- The noise amplifier fulfills $\sigma < \nu^2/|\log \nu|$.

Under those assumptions, Theorem 5.40 will show that a solution typically remains even in an environment of order ν around 0 up to time T_2/ν.

Remark 5.29. *The Bernstein-based approach from Subsection 5.2.2 can not so easily be extended to this regime. The main issue is that we can no longer reasonably assume that there is an adiabatic solution path in an environment of size ν around the equilibrium branch $\frac{1}{a(t)-|b(t)|}$, $t \in [T_0/\nu, T_1/\nu]$. That especially spoils the correctness of the estimates* (5.2.32) *and the ones from Lemma 5.23. The proof of Theorem 5.25 relies crucially on those.*

5.3.2. The End of Stability

This subsection studies the concentration of solution paths x of (5.1.3) along a ν-adiabatic solution in the transition phase where stability starts weak and continues ebbing away until it is entirely gone. The description in a close environment of the *transition point*, i.e. where stability is totally lost, here T_2/ν, requires a reinforcement of assumptions, or better, a clarification or choice of which kind of transition we want to study. We make the following assumptions.

Assumption 5.30.

a) $x^\star(t) = x^{\mathrm{ad}}(t) = 0$ *for all* $t \in [T_1/\nu, T/\nu]$,

b) Symmetry. The functions $f(\cdot, s), g(\cdot, s)$ *are odd in the spatial argument at* $s = T_2$*; in particular* $f_{xx}(x^\star(T_2/\nu), T_2) = g_{xx}(x^\star(T_2/\nu), T_2) = 0$,

c) The coefficients f *and* g *are supposed to lie in* $\mathcal{C}_b^{3,3}$ *and one of the following properties holds :*

- *c₁) Linearity. Both* $f(\cdot, s)$ *and* $g(\cdot, s)$ *are linear in the first argument. In particular* $f(y, \nu t) = a(t)y$ *and* $g(y, \nu t) = b(t)y$ *for all* $t \in [T_1/\nu, T_2/\nu]$ *and* $(y, \nu t) \in \mathbb{D}$*. Furthermore,* $\frac{d}{dt}\frac{|b(t)|}{a(t)} > 0$ *in* $t = T_2/\nu$.
- *c₂) Strictly positive third derivatives* $f_{xxx}(x^\star(T_2/\nu), T_2) + g_{xxx}(x^\star(T_2/\nu), T_2) > 0$*, and positive mixed derivatives* $f_{xt}(x^\star(T_2/\nu), T_2) + g_{xt}(x^\star(T_2/\nu), T_2) > 0$.

d) The remaining time $\frac{T_2 - T_1}{\nu}$ *is assumed to be of order* $\frac{|\log \nu|}{\sqrt{\nu}}$*. We assume that there is* $\Delta_{1,2} = \mathcal{O}(1)$ *such that* $T_2 - T_1 = \Delta_{1,2}\sqrt{\nu}|\log \nu|$.

Remark 5.31. *There are assumptions that are crucial for the typical behavior of solutions that we describe:*

- *The symmetry assumption on* $F + G$ *is crucial to characterize the delayed symmetric pitchfork bifurcation in case c₁). One ends up with a saddle-node bifurcation for the replacement system if that condition fails. Here, the assumption is strengthened by supposing both* F *and* G *to be symmetric at time* T_2/ν*. In particular, it provides that there are constants* $\tilde{N}_f, \tilde{N}_g, M_f, M_g \geq 0$ *such that*

$$\left.\begin{aligned} \mathcal{R}_f(y, \nu t) &\leq \sqrt{\nu}|\log \nu|\tilde{N}_f y^2 + M_f y^3, \\ \mathcal{R}_g(y, \nu t) &\leq \sqrt{\nu}|\log \nu|\tilde{N}_g y^2 + M_g y^3 \end{aligned}\right\} \text{ for all } t \in [T_1/\nu, T_2/\nu],\ (y, \nu t) \in \mathbb{D}. \quad (5.3.3)$$

  And we denote $\tilde{N} := \tilde{N}_f + \tilde{N}_g$*,* $M := M_f + M_g$*. Note, that the assumption is naturally fulfilled in case c₂).*

And there are rather technical assumptions involved to guarantee that the transition into the time interval $[T_2/\nu, T/\nu]$ *proceeds nicely through the transition point* T_2/ν.

- *Condition c₁) is self-explaining, and allows a broad variety of how a parameter combination leaves the stability area* S*. It excludes for instance a parameter-combination journey along the boundary on the stability area.*
- *Condition c₂) together with a), b) and d) characterizes a delayed pitchfork bifurcation.*
- *Assumption d) is again a minimal-transition-speed assumption.*

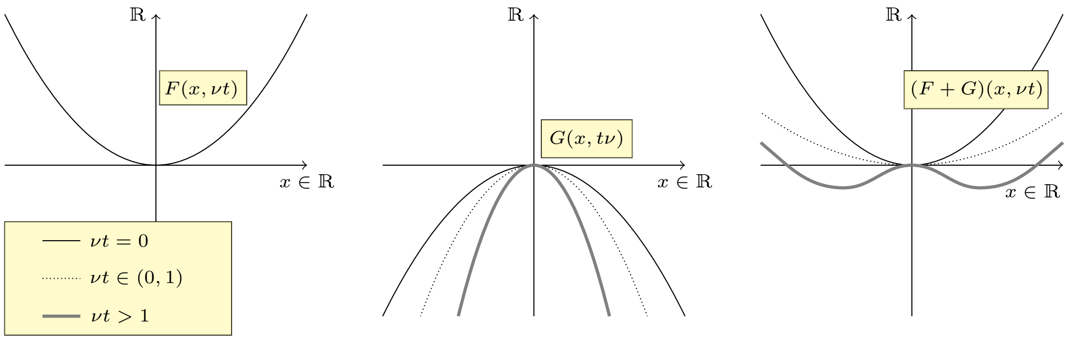

Figure 10: Illustration of Example 5.32 part a). For the replacement system, a symmetric pitchfork bifurcation manifests at $t\nu = 2$. Here, $\varepsilon > 0$ is some positive constant to illustrate the principle shape change of the potential $F + G$.

The rest of the Assumption 5.30 is mostly a comfort of notation

- *As we have seen in* (5.2.12)*, using a* ν*-adiabatic solution as reference provides a correction term at most of order* ν *in the differential law, see* (5.2.13)*,* (5.2.14)*. An additional term of that size can be dealt with similar to the treatment of* $2\sigma h\sqrt{t_\Delta}$ *within the proof of Theorem 5.35. This would lead to a modification of the family of successive upper bounds* $(\beta_i^\star)_{i\in I}$ *at most of order* $|\log\nu|$*. Setting* $x^{\mathrm{ad}_\nu} = x^\star$ *of* ν *obliterates this minor inaccuracy. In other words, the* ν*-adiabatic solution, which is defined with respect to the replacement system, from now on satisfies the delay differential law when neglecting nonlinear terms.*

- *Moreover, the inaccuracy in the definition of time points, see e.g. Remark 5.3, comes to an end.*

Example 5.32. *The following two examples serve as paragons for our study concerning the transition phase.*

a) An instance of a delayed symmetric pitchfork bifurcation that satisfies all requirements of Assumption 5.30 except for c_2*), and also uniformly vanishing quadratic nonlinearity, that we will consider in Subsection 5.3.4, is given by*

$$\left.\begin{aligned} F(x,t\nu) &:= \int_0^x f(y,t\nu)dy = x^4 + 2x^2, \\ G(x,t\nu) &:= \int_0^x g(y,t\nu)dy = -(1+\nu t)x^2 \end{aligned}\right\} \quad \text{for } x \in \mathbb{R}, t \in [0,T].$$

Here, $x^{\mathrm{ad}_\nu}(\cdot) = 0$ *is an adiabatic solution that follows the equilibrium branch* $x^\star(\cdot) = 0$ *over the whole time interval. The equilibrium branch* $x^\star$ *is uniformly stable up to any time* $t \in (0,1)$ *and becomes unstable at* $\nu t = 1$*. See Figure 10 for an illustration.*

b) Another example is given by the linear nonautonomous journey through the stability boundary which we have seen in Figure 4. Consider

$$\begin{cases} dx(t) = -a(t)x(t)dt + b(t)x(t-r)dt + \sigma dW(t) & \text{for all } t \in [0,T/\nu], \\ x_0 = \Upsilon, \end{cases}$$

where $(a(t),b(t))_{[0,T/\nu]}$ *leaves the area* $\tilde{S} := \{x,y \in \mathbb{R} : x < -|y|\}$ *at* $t = T_2/\nu$*. See Figure 11.*

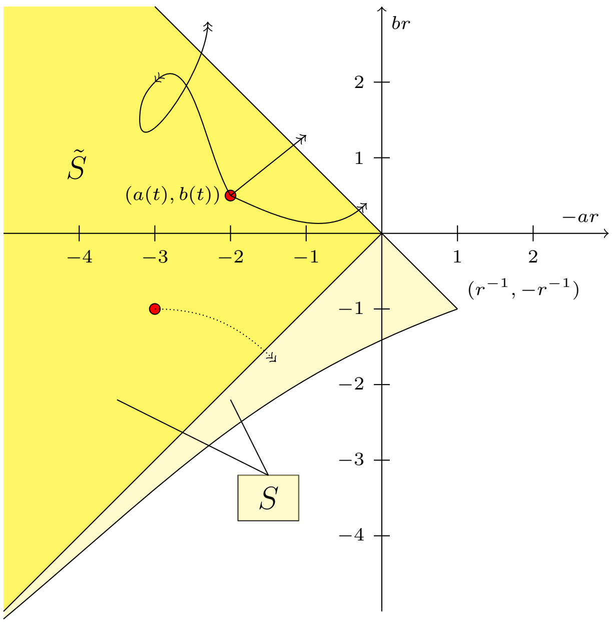

Figure 11: Illustration for Example 5.32 b). The arrow-headed lines represent possible shapes of the coefficient-combination paths $((a(t)r, b(t)r)_{t\in[0,T/\nu]}$ through the boundary of $\tilde{S}$. The pale yellow area and the labels have been taken over from Figure 4 for comparability. The parameter-combination journey, that corresponds to the dotted double-arrow headed line, is covered by our results, while the actual analytic area of stability S is not left in the case of the dotted line, when the combination escapes from $\tilde{S}$.

We introduce a time-dependent noise amplifier

$$F : [T_1/\nu, T_2/\nu] \to [0,1] \text{ adapted and bounded by } 1.$$

The reason for this slight extension is that it simplifies the work on the upcoming time interval, when the system turns slowly increasingly unstable. The Gaussian nature of linearizations stays untouched by this. Further, as a notational update, for this subsection we will denote

$$a_+ := \sup_{t\in[T_1/\nu, T_2/\nu]} a(t) \geq \sup_{t\in[T_1/\nu, T_2/\nu]} b(t) =: b_+.$$

Starting from T_1/ν, we keep denoting the deviation process as y, given as the unique solution of

$$\begin{cases} dy(t) = \big(-a(t)y(t) + b(t)y(t-r) + \mathcal{R}_f(y(t), \nu t) + \mathcal{R}_g(y(t-r), \nu t)\big)dt + \sigma F(t)dW(t) \\ \hfill \text{for } t \geq T_1/\nu, \\ y_{T_1/\nu} = \Upsilon, \end{cases}$$

for some appropriate $\Upsilon \in \mathcal{C}(J, \mathbb{R})$. And we keep thinking of y as the deviation process of a solution of (5.1.3) from the ν-adiabatic solution $x^{\mathrm{ad}_\nu} = 0$. Due to the previous subsection we conveniently assume that $\|\Upsilon\| \in \mathcal{O}(\sqrt{\nu})$. For some $n \in \mathbb{N}$, let $(\theta_i : i \in \{0, 1, \ldots, n-1\})$ with $T_1/\nu = \theta_0 < \theta_1 < \ldots < \theta_n = T_2/\nu$ denote the equidistant partition of $[T_1/\nu, T_2/\nu]$ into n pieces with step width $t_\Delta = \frac{T_2 - T_1}{n\nu}$, where n is chosen big enough such that $t_\Delta < r$ at least. Informally speaking, in time T_1/ν, there are $\frac{T_2-T_1}{\sqrt{\nu}}$ time units left before the stability is lost and the system has completed the transition to a possibly critical or unstable regime, i.e., $\frac{|b(s)|}{a(s)} \geq 1$. To be precise at that point, criticality or instability of the linearized system frozen in T_2/ν is only reached in case of positive delayed feedback $a(T_2/\nu) = b(T_2/\nu) > 0$, represented through the pale yellow overhang in Figure 11. But the deduced estimates apply just as well for the case $b(T_2/\nu) = -a(T_2/\nu) < 0$. For a nondecreasing family of constants $0 < \beta_0 \leq \beta_1 \leq \ldots \leq \beta_n$, let

$$\tau_\beta(x) = \inf\left\{ t \in [T_1/\nu, T_2/\nu] : |y(t)| > \sum_{i=0}^{n-1} \beta_i \mathbb{1}_{[\theta_i,\theta_{i+1}]}(t) \right\}. \tag{5.3.4}$$

We continue to denote $I := \{0, \ldots, n-1\}$. For an appropriate choice of $(\beta_i)_{i\in I}$ we are going to show that

$$|y(t)| \leq \beta_i \quad \text{for all } t \in [\theta_i, \theta_{i+1}] \text{ and } i \in I \text{ with high probability.}$$

For all $i \in I$ we define the linear nonautonomous piecewise approximation $(Y^{(i)}(t))_{t\in[\theta_i - r,\theta_{i+1}]}$ as the unique solutions of

$$\begin{cases} dY^{(i)}(t) = -a(t)Y^{(i)}(t) + b(t)Y^{(i)}(t-r)dt + \sigma F(t)dW(t) & \text{for } t \in [\theta_i, \theta_{i+1}], \\ Y^{(i)}_{\theta_i} = y_{\theta_i}, \end{cases} \tag{5.3.5}$$

each of which admits a representation through the variation-of-constants formula, see Theorem 3.5. To that end, let $(\check{Y}(t,u) : u \in [T_1/\nu - r, T_2/\nu],\ t \in [u-r, T_2/\nu])$ denote the fundamental solution corresponding to the differential law of (5.3.5) initiated at some u, and evaluated at t. Within a regime with $a(\cdot) > |b(\cdot)|$, through a simple contradiction argument as in Lemma 4.1, we may deduce that

$$|\check{Y}(t,u)| \leq 1 \quad \text{for all } u \in [T_1/\nu - r, T_2/\nu],\ t \in [u-r, T_2/\nu]. \tag{5.3.6}$$

Let further denote $(T^{\text{det}}(t,u) : u \in [T_1/\nu, T_2/\nu], u - r \leq t)$ denote the solution semi group, that maps from J to $\mathbb{R}$, of the deterministic counterpart of system (5.3.5), initiated at u and evaluated in t. Then, we may rewrite the approximation for each $i \in I$ as

$$Y^{(i)}(t) = T^{\text{det}}(t,\theta_i)y_{\theta_i} + \sigma\xi^{(i)}(t) \quad \text{with} \quad \xi^{(i)}(t) := \int_{\theta_i}^t \check{Y}(t,u)F(u)dW(u) \quad \text{for all } t \in [\theta_i, \theta_{i+1}]. \tag{5.3.7}$$

The first summand is the solution of the deterministic version of (5.3.5) and as long as $a(\cdot) > |b(\cdot)|$, analogue to (5.3.6), it is easy to check that

$$|T^{\text{det}}(t,\theta_i)y_{\theta_i}| \leq \|y_{\theta_i}\| = \sup_{u\in[-r,0]} |y(\theta_i + u)| \leq \sup_{u\in[T_1/\nu - r,\theta_i]} |y(u)| \quad \text{for all } t \in [\theta_i, \theta_{i+1}].$$

The second summand $(\xi^{(i)}(t))_{t\in[\theta_i,\theta_{i+1}]}$ solves

$$\begin{cases} d\xi^{(i)}(t) = \big(-a(t)\xi^{(i)}(t) + b(t)\xi^{(i)}(t-r)\big)dt + \sigma F(t)dW(t) & \text{for } t \in [\theta_i, \theta_{i+1}], \\ \xi^{(i)}_{\theta_i} = 0, \end{cases}$$

and actually forms a Gaussian process for every $i \in I$. Then, an application of the Fernique inequality provides the following lemma.

Lemma 5.33. *For*

$$h > \sqrt{1+4\log p} \quad \textit{where} \quad p \in \mathbb{N} \quad \textit{with} \quad \sqrt{p}\log p \geq 4(a_+ t_\Delta + 1), \tag{5.3.8}$$

we have that

$$\mathbb{P}\left\{\sup_{t\in[\theta_i,\theta_{i+1}]} |\xi^{(i)}(t)| > h2\sqrt{t_\Delta}\right\} \leq \frac{5p^2}{2}e^{-\frac{h^2}{2}} \qquad \textit{for all } i \in I.$$

Proof. We fix $i \in I$, write $Y^{(i)} = Y$, and focus on $(\check{Y}(t,u) : u \in [\theta_i, \theta_{i+1}], t \geq u - r)$. The Fernique parameters $Q = Q_{\xi^{(i)}}$ and $\Gamma = \Gamma_{\xi^{(i)}}$ are easily obtained. We find that

$$\begin{aligned}\|\Gamma\| = \sup_{s\in[0,t_\Delta]} \mathbb{E}\left[\left(\xi^{(i)}(\theta_i+s)\right)^2\right] &= \sup_{s\in[0,t_\Delta]} \int_0^s \check{Y}^2(\theta_i+s,\theta_i+u)F^2(\theta_i+u)du \\ &\leq \int_0^{t_\Delta} F^2(\theta_i+u)du \leq t_\Delta.\end{aligned}$$

Further, rewriting $\overline{s} = s - \theta_i, \overline{t} = t - \theta_i$ and formally substituting $v = u - \theta_i$, it is easy to see that

$$\begin{aligned}&\mathbb{E}\left[\left(\int_{\theta_i}^t \check{Y}(t,u)F(u)dW(u) - \int_{\theta_i}^s \check{Y}(s,u)F(u)dW(u)\right)^2\right] \\ &\quad = \int_{\theta_i}^s \left(\check{Y}(t,u) - \check{Y}(s,u)\right)^2 F^2(u)du + \int_s^t \check{Y}^2(t,u)F^2(u)du \\ &\quad \leq \int_0^{\overline{s}} \left(\check{Y}(\theta_i+\overline{t},\theta_i+v) - \check{Y}(\theta_i+\overline{s},\theta_i+v)\right)^2 dv + \int_{\overline{s}}^{\overline{t}} \check{Y}^2(\theta_i+\overline{t},\theta_i+v)dv \\ &\qquad\qquad \text{for all } s,t, \in [\theta_i,\theta_{i+1}], \text{ i.e., } \overline{s},\overline{t} \in [0,t_\Delta].\end{aligned}$$

Then, with (5.3.6) we find that

$$\begin{aligned}&\int_0^{\overline{s}} \left(\check{Y}(\theta_i+\overline{t},\theta_i+v) - \check{Y}(\theta_i+\overline{s},\theta_i+v)\right)^2 dv \\ &\quad = \int_0^{\overline{s}} \left(\int_0^{\overline{t}-\overline{s}} -a(\theta_i+\overline{s}+u)\check{Y}(\theta_i+\overline{s}+u,\theta_i+v)\right. \\ &\qquad\qquad \left. + b(\theta_i+\overline{s}+u)\check{Y}(\theta_i+\overline{s}+u-r,\theta_i+v)du\right)^2 dv \\ &\quad \leq (a_+ + b_+)^2 t_\Delta(\overline{t}-\overline{s})^2 \leq 4a_+^2 t_\Delta(\overline{t}-\overline{s})^2 \quad \text{for all } \overline{s},\overline{t} \in [0,t_\Delta],\ \overline{s} \leq \overline{t}.\end{aligned}$$

And it is easy to see that

$$\int_{\overline{s}}^{\overline{t}} \check{Y}^2(\theta_i+\overline{t},\theta_i+v)dv \leq \overline{t} - \overline{s} \qquad \text{for all } \overline{s},\overline{t} \in [0,t_\Delta],\ \overline{s} \leq \overline{t}.$$

Therefore, with a glimpse at the previously used notation in (3.4.10) and (3.4.11), and the formulas from the Appendix A.2, we obtain that

$$\mathcal{Q}_1 \leq 2a_+\sqrt{t_\Delta}\int_1^\infty t_\Delta p^{-u^2}du \leq \frac{2a_+ t_\Delta^{3/2}}{2p\log p} \quad \text{and} \quad \mathcal{Q}_2 \leq \int_1^\infty \sqrt{t_\Delta}\,p^{-\frac{u^2}{2}}du \leq \frac{\sqrt{t_\Delta}}{\sqrt{p}\log p}.$$

Finally, we may deduce from the condition on the minimal size of p, which we stated in (5.3.8), that

$$\begin{aligned}\sqrt{\|\Gamma\|}+Q(t_\Delta) &\le \sqrt{t_\Delta}+\left(2+\sqrt{2}\right)\left(\frac{\sqrt{t_\Delta}}{\sqrt{p}\log p}+\frac{a_+t_\Delta^{\frac{3}{2}}}{p\log p}\right)\\ &=\sqrt{t_\Delta}\left(1+(2+\sqrt{2})\left(\frac{1}{\sqrt{p}\log p}+\frac{a_+t_\Delta}{p\log p}\right)\right)\le 2\sqrt{t_\Delta}.\end{aligned}$$

And the claim follows through an application of the Fernique inequality. □

Let us denote

$$\begin{aligned}\tau^{(i)}_{2h\sqrt{t_\Delta}} &:= \inf\left\{t\in[\theta_i,\theta_{i+1}] : |\xi^{(i)}(t)|>2h\sqrt{t_\Delta}\right\} && \text{for all } i\in I,\ h>0,\\ \tau_{2h\sqrt{t_\Delta}}(\xi) &:= \min\left\{\tau^{(i)}_{2h\sqrt{t_\Delta}} : i\in I\right\} && \text{for all } h>0. \qquad (5.3.9)\end{aligned}$$

Then obviously, for h and p satisfying (5.3.8), we have that

$$\mathbb{P}\left\{\tau_{2h\sqrt{t_\Delta}}(\xi)<T_2/\nu\right\}\le n\frac{5p^2}{2}\exp\left(-\frac{h^2}{2}\right), \tag{5.3.10}$$

which makes the below Corollary is just as obvious. It states the result if we plug in the previous corollary into representation (5.3.7).

Corollary 5.34. *For $h>0$ and $p\in\mathbb{N}$ satisfying* (5.3.8)*, we have that*

$$|Y^{(i)}(t)|\le \sup_{u\in[T_1/\nu-r,\theta_i]}|y(u)|+h2\sigma\sqrt{t_\Delta}$$
$$\textit{for all } t\in[\theta_i,\theta_{i+1}],\ t\le\tau_{2h\sqrt{t_\Delta}}(\xi),\ i\in I.$$

And $\mathbb{P}\{\tau_{2h\sqrt{t_\Delta}}(\xi)<T/\nu\}\le 3np^2\exp(-h^2/2)$.

Further, for all $i\in I$ we define

$$Z^{(i)}(t) := y(t)-Y^{(i)}(t) \quad \text{for all } t\in[\theta_i,\theta_{i+1}]. \tag{5.3.11}$$

which $\mathbb{P}$-almost surely solves

$$\begin{cases}\frac{dZ^{(i)}(t)}{dt}=-a(t)Z^{(i)}(t)+b(t)Z^{(i)}(t-r)+\mathcal{R}_f(y(t),\nu t)+\mathcal{R}_g(y(t-r),\nu t) \text{ for } t\in[\theta_i,\theta_{i+1}),\\ Z^{(i)}_{\theta_i}=0.\end{cases}$$

By Theorem 3.5 the pieces $(Z^{(i)}(t))_{t\in[\theta_i,\theta_{i+1}]}$ also admit respective representations through the variation-of-constants formula in terms of the previously defined fundamental solution $\check{Z}(t,u)=\check{Y}(t,u)$, $u\in[\theta_i,T/\nu], t\in[u-r,T/\nu]$, characterized through its differential law (5.3.5). Namely, $(Z^{(i)}(t))_{t\in[\theta_i,\theta_{i+1}]}$ may be written as

$$Z^{(i)}(t)=\int_{\theta_i}^{t}\check{Z}(t,u)\Big(\mathcal{R}_f(y(u),\nu u)+\mathcal{R}_g(y(u-r),\nu u)\Big)du \quad \text{for all } u\in[\theta_i,\theta_{i+1}),\ i\in I.$$

Here, the fact that f and g are supposed to be odd functions at T_2/ν comes into play and serves through (5.3.3) that

$$
\begin{aligned}
Z^{(i)}(t) &< \int_{\theta_i}^{t} \sqrt{\nu}|\log\nu|\tilde{N}_f y^2(u) + \sqrt{\nu}|\log\nu|\tilde{N}_g y^2(u-r) + M_f y^3(u) + M_g y^3(u-r)du \\
&\leq t_\Delta\sqrt{\nu}|\log\nu|\left(\sup_{u\in[T_1/\nu-r,\theta_{i+1}]}|y(u)|\right)^2 + t_\Delta M\left(\sup_{u\in[T_1/\nu-r,\theta_{i+1}]}|y(u)|\right)^3 \\
&\qquad \text{for all } t\in[\theta_i,\theta_{i+1}],\ t\leq\tau_{2h\sqrt{t_\Delta}}(\xi),\ i\in I.
\end{aligned}
$$

And therefore,

$$
\begin{aligned}
|y(t)| < \sup_{u\in[T_1/\nu-r,\theta_i]}|y(u)| + 2h\sigma\sqrt{t_\Delta} + \sqrt{\nu}|\log\nu|\tilde{N}&\left(\sup_{u\in[T_1/\nu-r,\theta_{i+1}]}|y(u)|\right)^2 t_\Delta \\
+M&\left(\sup_{u\in[T_1/\nu-r,\theta_{i+1}]}|y(u)|\right)^3 t_\Delta \\
&\text{for all } t\in[\theta_i,\theta_{i+1}],\ t\leq\tau_{2h\sqrt{t_\Delta}}(\xi),\ i\in I.
\end{aligned}
\tag{5.3.12}
$$

Theorem 5.35. *In the situation of Assumption 5.30 let* $(\theta_i : i\in\{0,\dots,n-1\})$ *denote the equidistant partition of* $[T_1/\nu, T_2/\nu]$ *with step width* $t_\Delta = \frac{\Delta_{1,2}}{n\sqrt{\nu}}|\log\nu| < r$. *Assume that there is* $k>0$ *such that* $\|y_{T_1/\nu}\| < k\sqrt{\nu}$. *Assume further for* $h = \mathcal{O}\left(\log\left(\frac{\Delta_{1,2}}{\sqrt{\nu}}|\log\nu|\right)\right)$ *that* ν *and* $\sigma < \frac{\nu}{|\log\nu|^2}$ *are small enough such that there is* $K>k$, *independent of* ν, *with*

$$
e^{16\nu t_\Delta K(\tilde{N}|\log\nu|+MK)\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|}\left(k + \frac{2\sigma h\sqrt{t_\Delta}}{\sqrt{\nu}}\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|\right) < K. \tag{5.3.13}
$$

Define the family of increasing constants $(\beta_i^\star)_{i\in\{-1,0,1,\dots,n-1\}}$ *inductively through*

$$
\begin{cases}
\beta_{-1}^\star := k\sqrt{\nu}, \\
\beta_i^\star := \beta_{i-1}^\star\big(1+16\nu t_\Delta K(\tilde{N}|\log\nu|+MK)\big) + 2\sigma h\sqrt{t_\Delta} \quad \text{for } i\in I,
\end{cases}
\tag{5.3.14}
$$

and suppose that σ *and* ν *are small enough such that*

$$
4(\tilde{N}|\log\nu|+MK)\sqrt{\nu}t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}) < 1 \quad \text{for all } i\in I, \tag{5.3.15}
$$

$$
K\sqrt{\nu} < \frac{1}{2(\tilde{N}|\log\nu|+MK)\sqrt{\nu}t_\Delta}, \tag{5.3.16}
$$

$$
2h\sigma\sqrt{t_\Delta} \leq k\sqrt{\nu}. \tag{5.3.17}
$$

Then, the following assertions hold true:

a) We have that $\beta_{n-1}^\star \leq K\sqrt{\nu}$.

b) For all $i\in I$ *we have that* $\beta_i^\star \geq \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + \tilde{N}\sqrt{\nu}|\log\nu|(\beta_i^\star)^2 + M(\beta_i^\star)^3$.

c) The family $(\beta_i^\star)_{i\in\{-1,0,1,\dots,n-1\}}$ *constitutes an upper bound for* y, *i.e.* $|y(t)|\leq\beta_i^\star$ *for all* $t\in[\theta_i,\theta_{i+1}]$ *and* $\{-1,\dots,n-1\}$ *as long as* $t<\tau_{2h\sigma\sqrt{t_\Delta}}$ *and (5.3.10) holds true.*

Proof. *a)* The simple recursion formula for $(\beta_i^\star)_{i\in I}$ can be explicitly solved and serves

$$\beta_i^\star = \beta_{-1}^\star \left(1 + 16\nu t_\Delta K(\tilde{N}|\log\nu| + MK)\right)^{i+1} \\ + 2h\sigma\sqrt{t_\Delta}\sum_{j=0}^{i}(1 + 16\nu t_\Delta K(\tilde{N}|\log\nu| + MK))^j \text{ for all } i \in I.$$

Using that $(1+x) \le e^x$, we receive an upper boundary for $\beta_{n-1}^\star$ for $n = \frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|$ through

$$\beta_i^\star \le \beta_{n-1}^\star \le \beta_{-1}^\star e^{16\nu t_\Delta K(\tilde{N}|\log\nu|+MK)\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|} \\ + \frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu| 2h\sigma\sqrt{t_\Delta}\, e^{16\nu t_\Delta K(\tilde{N}|\log\nu|+MK)\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|}.$$

And therefore,

$$\beta_{n-1}^\star < \sqrt{\nu}\, e^{16\nu t_\Delta K(\tilde{N}|\log\nu|+MK)\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|}\left(k + \frac{2\sigma h\sqrt{t_\Delta}}{\sqrt{\nu}}\frac{\Delta_{1,2}}{t_\Delta\sqrt{\nu}}|\log\nu|\right).$$

Note that this was assumed to be less or equal to $K\sqrt{\nu}$ in (5.3.13).

b) In a first step, we show that $\beta_i^\star \ge \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + (\tilde{N}|\log\nu| + MK)\sqrt{\nu}(\beta_i^\star)^2 t_\Delta$. For $i \in I$ rewriting the desired inequality with the notation $R := (\tilde{N}|\log\nu| + MK)$ yields

$$\beta_i^\star \ge \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + R\sqrt{\nu}(\beta_i^\star)^2 t_\Delta \\ \Leftrightarrow (\beta_i^\star)^2 - \frac{1}{R\sqrt{\nu}t_\Delta}\beta_i^\star + \frac{1}{R\sqrt{\nu}t_\Delta}\left(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}\right) \le 0. \tag{5.3.18}$$

Then through inequality (5.3.15) on the size of σ, inequality (5.3.18) is true for $\beta_i^\star \in [\beta_i^{(\star 1)}, \beta_i^{(\star 2)}]$, where

$$\beta_i^{(\star 1)} = \frac{1}{2R\sqrt{\nu}t_\Delta} - \sqrt{\frac{1}{4R^2\nu t_\Delta^2} - \frac{1}{R\sqrt{\nu}t_\Delta}\left(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}\right)} \\ = \frac{1}{2R\sqrt{\nu}t_\Delta} - \frac{1}{2R\sqrt{\nu}t_\Delta}\sqrt{1 - \frac{4R^2\nu t_\Delta^2}{R\sqrt{\nu}t_\Delta}\left(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}\right)},$$

and

$$\beta_i^{(\star 2)} = \frac{1}{2R\sqrt{\nu}t_\Delta} + \sqrt{\frac{1}{4R^2\nu t_\Delta^2} - \frac{1}{R\sqrt{\nu}t_\Delta}\left(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}\right)} = \mathcal{O}\left(\frac{1}{\sqrt{\nu}t_\Delta}\right).$$

We use the fact that $\sqrt{1-x} > 1 - \frac{x}{2} - \frac{x^2}{2}$ for $x \in (0,1)$ and obtain that

$$\beta_i^{(\star 1)} < \frac{1}{2R\sqrt{\nu}t_\Delta} - \frac{1}{2R\sqrt{\nu}t_\Delta}\Big(1 - 2R\sqrt{\nu}t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}) \\ - 8R^2\nu t_\Delta^2(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta})^2\Big) \\ = \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + 4R\sqrt{\nu}t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta})^2.$$

Then, through (5.3.17) we know that σ is small enough such that

$$2h\sigma\sqrt{t_\Delta} \leq \beta^\star_{-1} \leq \beta^\star_{i-1} \quad \text{for all } i \in I.$$

Hence, the squared-parentheses term satisfies $(\beta^\star_{i-1} + 2\sigma h\sqrt{t_\Delta})^2 \leq 4(\beta^\star_{i-1})^2 \leq 4K\sqrt{\nu}\beta^\star_{i-1}$. Then, we receive that

$$\begin{aligned}\beta_i^{(\star 1)} &< \beta^\star_{i-1} + 2h\sigma\sqrt{t_\Delta} + 16KR\nu t_\Delta \beta^\star_{i-1} \\ &= \beta^\star_{i-1}(1 + 16K(\tilde{N}|\log\nu| + MK)\nu t_\Delta) + 2h\sigma\sqrt{t_\Delta} = \beta^\star_i.\end{aligned}$$

And by assumption (5.3.16), we have that

$$\beta^\star_i \leq \beta^\star_{n-1} \leq K\sqrt{\nu} < \frac{1}{2R\sqrt{\nu}t_\Delta} \leq \beta_i^{(\star 2)} \quad \text{for all } i \in I.$$

Therefore, $\beta^\star_i$ satisfies the desired inequality (5.3.18).
c) Let $\tau_{\beta^\star}(y)$ be defined as in (5.3.4). From (5.3.12) we know that

$$\begin{aligned}|y(t)| < \beta^\star_{i-1} + 2h\sigma\sqrt{t_\Delta} + \tilde{N}|\log\nu|\sqrt{\nu}(\beta^\star_i)^2 t_\Delta + M\left(\beta^\star_i\right)^3 t_\Delta \\ \text{for all } t < \tau_{\beta^\star_i}(y), t \in [\theta_i, \theta_{i+1}], t \leq \tau_{2h\sqrt{t_\Delta}}(\xi), i \in I,\end{aligned}$$

while

$$\begin{aligned}\beta^\star_{i-1} + 2h\sigma\sqrt{t_\Delta} + \tilde{N}|\log\nu|\sqrt{\nu}(\beta^\star_i)^2 + M(\beta^\star_i)^3 t_\Delta \leq \beta^\star_i \\ \text{for } t \in [\theta_i, \theta_{i+1}], t \leq \tau_{2h\sqrt{t_\Delta}}(\xi), i \in I,\end{aligned}$$

which actually shows that $\tau_{\beta^\star}(y) > \tau_{2h\sqrt{t_\Delta}}(\xi)$ providing that

$$|y(t)| \leq \beta^\star_i \text{ for all } t \in [\theta_i, \theta_{i+1}], t \leq \tau_{2h\sqrt{t_\Delta}}(\xi), i \in I.$$

□

The three assumptions (5.3.15), (5.3.16), (5.3.17) do not raise any further issue. The first one is in principle justified through a) since ν is considered to be small. Assumption (5.3.16) is also naturally fulfilled for sufficiently small ν, and so is Assumption (5.3.17), because we assumed that $\sigma < \nu$.

Remark 5.36. *In particular, $|y(t)|$ will not exceed a size of order $\sqrt{\nu}$ before stability is ultimately lost (with high probability) if $\sigma \leq \frac{\nu}{|\log\nu|^2}$.*

5.3.3. During a Small Time at Instability

In order to study the pathwise behavior at least for a short time after stability is gone in T_2/ν, we maintain the Assumptions 5.30 from the previous subsection for the time-interval $[T_2/\nu, T_3/\nu]$. In particular, the equilibrium branch $x^\star$ as well as the ν-adiabatic solution are identically zero. And in the variant including the c$_1$)-assumption, there can not be only one equilibrium branch but others will originate from the zero line in the transition point T_2/ν. The assumption that f and g are odd implies that there are at least two of them and they

are spawned symmetric and stable. But we will not pay them any more attention in this work.
We assume that $0 < a(t) < |b(t)|$ for all $t \in (T_2/\nu, T_3/\nu]$, which implies instability for the linearized system frozen in T_2/ν in case $b(\cdot) > 0$. In order to spare the supply of stars, tildes, bars and primes, we reuse the symbols from (5.3.3), and assume that there are constants $\tilde{N}_f, \tilde{N}_g, M_f, M_g$ such that

$$\left.\begin{aligned} \mathcal{R}_f(y,\nu t) &\leq \sqrt{\nu}\tilde{N}_f y^2 + M_f y^3, \\ \mathcal{R}_g(y,\nu t) &\leq \sqrt{\nu}\tilde{N}_g y^2 + M_g y^3 \end{aligned}\right\} \quad \text{for all } t \in [T_2/\nu, T_3/\nu] \text{ and } (y,\nu t) \in \mathbb{D},$$

and keep denoting $M = M_f + M_g$. As we do not have any idea of T_3/ν by now, the notation above is a bit odd in the sense that it includes a hidden assumption, namely that $T_3 - T_2$ is at most of order $\sqrt{\nu}$. This will be justified ex post. Starting from T_2/ν, we consider the solution of

$$\begin{cases} dy(t) = \big(- a(t)y(t) + b(t)y(t-r) + \mathcal{R}_f(y(t),\nu t) + \mathcal{R}_g(y(t-r),\nu t)\big)dt + \sigma F(t) dW(t) \\ \hspace{22em} \text{for } t \in [T_2/\nu, T_3/\nu], \\ y_{T_1/\nu} = \Upsilon, \end{cases} \tag{5.3.19}$$

between $[T_2/\nu, T_3/\nu]$, where T_3/ν is a point in time that is not too far away from T_2/ν. Again Υ is some suitable element of $\mathcal{C}(J,\mathbb{R})$. Actually, the main result of this subsection will show that paths typically remain in an environment of order $\sqrt{\nu}$ as long as $\frac{T_3-T_2}{\sqrt{\nu}}$ is at most of order 1; or in other words, only if $T_3 = T_2 + \mathcal{O}(\sqrt{\nu})$. We conveniently assume that $\|\Upsilon\| = \mathcal{O}(\sqrt{\nu})$. We will mostly reuse the ideas from the previous case, where the system has approached the point of instability. We continue to write $I = \{0,\dots,n-1\}$ for $n := \frac{T_3-T_2}{r\nu}$, $\theta_i = T_2/\nu + ir$, which means that we work with step width $t_\Delta = r$ for notional comfort, and we will construct an appropriate family of consecutive bounds $(\bar{\beta}_i)_{i\in I}$ inducing the stopping time $\tau_{\bar{\beta}}(y)$, defined as in (5.3.4). Let further

$$\begin{aligned} \lambda^+(t) &:= \sup\big\{\lambda \in \mathbb{R} : \text{ There is } u \in [T_2/\nu, t] \text{ with } a(u) + \lambda = b(u)e^{-\lambda r}\big\} \\ &\hspace{14em} \text{for all } t \in [T_2/\nu, T_3/\nu], \\ \lambda_i &:= \lambda^+(\theta_{i+1}) \ \text{ for all } i \in I. \end{aligned}$$

And finally, the conclusions, that we are going to derive, are formulated in the same manner as in the previous subsection. Namely, we will show that $|y(t)| \leq \bar{\beta}_i$ for all $t \in [\theta_i, \theta_{i+1}]$ for all $i \in I$ with high probability for an appropriate choice of $(\bar{\beta}_i)_{i\in\{-1,\dots,n-1\}}$.

For $i \in I$ let $(Y^{(i)}(t))_{t\in[\theta_i - r,\theta_{i+1}]}$ be the unique solutions of the nonautonomous linear SDDEs

$$\begin{cases} dY^{(i)}(t) = -a(t)Y^{(i)}(t)dt + b(t)Y^{(i)}(t-r)dt + \sigma F(t)dW(t) \quad \text{for } t \in [\theta_i, \theta_{i+1}], \\ Y^{(i)}_{\theta_i} = y_{\theta_i}. \end{cases} \tag{5.3.20}$$

Then, we have that $(\mathcal{Y}^{(i)}(t))_{t\in[\theta_i - r,\theta_{i+1}]}$, defined by $\mathcal{Y}^{(i)}(t) := e^{-\lambda_i(t-\theta_i)}Y^{(i)}(t)$ for all $t \in$

$[\theta_i - r, \theta_{i+1}]$, uniquely solves

$$\begin{cases} d\mathcal{Y}^{(i)}_t = -\tilde{a}(t)\mathcal{Y}^{(i)}(t)dt + \tilde{b}(t)\mathcal{Y}^{(i)}(t-r)dt + \sigma\tilde{F}^{(i)}(t)dW(t) & \text{for } t \in [\theta_i, \theta_{i+1}], \\ \mathcal{Y}^{(i)}(t) = e^{-\lambda_i(t-\theta_i)}y(t) & \text{for } t \in [\theta_i - r, \theta_i], \end{cases}$$

where $\tilde{a}(t) = a(t) + \lambda_i \geq b(t)e^{-\lambda_i r} = \tilde{b}(t)$ and $\tilde{F}^{(i)}(t) = e^{-\lambda_i(t-\theta_i)}F(t)$ for all $t \in [\theta_i, \theta_{i+1}]$ for all $i \in I$. The processes $(\mathcal{Y}^{(i)}(t))_{t\in[\theta_i - r, \theta_{i+1}]}$, $i \in I$, satisfy the requirements of the previous subsection, so we may adopt most of the results. Again, we use the notation of a fundamental solution $(\check{\mathcal{Y}}^{(i)}(t,u) : u \in [\theta_i, \theta_{i+1}], u - r \leq t)$ initiated at u, that corresponds to the differential law

$$d\check{\mathcal{Y}}^{(i)}(t) = -\tilde{a}(t)\check{\mathcal{Y}}^{(i)}(t)dt + \tilde{b}(t)\check{\mathcal{Y}}^{(i)}(t-r)dt \quad \text{for all } t \in [\theta_i, \theta_{i+1}], i \in I. \tag{5.3.21}$$

Then we may rewrite the process $(\mathcal{Y}^{(i)}(t))_{t\in[\theta_i,\theta_{i+1}]}$ as

$$\begin{aligned} &\mathcal{Y}^{(i)}(t) = \mathcal{T}^{\det(i)}_{t,\theta_i}\mathcal{Y}^{(i)}_{\theta_i}(0) + \sigma\xi^{(i)}(t), \quad \text{where} \\ &\bar{\xi}^{(i)}(t) := \int_{\theta_i}^{t} \check{\mathcal{Y}}^{(i)}(t,u)e^{-\lambda_i(u-\theta_i)}F(u)dW(u) \quad \text{for all } t \in [\theta_i, \theta_{i+1}],\ i \in I, \end{aligned} \tag{5.3.22}$$

where $(\mathcal{T}^{\det(i)}_{t,u} : u \in [\theta_i, \theta_{i+1}], t \in [u - r, \theta_i])$ denotes the solution semi group, that maps from $\mathcal{C}(J,\mathbb{R})$ to $\mathcal{C}(J,\mathbb{R})$, and corresponds to the differential law (5.3.21) for $i \in I$. Just like in the previous case, we can deduce that

$$\begin{aligned} \sup_{t\in[\theta_i,\theta_{i+1}]} |\mathcal{T}^{\det(i)}_{t,\theta_i}\mathcal{Y}^{(i)}_{\theta_i}(0)| \leq \|\mathcal{Y}^{(i)}_{\theta_i}\| &\leq \sup_{[\theta_i - r,\theta_i]} |\mathcal{Y}^{(i)}(u)| \\ &\leq e^{\lambda_i t_\Delta} \sup_{u\in[T_2/\nu - r,\theta_i]} |y(u)| \quad \text{for all } t \in [\theta_i, \theta_{i+1}]. \end{aligned}$$

The last of the above inequalities is due to the choice $t_\Delta = r$. Without this simplification the above inequality would raise additional intricate issues concerning intersections of intervals $[\theta_j, \theta_{j+1}]$ inside the initial interval $[\theta_i - r, \theta_i]$ of y_{θ_i} for potentially several $j < i$.

In the previous section we already achieved a useful concentration inequality concerning $\bar{\xi}^{(i)}$ given by

$$\mathbb{P}\left\{\sup_{s\in[\theta_i,\theta_{i+1}]} |\bar{\xi}^{(i)}(s)| > 2h\sqrt{t_\Delta}\right\} \leq 3p^2 e^{-\frac{h^2}{2}} \quad \text{for all } i \in I, \tag{5.3.23}$$

where h, p are chosen due to (5.3.8). We denote $\tau_{2h\sqrt{t_\Delta}}(\bar{\xi})$ analogue to $\tau_{2h\sqrt{t_\Delta}}(\xi)$ in (5.3.9). Then Retransforming $Y^{(i)}(t) = e^{\lambda_i(t-\theta_i)}\mathcal{Y}^{(i)}(t)$ for all $t \in [\theta_i - r, \theta_{i+1}]$, $i \in I$, we find that

$$|Y^{(i)}(t)| \leq e^{\lambda_i t_\Delta}\left(e^{\lambda_i t_\Delta} \sup_{u\in[T_2/\nu - r,\theta_i]} |y(u)| + 2h\sigma\sqrt{t_\Delta}\right) \quad \text{for all } t \in [\theta_i, \theta_{i+1}]\ t < \tau_{2h\sqrt{t_\Delta}}(\bar{\xi}).$$

If we fix one interval $[\theta_i, \theta_{i+1}]$, then the process

$$Z^{(i)}(t) := y(t) - Y^{(i)}(t) \quad \text{for all } t \in [\theta_i - r, \theta_{i+1}] \tag{5.3.24}$$

is $\mathbb{P}$-a.s. the unique solution of the ordinary, nonlinear DDE

$$\begin{cases} \frac{Z^{(i)}(t)}{dt} = -a(t)Z(t) + b(t)Z(t-r) + \mathcal{R}_f(y(t),\nu t) + \mathcal{R}_g(y(t-r),\nu t) & \text{for } t \in [\theta_i, \theta_{i+1}], \\ Z^{(i)}_{\theta_i} = 0. \end{cases} \tag{5.3.25}$$

Therefore, the transformed process $\mathcal{Z}^{(i)}(t) := e^{-\lambda_i(t-\theta_i)} Z^{(i)}(t)$ for $t \in [\theta_i - r, \theta_{i+1}]$ solves

$$\begin{cases} \frac{\mathcal{Z}^{(i)}(t)}{dt} = -\tilde{a}(t)\mathcal{Z}^{(i)}(t) + \tilde{b}(t)\mathcal{Z}^{(i)}(t-r) + \tilde{\mathcal{R}}^{(i)}_f(y(t),\nu t) + \tilde{\mathcal{R}}_g(y(t-r),\nu t), & t \in [\theta_i, \theta_{i+1}], \\ \mathcal{Z}^{(i)}_{\theta_i} = 0, \end{cases}$$

where $\tilde{\mathcal{R}}^{(i)}_f(x,\nu t) = e^{-\lambda_i(t-\theta_i)}\mathcal{R}_f(x,\nu t)$ for all tuples $(x,\nu t) \in \mathbb{D}$ with $t \in [\theta_i, \theta_{i+1}]$, and analogue for $\tilde{\mathcal{R}}_g$. The initial segments remain identical 0 which allows us to rewrite $\mathcal{Z}^{(i)}$ as

$$\mathcal{Z}^{(i)}(t) = \int_{\theta_i}^{t} \check{\mathcal{Z}}(t,u)\Big(\tilde{R}^{(i)}_f(y(u),\nu u) + \tilde{R}^{(i)}_g(y(u-r),\nu u)\Big)du \text{ for all } t \in [\theta_i, \theta_{i+1}],$$

and therefore,

$$|\mathcal{Z}^{(i)}(t)| \le t_\Delta \tilde{N}\sqrt{\nu}\left(\sup_{u\in[T_2/\nu - r,t]} |y(u)|\right)^2 + t_\Delta M\left(\sup_{u\in[T_2/\nu - r,t]} |y(u)|\right)^3 \quad \text{for all } t \in [\theta_i, \theta_{i+1}],$$

where $\check{\mathcal{Z}}^{(i)} = \mathcal{Y}^{(i)}$ denotes the corresponding fundamental solution which coincides with one of the previous ones. We may directly deduce

$$\begin{aligned} |y(t)| \le e^{\lambda_i t_\Delta}\Bigg(\sup_{u\in[T_2/\nu - r,\theta_i]} |y(u)|e^{\lambda_i t_\Delta} + 2h\sigma\sqrt{t_\Delta} + t_\Delta\tilde{N}\sqrt{\nu}\left(\sup_{u\in[T_2/\nu - r,t]} |y(u)|\right)^2 \\ + t_\Delta M\left(\sup_{u\in[T_2/\nu - r,t]} |y(u)|\right)^3\Bigg) \\ \text{for all } t \in [\theta_i, \theta_{i+1}],\ t \le \tau_{2h\sqrt{t_\Delta}}(\bar{\xi}). \end{aligned} \tag{5.3.26}$$

The following theorem is an analogue of Theorem 5.35 from the previous section, and the proofs are conceptually identical.

Theorem 5.37. *Consider the situation of Assumption 5.30 and let* $(\theta_i : i \in I)$ *denote the equidistant partition of* $[T_2/\nu, T_3/\nu]$ *with step width* $t_\Delta = r$*, i.e.* $\theta_i = T_2/\nu + ir$ *for* $i \in \{0, \dots, n\}$ *and* $n = \frac{T_3 - T_2}{\nu r}$*. Assume that there is a constant* $K > 0$ *such that* $\|y_{T_2/\nu}\| < K\sqrt{\nu}$*, and that* $h = \mathcal{O}\left(\log\left(\frac{T_3 - T_2}{\nu}\right)\right)$*. Denote* $\lambda_\star^{(T_3)} := \sup_{t \in [T_2/\nu, T_3/\nu]} \lambda^+(t)$*. Suppose that* $\sigma < \frac{\nu}{|\log \nu|^2}$*, and that there is* $\bar{K} > 0$ *for* ν *sufficiently small such that*

$$\begin{aligned} K \exp\left(2t_\Delta \sum_{j=0}^{n-1} \lambda_j\right) \exp\left(\frac{T_3 - T_2}{t_\Delta \nu} 16 t_\Delta \bar{K}(\tilde{N} + \bar{K}M)\nu e^{3\lambda_\star^{(T_3)} t_\Delta}\right) \\ + 2\frac{\sigma}{\sqrt{\nu}} h \sqrt{t_\Delta} e^{\lambda_\star^{(T_3)} t_\Delta} \frac{T_3 - T_2}{t_\Delta \nu} \exp\left(2t_\Delta \sum_{j=0}^{n-1} \lambda_j\right) \\ \cdot \exp\left(\frac{T_3 - T_2}{t_\Delta \nu} 16 t_\Delta \bar{K}(\tilde{N} + \bar{K}M)\nu e^{3\lambda_\star^{(T_3)} t_\Delta}\right) < \bar{K}. \end{aligned} \tag{5.3.27}$$

Define

$$\begin{cases} \bar{\beta}_{-1} := K\sqrt{\nu}, \\ \bar{\beta}_i := \bar{\beta}_{i-1} e^{2\lambda_i t_\Delta}\left(1 + 16 t_\Delta \bar{K}(\tilde{N} + \bar{K}M)\nu e^{3\lambda_i t_\Delta}\right) + 2\sigma h \sqrt{t_\Delta} e^{\lambda_i t_\Delta} \quad \text{for all } i \in I. \end{cases}$$

Assume further that ν *is small enough such that*

$$\begin{aligned} 4e^{2\lambda_i t_\Delta}(\tilde{N} + \bar{K}M)\sqrt{\nu}\, t_\Delta (e^{\lambda_i t_\Delta} \bar{\beta}_{i-1} + 2\sigma h \sqrt{t_\Delta}) < 1, \\ \bar{K}\sqrt{\nu} \le \frac{1}{2(\tilde{N}|\log \nu| + MK)\sqrt{\nu}\, t_\Delta}, \\ 2\sigma h \sqrt{t_\Delta} \le e^{\lambda_i t_\Delta} \bar{\beta}_{-1}, \end{aligned} \tag{5.3.28}$$

Then,

$$|y(t)| \le \bar{\beta}_i \le \bar{K}\sqrt{\nu} \quad \text{for all } t \in [\theta_i, \theta_{i+1}],\ t < \tau_{2h\sqrt{t_\Delta}}(\bar{\xi}),\ i \in I, \tag{5.3.29}$$

and

$$\mathbb{P}\left\{\tau_{2h\sqrt{t_\Delta}}(\bar{\xi}) < T_3/\nu\right\} \le n \frac{5p^2}{2} \exp\left(-\frac{h^2}{2}\right),$$

when h *and* p *satisfy* (5.3.8)

Remark 5.38. *a) Let* $\lambda : [T_2/\nu, T_3/\nu] \to \mathbb{R}$ *be defined through* $a(u) + \lambda(u) = b(u)e^{\lambda(u)r}$*. Obviously* $\lambda(T_2/\nu) = 0$*. Note, that this is only some mapping, and* $\lambda_+(\theta_i)$ *picks its value related to that one for* $i \in I$*. In particular, it does deliberately not provide a fancy transformation for the autonomous system. By the implicit function theorem*

$$\begin{aligned} \left.\frac{d\lambda(t)}{dt}\right|_{t = T_2/\nu} &= \frac{b'(T_2/\nu) - a'(T_2/\nu)}{1 - b(T_2/\nu)r} \\ &= -\nu \frac{f_{xt}(0, T_2/\nu) + g_{xt}(0, T_2/\nu)}{1 - g_x(0, T_2/\nu)r} =: \nu m_\lambda, \end{aligned}$$

where the prime denotes the time derivative, and m_λ *is independent of* ν*. Then, if* $T_3 - T_2$

is small enough such that $\lambda(\cdot)$ is monotone on the interval $[T_2/\nu, T_3/\nu]$, then

$$2t_\Delta \sum_{j=0}^{n-1} \lambda_j \approx \int_{T_2/\nu}^{T_3/\nu} \lambda(u)du \approx \int_0^{\frac{T_3-T_2}{\nu}} \nu m_\lambda u du = \frac{m_\lambda \nu}{2} \left(\frac{T_3 - T_2}{\nu} \right)^2,$$

where the approximations are rather informal in general, but provide a good approximation as soon as $T_3 - T_2$ is sufficiently small.

b) *For $\bar{K}/K > 1$ of order $\mathcal{O}(1)$ the choice (5.3.27) is possible, if $\sigma < \nu/|\log \nu|$ and if $2\nu T_\Delta \sum_{j=0}^{n-1} \lambda_j$ is not too big, which is true if $\frac{T_3-T_2}{\nu}$ is at most of order $\frac{1}{\sqrt{\nu}}$ for sufficiently small ν by a). Formulated in fast time, $T_3 - T_2$ may be of order $\sqrt{\nu}$.*

Proof of Theorem 5.37. The $\bar{\beta}_i$, $i \in I$, are explicitly given by

$$\bar{\beta}_i = \bar{\beta}_{-1} e^{2t_\Delta \sum_{j=0}^{i} \lambda_j} \prod_{j=0}^{i} \left(1 + 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_i t_\Delta} \right)$$
$$+ 2\sigma h \sqrt{t_\Delta} \sum_{j=0}^{i} e^{\lambda_j t_\Delta} e^{2t_\Delta \sum_{l=j+1}^{i} \lambda_l} \prod_{l=j+1}^{i} \left(1 + 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_i t_\Delta} \right) \quad \text{for all } i \in I.$$

We find that $\bar{\beta}_i \leq \bar{\beta}_{n-1}$ for all $i \in I$, where

$$\begin{aligned}
\bar{\beta}_{n-1} &\leq \bar{\beta}_{-1} \exp\left(2t_\Delta \sum_{j=0}^{n-1} \lambda_j \right) \prod_{j=0}^{n-1} \left(1 + 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_i t_\Delta} \right) \\
&\quad + 2\sigma h \sqrt{t_\Delta} \sum_{j=0}^{i} e^{\lambda_j t_\Delta} e^{2t_\Delta \sum_{l=j+1}^{i} \lambda_l} \prod_{l=j+1}^{i} \left(1 + 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_i t_\Delta} \right) \\
&\leq \bar{\beta}_{-1} \exp\left(2t_\Delta \sum_{j=0}^{n-1} \lambda_j \right) \exp\left(\frac{T_3 - T_2}{t_\Delta \nu} 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_\star^{(T_3)} t_\Delta} \right) \\
&\quad + 2\sigma h \sqrt{t_\Delta} e^{\lambda_\star^{(T_3)} t_\Delta} \frac{T_3 - T_2}{t_\Delta \nu} \exp\left(2t_\Delta \sum_{j=0}^{n-1} \lambda_j \right) \\
&\qquad \cdot \exp\left(\frac{T_3 - T_2}{t_\Delta \nu} 16 t_\Delta \bar{K} (\tilde{N} + \bar{K} M) \nu e^{3\lambda_\star^{(T_3)} t_\Delta} \right).
\end{aligned}$$

The crucial point is that $\bar{\beta}_{-1} > \|y_{T_2/\nu}\|$ and that

$$\begin{aligned}
\bar{\beta}_i &\geq e^{\lambda_i t_\Delta} \left(\bar{\beta}_{i-1} e^{\lambda_i t_\Delta} + 2\sigma h \sqrt{t_\Delta} + (\tilde{N} + \bar{K} M) \sqrt{\nu} t_\Delta \bar{\beta}_i^2 \right) \\
&\geq e^{\lambda_i t_\Delta} \left(\bar{\beta}_{i-1} e^{\lambda_i t_\Delta} + 2\sigma h \sqrt{t_\Delta} + \tilde{N} \sqrt{\nu} t_\Delta \bar{\beta}_i^2 + t_\Delta M \bar{\beta}_i^3 \right) \quad \text{for all } i \in I.
\end{aligned} \tag{5.3.30}$$

If we for a moment denote $R := \tilde{N} + \bar{K} M$, this can be seen through

$$\begin{aligned}
&\bar{\beta}_i \geq e^{\lambda_i t_\Delta} \left(\bar{\beta}_{i-1} e^{\lambda_i t_\Delta} + 2\sigma h \sqrt{t_\Delta} + R \sqrt{\nu} t_\Delta \bar{\beta}_i^2 \right) \\
\Leftrightarrow\ &\bar{\beta}_i^2 - \frac{\bar{\beta}_i}{e^{\lambda_i t_\Delta} R \sqrt{\nu} t_\Delta} + \frac{\bar{\beta}_{i-1} e^{\lambda_i t_\Delta} + 2\sigma h \sqrt{t_\Delta}}{R \sqrt{\nu} t_\Delta} \leq 0 \quad \text{for all } i \in I,
\end{aligned}$$

which is true if $\bar\beta_i \in [\bar\beta_i^{(1)}, \bar\beta_i^{(2)}]$ with

$$\begin{aligned}
\bar\beta_i^{(1)} &= \frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta} - \sqrt{\left(\frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta}\right)^2 - \frac{\bar\beta_{i-1}e^{\lambda_i t_\Delta} + 2\sigma h\sqrt{t_\Delta}}{R\sqrt{\nu}t_\Delta}} \\
&= \frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta} - \frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta}\sqrt{1 - \frac{4e^{2\lambda_i t_\Delta} R^2\nu t_\Delta^2(e^{\lambda_i t_\Delta}\bar\beta_{i-1} + 2\sigma h\sqrt{t_\Delta})}{R\sqrt{\nu}t_\Delta}} \\
&< \frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta} - \frac{1}{2e^{\lambda_i t_\Delta} R\sqrt{\nu}t_\Delta}\Bigg(1 - \frac{2e^{2\lambda_i t_\Delta} R^2\nu t_\Delta^2(e^{\lambda_i t_\Delta}\bar\beta_{i-1} + 2\sigma h\sqrt{t_\Delta})}{R\sqrt{\nu}t_\Delta} \\
&\qquad - 8e^{4\lambda_i t_\Delta} R^2\nu t_\Delta^2(e^{\lambda_i t_\Delta}\bar\beta_{i-1} + 2\sigma h\sqrt{t_\Delta})^2\Bigg) \\
&= e^{2\lambda_i t_\Delta}\bar\beta_{i-1} + 2e^{\lambda_i t_\Delta}\sigma h\sqrt{t_\Delta} + 4e^{3\lambda_i t_\Delta}R\sqrt{\nu}t_\Delta(e^{\lambda_i t_\Delta}\bar\beta_{i-1} + 2\sigma h\sqrt{t_\Delta})^2 \quad \text{for all } i \in I.
\end{aligned}$$

Then, with the assumption $2\sigma h\sqrt{t_\Delta} \le e^{\lambda_i t_\Delta}\bar\beta_{-1}$, we obtain that

$$\bar\beta_i^{(1)} < e^{2\lambda_i t_\Delta}\bar\beta_{i-1} + e^{\lambda_i t_\Delta}2\sigma h\sqrt{t_\Delta} + 16e^{3\lambda_i t_\Delta}R\sqrt{\nu}t_\Delta(e^{\lambda_i t_\Delta}\bar\beta_{i-1})^2 \quad \text{for all } i \in I.$$

Now, we use that $\beta_{i-1} \le \bar K\sqrt{\nu}$ and end up with

$$\begin{aligned}
\bar\beta_i^{(1)} &< e^{2\lambda_i t_\Delta}\bar\beta_{i-1} + e^{\lambda_i t_\Delta}2\sigma h\sqrt{t_\Delta} + 16e^{5\lambda_i t_\Delta}R\bar K\nu t_\Delta\bar\beta_{i-1} \\
&= e^{2\lambda_i t_\Delta}\bar\beta_{i-1}(1 + 16e^{3\lambda_i t_\Delta}R\bar K t_\Delta\nu) + 2e^{\lambda_i t_\Delta}\sigma h\sqrt{t_\Delta} \quad \text{for all } i \in I.
\end{aligned}$$

Moreover, it is obvious that $e^{2\lambda_i t_\Delta}\bar\beta_{i-1}(1 + 16e^{3\lambda_i t_\Delta}\bar K R\nu) + 2e^{\lambda_i t_\Delta}\sigma h\sqrt{t_\Delta} < \bar\beta_i^{(2)}$, because of assumption (5.3.28). With $\tau_{\bar\beta}(y)$ as defined in (5.3.4), we have that

$$\begin{aligned}
|y(t)| &\le e^{\lambda_i t_\Delta}\Bigg(\sup_{u\in[T_2/\nu - r,\theta_i]}|y(u)|e^{\lambda_i t_\Delta} + 2h\sigma\sqrt{t_\Delta} + t_\Delta\tilde N\sqrt{\nu}\left(\sup_{u\in[T_2/\nu - r,t]}|y(u)|\right)^2 \\
&\qquad + t_\Delta M\left(\sup_{u\in[T_2/\nu - r,\theta_{i+1}]}|y(u)|\right)^3\Bigg) \\
&< e^{\lambda_i t_\Delta}\left(\bar\beta_{i-1}e^{\lambda_i t_\Delta} + 2h\sigma\sqrt{t_\Delta} + t_\Delta\tilde N\sqrt{\nu}\bar\beta_i^2 + t_\Delta M\bar\beta_i^3\right) \\
&\le e^{\lambda_i t_\Delta}\left(\bar\beta_{i-1}e^{\lambda_i t_\Delta} + 2h\sigma\sqrt{t_\Delta} + t_\Delta\tilde N\sqrt{\nu}\bar\beta_i^2 + t_\Delta\bar K\sqrt{\nu}M\bar\beta_i^2\right) \le \bar\beta_i \\
&\qquad \text{for all } t \in [\theta_i, \theta_{i+1}], t \le \tau_{\bar\beta}(y) \wedge \tau_{2h\sqrt{t_\Delta}}(\bar\xi),
\end{aligned}$$

which is true by iteration starting with the interval $[\theta_0, \theta_1]$ and the assumption that $\|y_{T_2/\nu}\| < \bar\beta_{-1}$. Through (5.3.30) that actually provides that $\tau_{\bar\beta}(y) > \tau_{2h\sqrt{t_\Delta}}(\bar\xi)$ and the statement follows. □

5.3.4. Uniformly Symmetric Environment

As an addendum to the subsections 5.3.2 and 5.3.3 we consider a special case of reinforcements on the Assumptions 5.30, namely that the potentials F and G are symmetric throughout the transition phase rather than only at the point of transition T_2/ν. It is worth mentioning that both Example 5.32 a) and b) satisfy these conditions.

Assumption 5.39. *In addition to Assumptions 5.30 we assume that:*

- *Symmetry of the potentials F and G is fulfilled throughout $[T_0/\nu, T_2/\nu]$, i.e. that nonlinearity is of cubic order.*
- *Stricter assumption on the noise amplifier. We assume that $\sigma < \nu^2/|\log \nu|$.*

In that special case, we can apply a variant of the above argument of Theorem 5.35 to show that the deviation y remains in a neighborhood of size ν. Due to the similarity with the proofs of the previous two subsections, we will confine to repeat the main arguments only in the first case. Actually, it suffices to reconstruct the time interval, the step sizes and the definition of $(\beta_i^\star)_{i\in I}$. Then, the above proof of Theorem 5.35 can be used as a template. For notational comfort and by the flexibility for the choice of T_0, we decide to consider $(T_2 - T_0) = 1$. Note that by the general symmetry assumption, the estimate (5.3.12) simplifies to

$$|y(t)| < \sup_{u\in[T_0/\nu - r, \theta_i]} |y(u)| + 2h\sigma\sqrt{t_\Delta} + M\left(\sup_{u\in[T_0/\nu - r,\theta_{i+1}]} |y(u)|\right)^3 t_\Delta \qquad \text{for } t \in [\theta_i, \theta_{i+1}], t \le \tau_{2h\sqrt{t_\Delta}}(\xi), i \in I, \tag{5.3.31}$$

where $\tau_{2h\sqrt{t_\Delta}}(\xi)$ is defined analogue to (5.3.9) for ξ analogue to (5.3.7).

Corollary 5.40. *In the situation of Assumption 5.39 let $t_\Delta := \frac{T_2 - T_0}{n\nu} < r$, where $n = \frac{1}{t_\Delta \nu}$ is the number of steps from a uniformly stable regime to the point where stability is lost. Assume that there is $k > 0$ such that $\|y_{T_0/\nu}\| \le k\nu$. Assume further that $h = \mathcal{O}\left(\log\left(\frac{|\log\nu|}{\nu}\right)\right)$ and that ν is small enough such that*

$$ke^{16MK^2\nu^2 t_\Delta \frac{|\log\nu|}{t_\Delta\nu}} + \frac{\sigma}{t_\Delta\nu^2} 2h\sqrt{t_\Delta}\, e^{16MK^2\nu^2 t_\Delta \frac{|\log\nu|}{t_\Delta\nu}} < K. \tag{5.3.32}$$

We define

$$\beta_{-1}^\star := k\nu \quad \text{and} \quad \beta_i^\star := \beta_{i-1}^\star(1 + MK^2\nu^2 t_\Delta) + 2h\sigma\sqrt{t_\Delta} \quad \text{for all } i \in \{0, \dots, n-1\},$$

and assume that

$$4MK\nu t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}) < 1 \quad \text{for all } i \in I, \tag{5.3.33}$$

$$K\nu < \frac{1}{2MK\nu t_\Delta}, \tag{5.3.34}$$

$$2h\sigma\sqrt{t_\Delta} \le k\nu. \tag{5.3.35}$$

Then the following holds true:

a) We have that $\beta_{n-1}^\star \le K\nu$.

b) For all $i \in I$ we have that $\beta_i^\star \ge \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + M(\beta_i^\star)^3$.

c) The family $(\beta_i^\star)_{i\in\{-1,0,1,\dots,n-1\}}$ *constitutes an upper bound for* y*, i.e.* $|y(t)| \le \beta_i^\star$ *for all* $t \in [\theta_i, \theta_{i+1}]$*,* $t < \tau_{2h\sqrt{t_\Delta}}(\xi)$ *and* $\{-1,\dots,n-1\}$*. The corresponding probability is provided through Lemma 5.33, and given by*

$$\mathbb{P}\left\{\tau_{2h\sqrt{t_\Delta}}(\xi) < T_2/\nu\right\} \le n\frac{5p^2}{2}\exp\left(-\frac{h^2}{2}\right),$$

when h *and* p *satisfy* (5.3.8).

Remark 5.41. *In order to satisfy* (5.3.32)*, it is necessary to have* $\sigma < \frac{\nu^2}{|\log \nu|}$.

Proof. a) Due to the same differential law of the deviation y, we can take over the property (5.3.12), that provides a step-wise upper bound. Here, by absence of nonlinear terms of quadratic order, we obtain

$$|y(t)| < \sup_{u\in[T_0/\nu-r,\theta_i]} |y(u)| + 2h\sigma\sqrt{t_\Delta} + M\left(\sup_{u\in[T_0/\nu-r,\theta_{i+1}]} |y(u)|\right)^3 t_\Delta$$
$$\text{for } t\in[\theta_i,\theta_{i+1}], t\le \tau_{2h\sqrt{t_\Delta}}(\xi), i\in I.$$

Then, through the recursive definition, we have that

$$\begin{aligned}\beta_i^\star \le \beta_{n-1}^\star &\le \beta_{-1}^\star e^{16MK^2\nu^2 t_\Delta \frac{1}{t_\Delta\nu}|\log\nu|} + \frac{1}{t_\Delta\nu}|\log\nu|2h\sigma\sqrt{t_\Delta}\, e^{16MK^2\nu^2 t_\Delta \frac{1}{t_\Delta\nu}|\log\nu|}.\\ &< \nu e^{16MK^2\nu^2 t_\Delta \frac{1}{t_\Delta\nu}|\log\nu|}\left(k + \frac{1}{t_\Delta\nu^2}|\log\nu|2h\sigma\sqrt{t_\Delta}\right) < K\nu.\end{aligned}$$

b) First, we fix $i\in I$ and show that $\beta_i^\star \ge \beta_{i-1}^\star + 2\sigma h\sqrt{t_\Delta} + MK\nu(\beta_i^\star)^2 t_\Delta$. Reformulating the inequality yields

$$\begin{aligned}&\beta_i^\star \ge \beta_{i-1}^\star + 2\sigma h\sqrt{t_\Delta} + MK\nu(\beta_i^\star)^2 t_\Delta\\ &\Leftrightarrow (\beta_i^\star)^2 - \frac{1}{MK\nu t_\Delta}\beta_i^\star + \frac{1}{MK\nu t_\Delta}\left(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta}\right) \le 0.\end{aligned}$$

This is true through assumption (5.3.33) if $\beta_i^\star \in [\beta_i^{(\star1)}, \beta_i^{(\star2)}]$ with

$$\begin{aligned}\beta_i^{(\star1)} &= \frac{1}{2MK\nu t_\Delta} - \sqrt{\frac{1}{4M^2K^2\nu^2t_\Delta^2} - \frac{1}{MK\nu t_\Delta}\left(\beta_{i-1} + 2h\sigma\sqrt{t_\Delta}\right)}\\ &= \frac{1}{2MK\nu t_\Delta} - \frac{1}{2MK\nu t_\Delta}\sqrt{1 - \frac{4M^2K^2\nu^2t_\Delta^2}{MK\nu t_\Delta}\left(\beta_{i-1} + 2h\sigma\sqrt{t_\Delta}\right)}.\end{aligned}$$

Note that $\beta_i^{(\star2)} \in \mathcal{O}(\frac{1}{\nu t_\Delta})$. We use the fact that $\sqrt{1-x} > 1 - \frac{x}{2} - \frac{x^2}{2}$ for $x\in(0,1)$ and obtain that

$$\begin{aligned}\beta_i^{(\star1)} &< \frac{1}{2MK\nu t_\Delta} - \frac{1}{2MK\nu t_\Delta}\Big(1 - 2MK\nu t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta})\\ &\qquad - 8M^2K^2\nu^2t_\Delta^2(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta})^2\Big)\\ &= \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + 4MK\nu t_\Delta(\beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta})^2.\end{aligned}$$

With assumption (5.3.35) we obtain that

$$\begin{aligned}\beta_i^{(\star 1)} &< \beta_{i-1}^\star + 2h\sigma\sqrt{t_\Delta} + 16MK^2\nu^2 t_\Delta \beta_{i-1}^\star \\ &= \beta_{i-1}^\star(1 + 16MK^2\nu^2 t_\Delta) + 2h\sigma\sqrt{t_\Delta}.\end{aligned}$$

Furthermore, through assumption (5.3.34) we obtain that

$$\beta_{i-1}^\star(1 + 16MK^2\nu^2 t_\Delta) + 2h\sigma\sqrt{t_\Delta} < \beta_i^{(\star 2)},$$

and hence, as desired,

$$\beta_i^\star \geq \beta_{i-1}^\star + 2\sigma h\sqrt{t_\Delta} + MK\nu(\beta_i^\star)^2 t_\Delta \geq \beta_{i-1}^\star + 2\sigma h\sqrt{t_\Delta} + M(\beta_i^\star)^3 t_\Delta.$$

c) Completely analogue to Part c) of Theorem 5.35 □

Remark 5.42.

- *Assumptions (5.3.33), (5.3.34) and (5.3.35) are naturally fulfilled if ν is small enough through the fact that $\beta_{n-1}^\star \leq K\nu$.*
- *Regarding the necessary smallness of σ, one realizes that an additional correction term of order ν in formula (5.3.31) cannot be compensated through the procedure in the same way Remark 5.31 indicated for Theorem 5.35. It is therefore necessary to assume that $x^\star = x^{\mathrm{ad}_\nu} = 0$ throughout $[T_0/\nu, T_2/\nu]$.*
- *The assumption of absence of quadratic nonlinear terms is crucial, because starting at T_0/ν in the uniformly stable phase invalidates plausible attempts providing quadratic nonlinearity to be small in terms of ν as we did in (5.3.3) when starting from T_1/ν.*

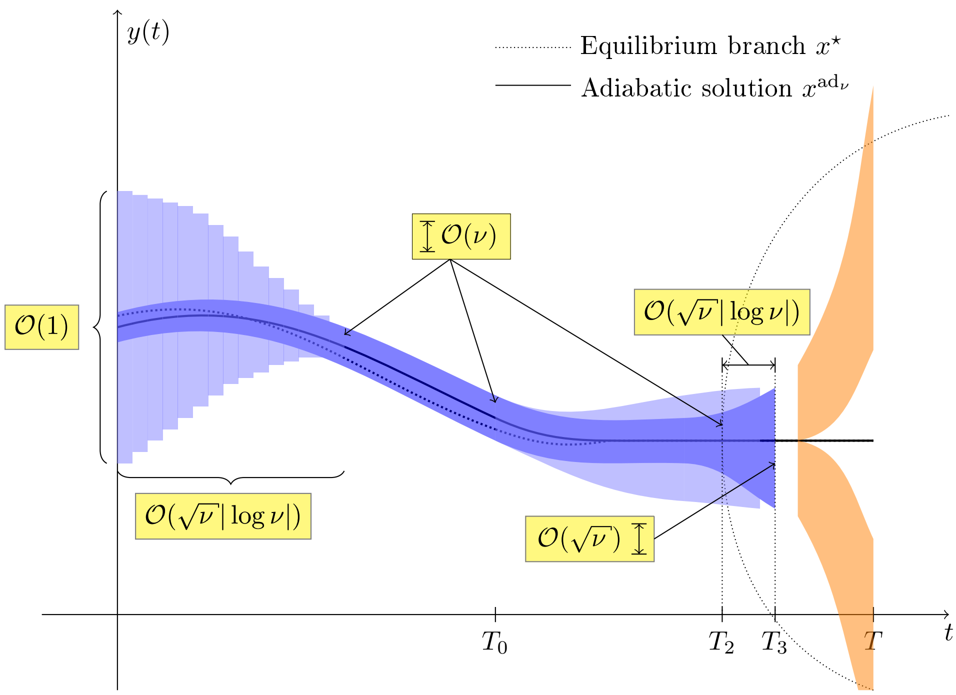

Figure 12: Sketch of the typical path behavior during different transition phases for the special case of a pitchfork bifurcation with uniformly symmetric F, G. The lighter blue area is a reference from the previous more general case. Time points and durations are formulated in fast time.

Instead of formulating an analogue theorem for the phase between T_2/ν and T_3/ν for the uniformly symmetric special case, we provide the respective results in the following remark.

Remark 5.43 (Uniformly symmetric pitchfork continued from T_2/ν)**.** *As we have seen, Corollary 5.40 can be easily deduced from Theorem 5.35. And in the same way, Theorem 5.37 can be modified to the case of uniformly symmetric potentials. We will not formulate such a corollary, but only have a look at the implications. To that end, we continue from Corollary 5.40 where y remains in a distance at most of order ν around $x^\star = x^{\mathrm{ad}_\nu} = 0$, if $\sigma < \frac{\nu^2}{|\log \nu|}$. In that case the results from Remark 5.38 b) can be improved a bit, because Condition 5.3.27 is a bit easier to fulfill, which is to say, that the maximal time T_3/ν for which the process remains at most of order $\sqrt{\nu}$ increases a bit. Regarding the Key condition* (5.3.27)*, in order to have y remaining in a distance of at most of order $\sqrt{\nu}$, the distance $T_3 - T_2$ must allow for a small parameter ν to fulfill*

$$K\sqrt{\nu}\exp\left(2t_\Delta\sum_{j=0}^{n-1}\lambda_j\right)\exp\left(\frac{T_3-T_2}{t_\Delta\nu}16t_\Delta\bar{K}^2M\nu e^{3\lambda_\star^{(T_3)}t_\Delta}\right)$$

$$+2\frac{\sigma}{\sqrt{\nu}}h\sqrt{t_\Delta}e^{\lambda_\star^{(T_3)}t_\Delta}\frac{T_3-T_2}{t_\Delta\nu}\exp\left(2t_\Delta\sum_{j=0}^{n-1}\lambda_j\right)\exp\left(\frac{T_3-T_2}{t_\Delta\nu}16t_\Delta\bar{K}^2M\nu e^{3\lambda_\star^{(T_3)}t_\Delta}\right)<\bar{K}.$$

In reducing the size of ν this is satisfiable as long as $2t_\Delta\sum_{j=0}^{n-1}\lambda_j$ is not yet big. With the above considerations in Remark 5.38, we may conclude that one finds ν small enough such that the above inequality holds as long as $\frac{T_3-T_2}{\nu}$ is at most of size $\frac{|\log\nu|}{\sqrt{\nu}}$ meaning that we gain an additional logarithmic factor of time in which the deviation process does not leave a distance of order $\sqrt{\nu}$. We obtain a modified version of Figure 8 through Figure 12.

5.3.5. On the Choice of Delay Influence

All the presented approaches to concentration inequalities or contraction-like behavior including the Subsections 5.2.1 to 5.3.4 actually have a Halanay-type inequality as their core arguments. That means that – in one way or the other – an estimate for the future evolution of a solution path is constructed using the supremum of the solution path over the preceding delay-length interval. Of course, that technique covers several formulations of time-delayed influence at once and we initially decided for a very simple one due to notational comfort. Other possible formulations in place of (5.1.3) that would still satisfy all of the presented results in the subsections mentioned above are for instance:

- $$dx(t)=\Big[f(x(t),\nu t)+g(x(t-\tilde{r}(t/\nu)),\nu t)\Big]dt+\sigma dW(t)\text{ for }t\in[0,T/\nu],$$

  where $\tilde{r}:[0,T]\to[0,r]$ is a mapping with values in $[0,r]$.
- $$dx(t)=\Big[f(x(t),\nu t)+\sum_{i=1}^{n}b_i(t/\nu)x(t-r_i)\Big]dt+\sigma dW(t)\text{ for }t\in[0,T/\nu],$$

  where $r_1,\dots,r_n\in[0,r]$ and continuous $b_i:[0,T]\to\mathbb{R}$ for all $i\in\{1,\dots,n\}$ such that $\sum_{i=1}^{n}|b_i(t)|<a(t)$.
- $$dx(t)=\Big[f(x(t),\nu t)+L(t)x_t\Big]dt+\sigma dW(t)\text{ for }t\in[0,T/\nu],$$

  where $L:[0,T/\nu]\times\mathcal{C}(J,\mathbb{R})\to\mathbb{R}$ continuous and linear with $\|L(t)\|\leq b$.

5.4. Departure From Instability

This work has predominantly focused on concentration inequalities to limit the probability of an escape from an area that usually was formulated in terms of standard deviations or closely related quantities. An exception is Section 4.4 that has provided insight to the small-ball probabilities of SDDEs in the critical regime, where an analogy to a properly rescaled Brownian motion was established. Key ingredients have been the well-known small-ball probabilities of Brownian motion and the convergence of fundamental solutions that we established in the earlier Section 4.1. In this final part of the transition of a solution to an SRFDE into an unstable regime, we address ourselves to the question how much time is sufficient for the solution to escape from a neighborhood of order 1 along the equilibrium branch. We will further simplify the considered system through the following set of assumptions:

Assumption 5.44.

- *We keep the assumption that* $0 < b_- \leq b(t) \leq b_+$ *and* $0 < a_- \leq a(t) \leq a_+$ *for all* $t \in [T_3/\nu, T/\nu]$*, where* a_-, a_+, b_-, b_+ *are independent of* ν.
- *With regard to the results of Section 5.3.3 as well as Section 5.3.4, we assume that* $T_3/\nu - T_2/\nu = \mathcal{O}(\sqrt{\nu})$*, and that there is* $c_3, \bar{c}_3 > 0$ *such that*

$$c_3\sqrt{\nu} \leq b(T_3/\nu) - a(T_3/\nu) \leq \bar{c}_3\sqrt{\nu}.$$

- *The systems keeps turning more and more unstable immediately after passing through* T_3/ν*. We assume that there are positive constants* m_b *and* m_a *such that:*

$$0 \leq \frac{db(t)}{dt} \leq m_b\nu \quad \text{and} \quad -m_a\nu \leq \frac{da(t)}{dt} \leq 0 \quad \text{for all } t \geq T_3/\nu,$$

  which includes that $b(t) > a(t)$ *for all* $t \geq T_3/\nu$.

- *All nonlinear terms will be neglected, in particular we will study the solution of*

$$\begin{cases} dx(t) = -a(t)x(t)dt + b(t)x(t-r)dt + \sigma dW(t) & \text{for } t \geq T_3/\nu, \\ x_{T_3/\nu} = \Upsilon. \end{cases} \tag{5.4.1}$$

  Note that due to the absence of nonlinear terms, we implicitly keep the assumption that $x^\star = x^{\mathrm{ad}_\nu} = 0$ *for the remaining time interval* $[T_3/\nu, T/\nu]$.

- *In foresight we assume that* $T - T_3$ *is at most of order* $\sqrt{\nu}|\log\sigma|$.

The results so far have revealed that typically $\|\Upsilon\| \in \mathcal{O}(\sqrt{\nu})$, for sufficiently small ν and $\sigma < \frac{\nu}{|\log\nu|}$. We keep denoting the fundamental solution by $\check{x} = (\check{x}(t,u), u \geq T_3/\nu, t \geq u-r)$, and let the corresponding deterministic solution semi group $(T^{\mathrm{det}}_{t,u} : u \geq T_3/\nu, t \geq u-r)$ map from $\mathcal{C}(J,\mathbb{R})$ to $\mathcal{C}(J,\mathbb{R})$. Then, by Theorem 3.5, we may represent the solution as

$$x(t) = T^{\mathrm{det}}_{t,T_3/\nu}\Upsilon(0) + \sigma\int_{T_3/\nu}^{t} \check{x}(t,u)dW(u) \quad \text{for all } t \in [T_3/\nu, T/\nu].$$

For the generalized fundamental solution $\check{x}$ we have so far only developed upper-bound estimates while in this section we will need a lower-bound estimate. This will be achieved by transforming the fundamental solution so that it solves a nonautonomous DDE with coefficients that coincide in every point in time. Then, we will make use of slow system evolution and our knowledge on the constant-coefficient case. The transformation is not as obvious as its kind have been in the constant-coefficient case. As a key result of this section, the following lemma provides the existence of a *nice* process $c : [T_3/\nu, T/\nu] \to \mathbb{R}$ such that

$$\tilde{a}(t) := a(t) + c(t) = b(t) \exp\left(-\int_{t-r}^{t} c(s)ds\right) =: \tilde{b}(t) \quad \text{for all } t \in [T_3/\nu, T/\nu]. \tag{5.4.2}$$

The importance lies in the fact that for $\gamma(t,s) = \int_s^t c(u)du$ the transformed fundamental solution

$$\check{X}(t,u) : = \exp\left(-\gamma(t,u)\right) \check{x}(t,u) \quad \text{for all } u \geq T_3/\nu, t \geq u - r \tag{5.4.3}$$

again constitutes a fundamental solution and solves

$$\begin{cases} d\check{X}(t,u) = -\tilde{a}(t)\check{X}(t,u)dt + \tilde{b}(t)\check{X}(t-r,u)dt & \text{for } t \geq T_3/\nu, u \leq t, \\ \check{X}(t,u) = \mathbb{1}_{\{u\}}(t) & \text{for } t \in [u-r,u]. \end{cases} \tag{5.4.4}$$

It is worth emphasizing that the simplified notation, for instance $a(t) = f_x(x^\star(t), \nu t)$, tends to hide the fact that the process c depends on ν.

Lemma 5.45. *Let $a, b : [t_0, t_1] \to \mathbb{R}$ be nonnegative and continuous for arbitrary finite $0 < t_0 < t_1$, denote $b_+ := \|b\|_{[t_0,t_1]}$ and $a_+ := \|a\|_{[t_0,t_1]}$. Define $\mathcal{H} : \mathcal{C}\big([t_0 - r, t_1], \mathbb{R}\big) \to \mathcal{C}\big([t_0 - r, t_1], \mathbb{R}\big)$, $h \mapsto \mathcal{H}(h)$, pointwisely through*

$$\mathcal{H}(h)(t) := \begin{cases} b(t) \exp\left(-\int_{t-r}^{t} h(u)du\right) - a(t) & \text{for all } t \in [t_0, t_1], \\ h(t) & \text{for } t \in [t_0 - r, t_0]. \end{cases} \tag{5.4.5}$$

Assume that $h \in \mathcal{C}\big(J, [-a_+, b_+e^{a_+r}]\big)$ satisfies

$$h(0) + a(t_0) = b(t_0) \exp\left(-\int_{-r}^{0} h(u)du\right), \tag{5.4.6}$$

and one example of such h is given by the constant mapping $h_{\text{const}} \in \mathcal{C}\big(J, \{h_{\text{const}}(0)\}\big)$, where $h_{\text{const}}(0)$ solves $h_{\text{const}}(0) + a(t_0) = b(t_0)e^{-h_{\text{const}}(0)r}$.

a) Then, there is a unique continuation $\overline{h} \in \mathcal{C}([t_0 - r, t_1], [-a_+, b_+e^{a_+r}])$ of h, i.e. satisfying $h(u) = \overline{h}(t_0 + u)$ for all $u \in J$, such that $\mathcal{H}(\overline{h}) = \overline{h}$.

b) The continuation $\overline{h}$ from a) is continuously differentiable over (t_0, t_1) and right continuous in t_0.

Proof. It is actually easy to see that $-a_+ \leq h_{\text{const}} \leq b_+e^{a_+}$. Therefore, h_{const} is a valid initial segment in the sense that it is an element of $\mathcal{C}(J, [-a_+, b_+e^{a_+r}])$.

a) For $h_0 \in \mathcal{C}\big(J, \mathbb{R}\big)$ we denote $\mathcal{C}^{(h_0)}$ for the set of continuous functions with initial segment h_0, i.e. $f \in \mathcal{C}^{(h_0)}\big([t_0, t_1], \mathbb{R}\big)$, if $f \in \mathcal{C}\big([t_0 - r, t_1], \mathbb{R}\big)$ and

$$f(t_0 + u) = h_0(u) \quad \text{for all } u \in J.$$

Then, it is easy to check that

$$\mathcal{H}\Big(\mathcal{C}^{(h_0)}\big([t_0,t_1],[-a_+,b_+e^{a+r}]\big)\Big) \subset \mathcal{C}^{(h_0)}\big([t_0,t_1],[-a_+,b_+e^{a+r}]\big),$$
$$\text{if } h_0(u) \in [-a_+,b_+e^{a+r}] \text{ for all } u \in J.$$

The space $C := \mathcal{C}^{(h_0)}\big([t_0,t_1],[-a_+,b_+e^{a+r}]\big)$, equipped with the $\|\cdot\|$-norm, or topology of uniform convergence, is complete, i.e. it is a Banach space. Further, it is easy to see that it is bounded and convex. To justify the application of the Schauder fixed-point theorem, see [HVL93, Lemma 2.4, Section 2], it remains to show that $\mathcal{H}$ is completely continuous, which means that it takes weakly convergent sequences in C to (norm) convergent sequences in C. To this end we assume that $h, h_k \in C$, $k \in \mathbb{N}$, and that h_k weakly converges to h, i.e. for any continuous linear functional $f : C \to \mathbb{R}$, we have that $\lim_{k\to\infty} f(h_k) = f(h)$. So, for $f^{(t)}(g) := \int_{t-r}^{t} g(u)du$, $g \in C$, we know that

$$\lim_{k\to\infty} \int_{t-r}^{t} h_k(u)du = \int_{t-r}^{t} h(u)du \qquad \text{for all } t \in [t_0,t_1],$$

which shows that for every $t \in [t_0,t_1]$, we have $\lim_{n\to\infty} \mathcal{H}(h_n)(t) = \mathcal{H}(h)(t)$ (pointwise). To show that $\mathcal{H}(h_n)$ converges even uniformly to $\mathcal{H}(h)$, we let $\varepsilon > 0$ be arbitrary and let $(t_i)_{i\in\{1,\dots,n\}}$ denote a partition of $[t_0,t_1]$ defined such that

$$t_i = t_0 + \frac{i}{n}(t_1 - t_0) \qquad \text{for all } i \in \{0,1,\dots,n\}.$$

Then, for arbitrary $\delta > 0$ and every $n \in \mathbb{N}$ there is an $N = N(\delta, n)$ such that

$$\left| \int_{t_i-r}^{t_i} h_k(u) - h(u)du \right| < \delta \qquad \text{for all } i \in \{0,1\dots,n\},\ k \geq N. \tag{5.4.7}$$

And for some arbitrary $t \in (t_i, t_{i+1})$ for some $i \in \{0,1,\dots n\}$ we obtain that

$$\begin{aligned}
\big|\mathcal{H}(h_k)(t) - \mathcal{H}(h)(t)\big| &= b(t)\exp\left(-\int_{t-r}^{t} h(u)du\right)\left|1 - \exp\left(\int_{t-r}^{t} h(u) - h_k(u)du\right)\right| \\
= b(t)\exp\left(-\int_{t-r}^{t} h(u)du\right)\Bigg|1 - \exp\Bigg(&-\int_{t_i-r}^{t-r} h(u) - h_k(u)du \\
&+ \int_{t_i}^{t} h(u) - h_k(u)du \\
&+ \int_{t_i-r}^{t_i} h(u) - h_k(u)du\Bigg)\Bigg|
\end{aligned} \tag{5.4.8}$$

for all $k \geq N$. For an illustration of the integral decomposition see Figure 13.

Denoting $M_+ := \max\{a_+, b_+e^{a+r}\}$ boundedness of $h, h_n \in C$ representatively allows the following upper-bound estimate:

$$\left| \int_{t_i-r}^{t-r} h(u)du \right| \leq \frac{M_+}{n}.$$

Analogue estimates can be applied to three more terms in (5.4.8). By continuing from (5.4.8)

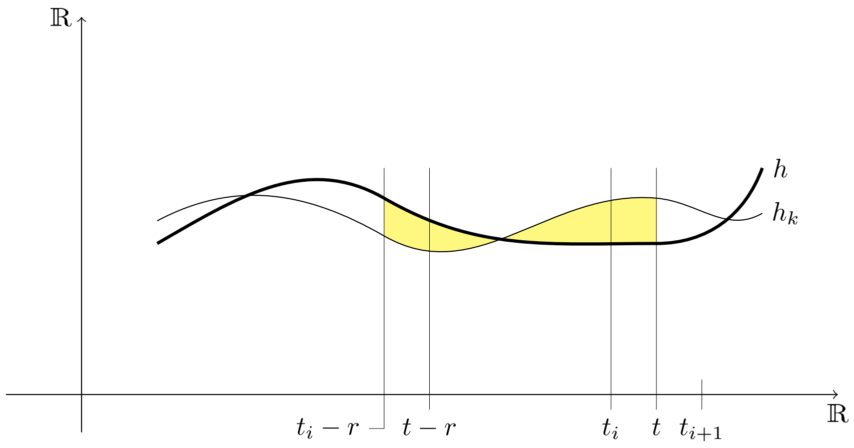

Figure 13: Illustration of the integral decomposition in (5.4.8).

and using (5.4.7) we find that

$$\big|\mathcal{H}(h_k)(t) - \mathcal{H}(h)(t)\big| \le b_+ e^{M_+ r} \max_{j\in\{-1,1\}} \left\{ \left| 1 - \exp\left(j\left(\frac{2M_+}{n} + \delta\right)\right)\right| \right\} \quad \text{for all } k \ge N.$$

That settles uniform convergence. Hence, we may apply the Schauder fixed-point theorem to obtain the existence of a continuation of h_0 in C, i.e.

$$\begin{aligned} & h_0 \in \mathcal{C}\big(J, [-a_+, b_+ e^{a_+ r}]\big) \\ & \quad \Rightarrow \quad \text{There is } \overline{h} \in \mathcal{C}^{(h_0)}\big([t_0,t_1], [-a_+, b_+e^{a_+ r}]\big) \text{ with } \overline{h}_{t_0} = h_0, \mathcal{H}(\overline{h}) = \overline{h}. \end{aligned} \tag{5.4.9}$$

b) The fact, that $\overline{h} = \mathcal{H}(\overline{h})$, also yields that $\overline{h}$ is differentiable over (t_0, t_1) with

$$\frac{d}{dt}\overline{h}(t) = \Big(b'(t) + b(t)(\overline{h}(t-r) - \overline{h}(t))\Big)e^{-\int_{t-r}^{t}\overline{h}(u)du} - a'(t) \quad \text{for all } t \in (t_0,t_1). \tag{5.4.10}$$

Here, $b'(t) = \frac{db(t)}{dt}$ and $a'(t) = \frac{da(t)}{dt}$. And the differential quotient's limit from the right also exists in t_0 due to continuity of h. It remains to show that the continuation $\overline{h} \in \mathcal{C}\big([t_0 - r, t_1], \mathbb{R}\big)$ is unique. To this end we assume that $h^{(1)}$ and $h^{(2)}$ are two continuations of h_0 that we assume to coincide up to some time $\hat{t} \in [t_0, t_1]$ and to differ on the interval $(\hat{t}, \hat{t} + \varepsilon)$ for some $\varepsilon > 0$, and without loss of generality $\varepsilon < r$, namely we assume that

$$h^{(1)}(t) = h^{(2)}(t) \text{ for all } t \in [t_0 - r, \hat{t}\,] \quad \text{and} \quad h^{(1)}(t) < h^{(2)}(t) \text{ for all } t \in (\hat{t}, \hat{t}+\varepsilon], \tag{5.4.11}$$

which is possible due to differentiability of $h^{(1)}$ and $h^{(2)}$. But then

$$\int_{\hat{t}-r+\varepsilon}^{\hat{t}+\varepsilon} h^{(2)}(u)du > \int_{\hat{t}-r+\varepsilon}^{\hat{t}+\varepsilon} h^{(1)}(u)du,$$

and therefore, because $b(t) > 0$ for all $t \ge T_3/\nu$, we find that

$$h^{(1)}(\hat{t}+\varepsilon) < h^{(2)}(\hat{t}+\varepsilon) = b(\hat{t}+\varepsilon)\exp\left(-\int_{\hat{t}-r+\varepsilon}^{\hat{t}+\varepsilon} h^{(2)}(u)du\right) - a(\hat{t}+\varepsilon) < h^{(1)}(\hat{t}+\varepsilon).$$

This contradiction settles uniqueness and the proof is complete. □

Actually, the continuation is not restricted to finite time interval as long as $a(\cdot)$ and $b(\cdot)$ remain nonnegative and continuously differentiable over $[0,\infty)$, but the quantities a_+ and b_+ are possibly no longer well-defined.

Corollary 5.46. *Given that $a, b : [t_0,\infty) \to (0,\infty)$ are continuously differentiable,*

$$h_0 \in \mathcal{C}([t_0 - r, t_0], [-\|a\|_{[t_0,T]}, \|b\|_{[t_0,T]} e^{\|a\|_{[t_0,T]} r}]) \quad \text{for some } T > t_0, \tag{5.4.12}$$

and $\mathcal{H}$ is defined analogously to (5.4.5) for all $t \in [t_0,\infty)$, then there is a unique continuation $\overline{h} \in \mathcal{C}^{(h_0)}([t_0,\infty),\mathbb{R})$ with $\overline{h} = \mathcal{H}(\overline{h}) : [0,\infty) \to \mathbb{R}$.

Proof. For $a_+ := \|a\|_{[t_0,T]}$, $b_+ := \|b\|_{[t_0,T]}$ the previous Lemma yields a continuation $\overline{h}$ over $[t_0 - r, T]$. We know that $\|\overline{h}\|_{[t_0-r,T]} \in [-a_+, b_+ e^{a_+ r}]$ and define $\tilde{h}(u) := \overline{h}(T+u)$ for all $u \in J$. Then for given $\hat{T} > T$, $\|\tilde{h}\|_J \in [-\|a\|_{[0,\hat{T}]}, \|b\|_{[0,\hat{T}]} e^{\|a\|_{[0,\hat{T}]} r}]$ and therefore, $\tilde{h}$ is a feasible initial segment and the previous lemma implies a unique continuation to the interval $[0,\hat{T}]$. Repeating this argument yields the claim. □

Remark 5.47. *Note that the previous Lemma 5.45 and Corollary 5.46 apply in relatively general situations. But, in order to gather an upper and a lower bound as well as a uniformly upper bound for differential, the below lemma will require the entire scope of Assumption 5.44.*

Lemma 5.48. *Let the Assumptions 5.44 hold. For given ν let $\overline{h}$ denote the continuation of the constant mapping $h_{\text{const}}^{(T_3/\nu)} \in \mathcal{C}(J,\mathbb{R})$, suggested in Lemma 5.45, with $h_{\text{const}}^{(T_3/\nu)}(t) := c_\star$ for all $t \in J$, where $c_\star$ is uniquely defined as the solution of*

$$a(T_3/\nu) + c_\star = b(T_3/\nu) e^{-c_\star r}.$$

Then,

a) The continuation $\overline{h}$ never falls below the level $c_\star$, i.e.

$$\overline{h}(t) \geq c_\star \quad \text{for all } t \in [T_3/\nu, T/\nu].$$

b) The continuation $\overline{h}$ never overcomes $b(t) - a(t)$, i.e.

$$\overline{h}(t) \leq b(t) - a(t) \quad \text{for all } t \in [T_3/\nu, T/\nu].$$

c) There is a constant $\bar{m}_+ > 0$ such that

$$\sup_{t \in [T_3/\nu, T/\nu]} \left| \frac{d\overline{h}(t)}{dt} \right| < \sqrt{\nu}\, \bar{m}_+,$$

and $\bar{m}_+$ is independent of ν and at most of order $|\log\sigma|$.

Proof. a) Let $\tau_{c_\star} := \inf\{t \geq T_3/\nu : \overline{h}(t) < c_\star\}$ denote the deterministic first exit time of the continuation $\overline{h}$ from the nonnegative half line $[c_\star,\infty)$. Suppose that $\tau_{c_\star} < T/\nu$. Then due to the fact that $\overline{h}'(T_3/\nu) > 0$ because of (5.4.10), there is $\varepsilon > 0$ with $\overline{h}(\tau_{c_\star} + s) < c_\star$ for all $s \in (0,\varepsilon]$. Without loss of generality we let $\varepsilon < r/2$. But then, on the one hand

$$a(\tau_{c_\star} + \varepsilon) - a(\tau_{c_\star}) + \overline{h}(\tau_{c_\star} + \varepsilon) - \overline{h}(\tau_{c_\star}) < 0,$$

because $a(\cdot)$ is nonincreasing and $\overline{h}(\tau_{c_\star}+\varepsilon)-\overline{h}(\tau_{c_\star})<0$ by construction. On the other hand,

$$b(\tau_{c_\star}+\varepsilon)\exp\left(-\int_{\tau_{c_\star}+\varepsilon-r}^{\tau_{c_\star}+\varepsilon}\overline{h}(u)du\right)-b(\tau_{c_\star})\exp\left(-\int_{\tau_{c_\star}-r}^{\tau_{c_\star}}\overline{h}(u)du\right)>0,$$

because $b(\cdot)$ is nondecreasing and

$$\int_{\tau_{c_\star}+\varepsilon-r}^{\tau_{c_\star}+\varepsilon}\overline{h}(u)du<\int_{\tau_{c_\star}-r}^{\tau_{c_\star}}\overline{h}(u)du.$$

But that is a contradiction to the fixed-point property that guarantees that

$$a(t)+\overline{h}(t)=b(t)\exp\left(-\int_{t-r}^{t}\overline{h}(u)du\right)\qquad\text{especially for } t\in\{\tau_{c_\star},\tau_{c_\star}+\varepsilon\}.$$

b) After we know from part a) that $\overline{h}$ is actually nonnegative, this can easily be seen from the fixed-point property.
c) As we assumed that $T/\nu-T_3/\nu=\mathcal{O}(\sqrt{\nu}|\log\sigma|)$, we know that there is a constant $m_{b,a}$ at most of order $|\log\sigma|$ and independent of ν, such that

$$\sup_{t\in[T_3/\nu,T/\nu]} b(t)-a(t)\le c_3\sqrt{\nu}+(m_b+m_a)\nu\frac{T-T_3}{\nu}\le m_{b,a}\sqrt{\nu}.$$

Then

$$\left|\frac{d}{dt}\overline{h}(t)\right|=\left|\left(b'(t)+b(t)(\overline{h}(t-r)-\overline{h}(t))\right)\right|e^{-\int_{t-r}^{t}\overline{h}(u)du}+|a'(t)|<\bar{m}_+\sqrt{\nu}$$
$$\text{for all } t\in[T_3/\nu,T/\nu]$$

for some appropriate constant $\bar{m}_+>0$ at most of order $|\log\sigma|$ and independent of ν. This is because $\left|\overline{h}(t-r)-\overline{h}(t)\right|<|b(t)-a(t)-c_\star|$ for all $t\in[T_3/\nu,T/\nu]$. □

Returning to the solution of (5.4.1), we let $(\check{x}(t,u):u\ge T_3/\nu,\ t\ge u-r)$ denote the corresponding fundamental solution, and let $(T^{\text{det}}_{t,u}:u\ge T_3/\nu,\ t\ge u-r)$ denote the solution semi group of the corresponding deterministic system. The solution process of (5.4.1) admits the representation

$$x(t)=T^{\text{det}}_{t,T_3/\nu}\Upsilon(0)+\xi(t)\quad\text{where}\quad\xi(t)=\sigma\int_{T_3/\nu}^{t}\check{x}(t,u)dW(u)\qquad\text{for all } t\ge T_3/\nu.\quad(5.4.13)$$

As usual, the deterministic term is ignored and we focus on the stochastic term ξ.

End-Point Estimate. It is a technically simple while natural attempt to use the normal one-dimensional distribution and easily derived variance of the process to deduce an estimate on the first-exit tail distribution only through observation of the end-point distribution. The variance at the end point is given through

$$\operatorname{var}x(T/\nu)=\sigma^2\int_{T_3/\nu}^{T/\nu}\check{x}^2(T/\nu,u)du$$

which, after we assure that fundamental solutions of slowly evolving systems with pointwisely identical coefficients behave virtually brave, directly shows the impact of Lemma 5.45. By that one we have the existence of a continuous mapping $c : [T_3/\nu - r, T/\nu] \to \mathbb{R}$ satisfying the initial condition

$$c(u) = c(T_3/\nu) = b(T_3/\nu)\exp\left(-c(T_3/\nu)r\right) - a(T_3/\nu) \quad \text{for all } u \in [T_3/\nu - r, T_3/\nu],$$

and the fixed-point property (5.4.2) holds true. Of course, the fixed-point property of $c(\cdot)$ has been invented to justify the transformation in (5.4.3) to take the analysis into a regime of pointwisely identical coefficients, see (5.4.4). The below schedule contains a brief reminder of the convergence of fundamental solutions in the autonomous case, and gives an outlook what implications can be carried over due to the system's small evolution speed.

- From Theorem 4.5, we know that, given some $a_0 > 0$, the fundamental solution $(\check{z}(t))_{t\in[-r,\infty)}$ corresponding to a linear autonomous delay differential law $dz(t) = -a_0 z(t)dt + a_0 z(t-r)dt$ converges to $\frac{1}{1+a_0 r}$ exponentially fast. In particular,

$$\left|\check{z}(t) - \frac{1}{1+a_0 r}\right| \leq e^{-\kappa t} \quad \text{for all } t \geq 0, \quad \kappa = \frac{|\log(1-e^{-a_0 r})|}{2r}. \tag{5.4.14}$$

- On finite time intervals, slowly varying coefficients lead to fundamental solutions that also change their behavior only slightly:
  Let $\check{X}^{(t_0)} = (\check{X}^{(t_0)}(t))_{t\in[t_0-r,T/\nu]}$ denote the nonautonomous fundamental solution, defined through (5.4.4), with start in t_0. It has pointwisely identical coefficients $\tilde{a}(t) = \tilde{b}(t)$ for all $t \in [T_3/\nu, T/\nu]$, which are defined in (5.4.2). Let further denote $\check{\mathcal{X}}^{(t_0)} = (\check{\mathcal{X}}^{(t_0)}(t))_{t\in[t_0-r,T/\nu]}$ the autonomous fundamental solution initiated in t_0 with coefficients frozen in t_0. Then

$$|\check{X}^{(t_0)}(t) - \check{\mathcal{X}}^{(t_0)}(t)| \leq 2(t-t_0)^2 \sup_{u\in[t_0,t]} \left|\frac{d\tilde{a}(u)}{du}\right| \quad \text{for all } t \in [t_0, T/\nu],$$

  see Lemma 5.49.

- Due to the first two points, for every t_0, there is $\varepsilon^{(t_0)}$ reasonably small such that the nonautonomous $\check{X}^{(t_0)}$ gets close to the point of convergence $\frac{1}{1+\tilde{a}(t_0)r}$ of its autonomous fellow $\check{\mathcal{X}}^{(t_0)}$; in particular

$$\check{X}^{(t_0)}(t) \geq \frac{1}{1+\tilde{a}(t_0)r} - \varepsilon^{(t_0)} \quad \text{for all } t \in [t_0+s_0, t_0+s_0+r], \tag{5.4.15}$$

  and the quantities s_0 and $\varepsilon^{(t_0)}$ may be chosen to be uniformly bounded,

$$s_0 \leq \hat{s} = \mathcal{O}(|\log\nu|) \quad \text{and} \quad \varepsilon^{(t_0)} \leq \varepsilon_+ = \mathcal{O}(\sqrt{\nu}|\log\nu|),$$

  see Lemma 5.50.

- Once, a segment of a solution with pointwisely identical coefficients, not necessarily autonomous, remains above a certain level, pointwisely identical and nonnegative coefficients will not change that. The details are given in Lemma 5.51.

Summarizing we will show that $\check{X}^{(t_0)}$ never falls below $\frac{1}{1+\tilde{a}(t_0)r} - \varepsilon_+$ after an initial cool-down phase of duration at most $\hat{s} = \mathcal{O}(|\log \nu|)$. Remember that

$$\frac{d\tilde{a}(t)}{dt} = \frac{da(t)}{dt} + \frac{dc(t)}{dt} \qquad \text{for all } t \in [T_3/\nu, T/\nu],$$

and with regard to Lemma 5.48 c), we conclude that there is $\tilde{m}_+$ such that

$$\sup_{t\in[T_3/\nu,T/\nu]} \left|\frac{d\tilde{a}(t)}{dt}\right| \le \frac{\tilde{m}_+\sqrt{\nu}}{2}, \tag{5.4.16}$$

and the constant $\tilde{m}_+$ is at most of order $|\log \sigma|$ independent of ν.

Lemma 5.49. *Under the Assumptions 5.44 let* $(\check{X}(t,u), u \in [T_3/\nu, T/\nu], t \in [u-r, T/\nu])$ *denote the fundamental solution of* (5.4.1). *For arbitrary* $t_0 \in [T_3/\nu, T/\nu]$ *denote* $\check{X}^{(t_0)}(t) := \check{X}(t, t_0)$ *for all* $t \in [t_0 - r, T/\nu]$, *and let* $(\check{\mathcal{X}}^{(t_0)}(t) : t \in [t_0, T/\nu])$ *be the autonomous fundamental solution initiated at* t_0, *defined as*

$$\begin{cases} d\check{\mathcal{X}}^{(t_0)}(t) = -\tilde{a}(t_0)\check{\mathcal{X}}^{(t_0)}(t)dt + \tilde{a}(t_0)\check{\mathcal{X}}^{(t_0)}(t-r)dt & \textit{for } t \in [t_0, T/\nu], \\ \check{\mathcal{X}}^{(t_0)}(t) = \mathbb{1}_{\{t_0\}}(t) & \textit{for } t \in [t_0 - r, t_0]. \end{cases}$$

And consider the deviation $Y^{(t_0)}(t) := \check{X}^{(t_0)}(t) - \check{\mathcal{X}}^{(t_0)}(t)$, $t \in [t_0 - r, T/\nu]$, *that satisfies*

$$\begin{cases} dY^{(t_0)}(t) = -\tilde{a}(t)Y^{(t_0)}(t)dt + \tilde{a}(t)Y^{(t_0)}(t-r)dt \\ \qquad\qquad - \Delta_{\tilde{a}}(t,t_0)\check{\mathcal{X}}^{(t_0)}(t)dt + \Delta_{\tilde{a}}(t,t_0)\check{\mathcal{X}}^{(t_0)}(t-r)dt & \textit{for } t \in [t_0, T/\nu], \\ Y^{(t_0)}(t) = 0 & \textit{for } t \in [t_0 - r, t_0], \end{cases}$$

where $\Delta_{\tilde{a}}(t_0, t) := \tilde{a}(t_0) - \tilde{a}(t)$ *for all* $t \in [t_0, T/\nu]$. *Then* $|Y^{(t_0)}(t)| \le \sqrt{\nu}\,\tilde{m}_+(t - t_0)^2$.

Proof. The deviation process $Y^{(t_0)}(\cdot)$ may be represented as

$$Y^{(t_0)}(t) = \int_{t_0}^{t} \check{X}(t,u)\Big(- \Delta_{\tilde{a}}(u,t_0)\check{\mathcal{X}}^{(t_0)}(u) + \Delta_{\tilde{a}}(u,t_0)\check{\mathcal{X}}^{(t_0)}(u-r)\Big)du \qquad \text{for all } t \in [t_0 - r, T/\nu].$$

And for the usual arguments, $|\check{X}^{(t_0)}(t)| \le 1$, and also $|\check{\mathcal{X}}^{(t_0)}(t)| \le 1$ for all $t \in [t_0, T/\nu]$. Then, together with the estimate (5.4.16), the claim is obvious. □

Lemma 5.50. *Under Assumptions 5.44 let* $\tilde{m}_+$ *be the constant characterized in* (5.4.16) *and let*

$$\kappa_{t_0} := \frac{|\log(1 - e^{-\tilde{a}(t_0)r})|}{2r}, \qquad \hat{\kappa} := \min_{t_0\in[T_3/\nu,T/\nu]} \{\kappa_{t_0}\}$$

For arbitary $t_0 \in [T_3/\nu, T/\nu]$, *let* s_0 *be the unique positive solution of*

$$e^{-\kappa_{t_0}(s_0+r)} = (s_0 + r)^2\sqrt{\nu}\,\tilde{m}_+. \tag{5.4.17}$$

Let further ν be sufficiently small such that

$$4\sqrt{\nu}\,\tilde{m}_+ r^2 \le \exp\left(-2\kappa_{t_0} r\right) \qquad \Big(\Rightarrow s_0 \ge r\Big), \tag{5.4.18}$$

$$4\sqrt{\nu}\,\tilde{m}_+ r^2 \le e^{-1}, \tag{5.4.19}$$

$$\frac{\sqrt{\nu}\,\tilde{m}_+}{\kappa_{t_0}} \le \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|. \tag{5.4.20}$$

we find that the solution s_0 of (5.4.17) *is bounded above through the following expression which includes the definition of $s_0^{(+)}$:*

$$s_0 + r \le s_0^{(+)} + r := \frac{1}{\kappa_{t_0}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|, \tag{5.4.21}$$

and a lower bound $s_0^{(-)}$ is determined by

$$s_0 + r \ge s_0^{(-)} + r = \frac{1}{\kappa_{t_0}} \left|\log\left(\frac{\sqrt{\nu}\,\tilde{m}_+}{\kappa_{t_0}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|\right)\right|. \tag{5.4.22}$$

Furthermore, $\varepsilon^{(t_0)}$ is uniformly bounded above in t_0 by

$$\varepsilon_+ := 2\left(\frac{\sqrt{\nu}\,\tilde{m}_+}{\hat{\kappa}}\right)^{|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)|},$$

which satisfies $\varepsilon_+ = \mathcal{O}\Big(|\log\sigma|\sqrt{\nu}\Big)$. And $s_0^{(+)}$ is uniformly bounded above in t_0 by

$$\hat{s} + r := \frac{1}{\hat{\kappa}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|,$$

which satisfies $\hat{s} = \mathcal{O}(|\log\nu|)$.

Proof. By (5.4.14), or Theorem 4.5 respectively, we have that

$$\left|\check{\mathcal{X}}^{(t_0)}(t) - \frac{1}{1 + \tilde{a}(t_0) r}\right| < e^{-\kappa_{t_0}(t-t_0)} \qquad \text{for all } t > t_0.$$

Therefore, with regard to Lemma 5.49, we know that

$$\check{X}^{(t_0)}(t) \ge \frac{1}{1 + \tilde{a}(t_0) r} - e^{-\kappa_{t_0}(t-t_0)} - \sqrt{\nu}\,\tilde{m}_+ (t - t_0)^2 \qquad \text{for all } t \in [t_0, T/\nu]. \tag{5.4.23}$$

If we understand each side of the equation in (5.4.17) as mappings in s_0, then, the left-hand side is strictly decreasing in s_0 with start in 1 while the right-hand side is strictly increasing with start in 0. Therefore, the intersection point s_0 exists and is unique over the positive half line $[0, \infty)$. Observe that through the assumption in (5.4.17) the following estimates hold true:

$$\left.\begin{aligned} e^{-\kappa_{t_0}(t-t_0)} &\le e^{-\kappa_{t_0} s_0} \\ \sqrt{\nu}\,\tilde{m}_+ (t - t_0)^2 &\le e^{-\kappa_{t_0} s_0} \end{aligned}\right\} \quad \text{for all } t \in [t_0 + s_0, t_0 + s_0 + r], \tag{5.4.24}$$

because $e^{-\kappa_{t_0}(t-t_0)}$ is decreasing in t and $(t - t_0)^2$ is increasing in t. See Figure 14 for an illustration of the idea.

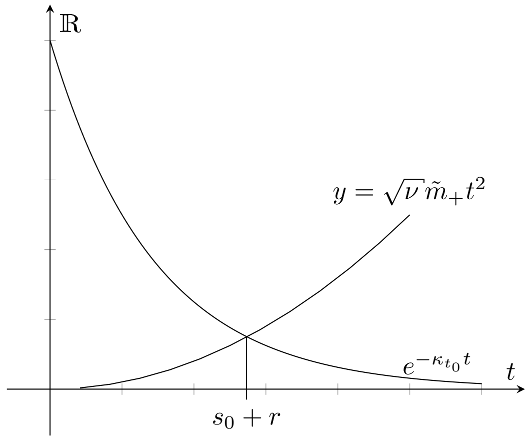

Figure 14: Illustration of the estimates (5.4.24) due to the choice in (5.4.17).

Continuing from (5.4.23) we observe that

$$\check{X}^{(t_0)}(t) \geq \frac{1}{1+\tilde{a}(t_0)r} - 2e^{-\kappa_{t_0}(t-t_0)} \qquad \text{for all } t \in [t_0 + s_0, t_0 + s_0 + r].$$

And we conveniently define

$$2e^{-\kappa_{t_0} s_0} =: \varepsilon^{(t_0)} \text{ for all } t_0 \geq T_3/\nu \text{ such that } t_0 + s_0 \leq T/\nu. \tag{5.4.25}$$

Assumption (5.4.18) provides that the intersection point s_0 must be greater or equal to r. But then, it must be smaller than the intersection point of the left-hand side of (5.4.17) with the constant niveau $2\sqrt{\nu}\,\tilde{m}_+ r$, where we plugged in r for s_0 on the right-hand side of the equation. That reveals that

$$s_0 + r \leq s_0^{(+)} + r = \frac{1}{\kappa_{t_0}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|.$$

But then, s_0 must be greater or equal than the right-hand side with $s_0^{(+)}$ plugged into it;

$$\begin{aligned} s_0 \geq s_0^{(-)} &= \frac{1}{\kappa_{t_0}} \left|\log(\sqrt{\nu}\,\tilde{m}_+ (s^{(+)} + r)^2)\right| \\ &= \frac{1}{\kappa_{t_0}} \left|\log\left(\frac{\sqrt{\nu}\,\tilde{m}_+}{\kappa_{t_0}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|\right)\right|. \end{aligned}$$

Therefore, we find an upper bound for $\varepsilon^{(t_0)}$ by plugging $s_0^{(-)}$ into the definition in (5.4.25). That provides that

$$\varepsilon^{(t_0)} \leq 2\exp\left(-\kappa_{t_0} \frac{1}{\kappa_{t_0}} \left|\log\left(\frac{\sqrt{\nu}\,\tilde{m}_+}{\kappa_{t_0}} \left|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)\right|\right)\right|\right) = 2\left(\frac{\sqrt{\nu}\,\tilde{m}_+}{\kappa_{t_0}}\right)^{|\log(4\sqrt{\nu}\,\tilde{m}_+ r^2)|}$$

The claimed form and order of ε_+ follows from the assumption (5.4.19) acting as a minimal condition on the exponent in the above estimate. The claim concerning the uniformly upper bound $\hat{s}$ is obvious.

□

Lemma 5.51. *Let* $a : [t_0, \infty) \to (0, \infty)$ *and consider the solution* $(x(t))_{t \geq t_0}$ *of*

$$\begin{cases} dx(t) = -a(t)x(t)dt + a(t)x(t-r)dt & \text{for all } t \in [t_0, \infty), \\ x_{t_0} = \Upsilon \in \mathcal{C}(J, [l_1, l_2]). \end{cases} \tag{5.4.26}$$

Then, the bound holds for all times subsequent to t_0*, i.e.*

$$x(t_0 + u) \in [l_1, l_2] \quad \text{for all } u \geq -r.$$

Remark 5.52. *The multiply used contradiction argument, that before has shown the boundedness of critical-regime fundamental solutions by* 1*, works just as well in the opposite direction. Since the argument is rather standard by now, we slightly modify it and cover both contradictions almost at once:*

Proof of Lemma 5.51. Consider the deterministic first-exit time from the interval $[l_1, l_2]$ after t_0, defined as

$$\tau_{[l_1,l_2]} := \inf\{t \geq t_0 : x(t) \notin [l_1, l_2]\}.$$

For the purpose of a contradiction, we assume $\tau_{[l_1,l_2]}$ to be finite. By absolute continuity of the solution path, $\tau_{[l_1,l_2]} > 0$ we know that there is an $\varepsilon > 0$ and an interval $N_\varepsilon = (\tau_{[l_1,l_2]}, \tau_{[l_1,l_2]} + \varepsilon)$ such that

$$x(t) \notin [l_1, l_2] \quad \text{for all } t \in N_\varepsilon.$$

Choose an arbitrary $t_1 \in N_\varepsilon$, then

$$x(t_1) = x(\tau_{[l_1,l_2]})e^{-\alpha(t_1,\tau_{[l_1,l_2]})} + e^{-\alpha(t_1)} \int_{\tau_{[l_1,l_2]}}^{t_1} e^{\alpha(u)} a(u)x(u-r)du,$$

where again $\alpha(t, s) = \int_s^t a(u)du$ and $\alpha(t) = \alpha(t, 0)$ for all $s, t \in [0, \infty)$. As $\int_s^t a(u)e^{\alpha(u)}du = e^{\alpha(t)} - e^{\alpha(s)}$, together with the initial condition (5.4.26), we obtain on the one hand,

$$\begin{aligned} x(t_1) &\leq l_2 e^{-\alpha(t_1,\tau_{[l_1,l_2]})} + e^{-\alpha(t_1)} \int_{\tau_{[l_1,l_2]}}^{t_1} e^{\alpha(u)} a(u) l_2 du \\ &= l_2 e^{-\alpha(t_1,\tau_{[l_1,l_2]})} + l_2 e^{-\alpha(t_1)} \left(e^{\alpha(t_1)} - e^{\alpha(\tau_{[l_1,l_2]})} \right) = l_2. \end{aligned}$$

And on the other hand,

$$\begin{aligned} x(t_1) &\geq l_1 e^{-\alpha(t_1,\tau_{[l_1,l_2]}} + e^{-\alpha(t_1)} \int_{\tau_{[l_1,l_2]}}^{t_1} e^{\alpha(u)} a(u) l_1 du \\ &= l_1 e^{-\alpha(t_1,\tau_{[l_1,l_2]}} + l_1 e^{-\alpha(t_1)} \left(e^{\alpha(t_1)} - e^{\alpha(\tau_{[l_1,l_2]})} \right) = l_1. \end{aligned}$$

Which settles the contradiction to $\tau_{[l_1,l_2]} < \infty$, and the proof is done. □

The following theorem constitutes the main result of this section and its content summarizes the result we have achieved through Lemmas 5.49, 5.50 and 5.51.

Theorem 5.53. *Consider the situation of Lemma 5.50, and let* $\check{x}$ *denote the fundamental solution of* (5.4.1) *and* $\check{X}(t, u)$ *be defined as in* (5.4.3)*, where we denote* $c : [T_3/\nu - r, T/\nu]$

the unique continuation of the constant initial segment $c_\star$ with $c_\star$ is as in Lemma 5.48. Then, if ν and $\nu|\log\sigma|$ are sufficiently small,

$$\check{X}(t,t_0) \geq \frac{1}{1+\tilde{a}(t_0)r} - \varepsilon_+ \quad \text{for } t \geq t_0 + \hat{s} \text{ and } t_0 \text{ such that } t_0 + \hat{s} \leq T_3/\nu.$$

Therefore,

$$\int_{T_3/\nu}^{T/\nu} \exp\Big(2\gamma(T/\nu,u)\Big)\check{X}^2(T/\nu,u)du \geq \int_{T_3/\nu}^{T/\nu-\hat{s}} \exp\Big(2\gamma(T/\nu,u)\Big)\left(\frac{1}{1+\tilde{a}(u)r} - \varepsilon_+\right)^2 du.$$

And a lower boundary is given through

$$\frac{\operatorname{var} x(T/\nu)}{\sigma^2} \geq \left(\frac{1}{1+(a(T/\nu)+c_\star)r} - \varepsilon_+\right)^2 \exp\left(2c_\star\frac{T-T_3}{\nu}\right)\left(1-\exp\left(-2c_\star\frac{T-T_3-\nu\hat{t}}{\nu}\right)\right).$$

One more representation of this estimate:

$$\operatorname{var} x(T/\nu) \geq \left(\frac{\sigma}{1+a_-r+c_\star r}\right)^2 \exp\left(2c_\star\frac{T-T_3}{\nu}\right)(1-\varepsilon_1)^2(1-\varepsilon_2),$$

where $\varepsilon_1 = (1+(a(T/\nu)+c_\star)r)\varepsilon_+ = \mathcal{O}(\sqrt{\nu}|\log\sigma|)$, $\varepsilon_2 = \exp\left(-2c_\star\frac{T-T_3-\nu\hat{t}}{\nu}\right)$.

Proof. All of which has been shown in advance. □

As a centered normal distribution with standard deviation $\Sigma > 0$ aggregates most of its mass outside $[-\beta,\beta]$, i.e.

$$\mathcal{N}_{0,\Sigma^2}\{[-\beta,\beta]^c\} \geq 1 - \frac{2\beta}{\sqrt{2\pi}}\Sigma^{-1},$$

if $\Sigma \gg \beta$, the previous theorem implies that

$$\mathbb{P}\left\{\sup_{s\in[T_3/\nu,T/\nu]}|x(s)| < \beta\right\} = \mathcal{O}\left(\frac{\beta}{\sqrt{\operatorname{var} x_{T/\nu}}}\right),$$

which is helpful, if

$$\sqrt{\operatorname{var} x(T/\nu)} > \beta \quad \Leftrightarrow \quad \frac{T-T_3}{\nu} > \frac{1}{c_\star}\log\left(\frac{\beta}{\sigma^2}\big(1+(a(T/\nu)+c_\star)r\big)\right).$$

Remark 5.54.

- *In order to observe an escape from an environment of diameter β of order 1 over $[T_3/\nu, T/\nu]$ it suffices to have $T-T_3$ of order $\sqrt{\nu}|\log\sigma|$. In particular that justifies the fourth item in the Assumptions 5.44.*
- *Small-Ball-Probability Approach. To make use of the small-ball probabilities of Brownian motion as we have in Subsection 4.4, we an improved understanding of the fixed point c seems necessary. Having achieved that, a procedure may be accomplished that generalizes the one, we have seen in the autonomous case.*

A. Auxiliaries

A.1. A Concentration Result for linear SDEs

The following presentation follows [BG06, Section 3.1]. Let $\nu > 0$ be some small parameter. Consider continuous differentiable $\tilde{a} : [0,T] \to [a_-, a_+]$ where $0 < a_- < a_+$. Denote $a(t) = \tilde{a}(\nu t)$ for all $t \in [0, T/\nu]$. Then, the solution of the SDE $dy(t) = -\tilde{a}(\nu t)y(t)dt + \sigma dW(t)$ for $t \in [0, T/\nu]$ and $y(0) = 0$ is $\mathbb{P}$-almost surely given by

$$y(t) = \sigma \int_0^t e^{-\alpha(t,u)} dW(u) \quad \text{for } t \in [0, T/\nu],$$

where $\alpha(t,s) = \alpha(t) - \alpha(s)$ with $\alpha(t) := \int_0^t a(u)du$ for $s, t \in [0, T/\nu]$, $s < t$. The differential law may be formulated in *fast time* $t = s/\nu$ as

$$d\tilde{y}(s) = -\frac{1}{\nu}\tilde{a}(s)\tilde{y}(s)ds + \frac{\sigma}{\sqrt{\nu}} d\tilde{W}(s) \quad \text{for } s \in [0, T],$$

where $\tilde{W}$ is again a Brownian motion. The according rescaled variance process $\tilde{v}(\cdot) := \frac{1}{\sigma^2} \operatorname{var} \tilde{y}(\cdot)$ solves $\frac{\varepsilon d\tilde{v}(s)}{ds} = -2\tilde{a}(s)\tilde{v}(s) + 1$. That differential law features the equilibrium branch $\tilde{v}^\star(s) = \frac{1}{2\tilde{a}(s)}$ and by the uniform stability property $\tilde{a}(\cdot) > a_-$, there is an adiabatic solution path $\tilde{v}^{\mathrm{ad}_\nu}$ that solves the differential and satisfies $\|\tilde{v}^{\mathrm{ad}_\nu} - \tilde{v}^\star\| \in \mathcal{O}(\nu)$. Retranslation into the slow-time setting provides the existence of an adiabatic solution ζ, given by $\zeta(t) = \tilde{v}^{\mathrm{ad}_\nu}(\nu t)$ for $t \in [0, T/\nu]$ with $\|\zeta - v^\star\| \in \mathcal{O}(\nu)$, where $v^\star(t) = \frac{1}{2a(t)} = \frac{1}{2\tilde{a}(\nu t)} = \tilde{v}^\star(\nu t)$ for $t \in [0, T/\nu]$. In this situation, we have the following concentration inequality formulated in slow time, i.e. when ν is small:

$$\mathbb{P}\left\{ \sup_{t\in[0,T/\nu]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta \right\} \leq \frac{2eT\beta^2\big(1+\mathcal{O}(\nu)\big)}{\sigma^2 \nu \alpha(T/\nu)} \exp\left(-\frac{\beta^2}{2\sigma^2}\right) \quad \text{for } \beta > 0,$$

where integer-value restrictions are ignored.

Proof. Consider a partition $0 = t_0 < t_1 < \ldots < t_n = T/\nu$ with step sizes defined by $\alpha(t_i, t_{i-1}) = \gamma$ for all $i \in \{1, \ldots, n\}$ and some arbitrary $\gamma > 0$. Then, the number of steps is given by $n = \frac{T}{\gamma\alpha(T/\nu)\nu}$ if we ignore integer-value restrictions. Then,

$$\mathbb{P}\left\{ \sup_{t\in[0,T/\nu]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta \right\} \leq \sum_{i=1}^n \mathbb{P}\left\{ \sup_{t\in[t_{i-1},t_i]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta \right\} \quad \text{for all } \beta > 0. \tag{A.1.1}$$

Through appropriate estimates, one can isolate the martingale parts on the right-hand side in order to apply the Bernstein-type inequality, [BG06, Appendix B.1]. For every $i \in \{1, \ldots, n\}$ we obtain that

$$\begin{aligned}
\mathbb{P}\left\{ \sup_{t\in[t_{i-1},t_i]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta \right\} &\leq \mathbb{P}\left\{ \sup_{t\in[t_{i-1},t_i]} \left| \int_0^t \sigma e^{\alpha(u)} dW(u) \right| > \beta \inf_{t\in[t_{i-1},t_i]} \sqrt{\zeta(t)}\, e^{\alpha(t)} \right\} \\
&\leq 2\exp\left(-\frac{\beta^2}{2\sigma^2 \int_0^{t_i} e^{2\alpha(u)}du} \inf_{t\in[t_{i-1},t_i]} \zeta(t) \inf_{t\in[t_{i-1},t_i]} e^{2\alpha(t)} \right) \\
&= 2\exp\left(-\frac{\beta^2}{2\sigma^2 \int_0^{t_i} e^{-2\alpha(t_i,u)}du} \inf_{t\in[t_{i-1},t_i]} \zeta(t) \inf_{t\in[t_{i-1},t_i]} e^{-2\alpha(t_i,t)} \right).
\end{aligned}$$

Then, we observe that

$$\int_0^{t_i} e^{-2\alpha(t_i,u)}du = \frac{1}{\sigma^2}\operatorname{var} y(t_i) \quad \text{for all } i \in \{1,\dots,n\}.$$

Just like we have seen above, the rescaled variance process $v(\cdot) = \frac{1}{\sigma^2}\operatorname{var} y(\cdot)$ solves the differential equation $dv(t) = -2a(t)v(t)dt + 1$ for all $t \in [0, T/\nu]$. The process $\zeta(\cdot)$ solves the same differential equation and $\zeta(0) > v(0) = 0$ for sufficiently small ν. As solution paths must not intersect, it is generally true that $v(t) \le \zeta(t)$ for all $t \in [0, T/\nu]$. In other words,

$$\int_0^{t_i} e^{-2\alpha(t_i,u)}du \le \zeta(t_i) \quad \text{for all } i \in \{1,\dots,n\}.$$

Hence,

$$\mathbb{P}\left\{\sup_{t\in[t_{i-1},t_i]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta\right\} \le 2\exp\left(-\frac{\beta^2}{2\sigma^2}\inf_{t\in[t_{i-1},t_i]}\frac{\zeta(t)}{\zeta(t_i)}e^{-2\alpha(t_i,t_{i-1})}\right) \text{ for all } i \in \{1,\dots,n\}.$$

And as ζ varies slowly, is bounded above by $\frac{1}{2a_- + \mathcal{O}(\nu)}$, and below by $\frac{1}{2a_+ + \mathcal{O}(\nu)}$, we obtain that

$$\mathbb{P}\left\{\sup_{t\in[t_{i-1},t_i]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta\right\} \le 2\exp\left(-\frac{\beta^2}{2\sigma^2}e^{-2\alpha(t_i,t_{i-1})}\big(1+\mathcal{O}(\nu t_\Delta)\big)\right) \text{ for all } i \in \{1,\dots,n\}.$$

Using that $e^{-x} \ge 1 - x$ for all $x \in \mathbb{R}$, and continuing from (A.1.1) shows that

$$\begin{aligned}\mathbb{P}\left\{\sup_{t\in[0,T/\nu]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta\right\} &\le 2\frac{T}{\gamma\nu\alpha(T/\nu)}\exp\left(-\frac{\beta^2}{2\sigma^2}e^{-2\gamma}(1+\mathcal{O}(\nu t_\Delta))\right)\\ &\le 2\frac{T}{\gamma\nu\alpha(T/\nu)}\exp\left(-\frac{\beta^2}{2\sigma^2}(-2\gamma)\big(1+\mathcal{O}(\nu t_\Delta)\big)\right)\exp\left(-\frac{\beta^2}{2\sigma^2}\right).\end{aligned}$$

Optimization over γ leads to the choice $\gamma = \frac{\sigma^2}{\beta^2(1+\mathcal{O}(\nu))}$ and thus,

$$\mathbb{P}\left\{\sup_{t\in[0,T/\nu]} \frac{|y(t)|}{\sqrt{\zeta(t)}} > \beta\right\} \le \frac{2T\beta^2\big(1+\mathcal{O}(\nu)\big)}{\sigma^2\nu\alpha(T/\nu)}\exp(1)\exp\left(-\frac{\beta^2}{2\sigma^2}\right).$$

□

A.2. Estimates for Q-integrals

Theorem A.1. *For arbitrary $\alpha, \gamma > 0$ and $p \in \mathbb{N}, p \ge 2$, we have that*

$$\int_\alpha^\infty p^{-\gamma u^2}du \le \frac{p^{-\alpha^2\gamma}}{2\alpha\gamma\log p}. \tag{A.2.1}$$

Proof. With a substitution $v = \sqrt{2\gamma\log p}\,u \Leftrightarrow u = \frac{v}{\sqrt{2\gamma\log p}}$ and so,

$$\begin{aligned}\int_\alpha^\infty p^{-\gamma u^2}du &= \int_\alpha^\infty \exp\left(-\frac{u^2}{2}2\gamma\log p\right)du\\ &= \int_{\alpha\sqrt{2\gamma\log p}} \exp\left(-\frac{v^2}{2}\right)\frac{1}{\sqrt{2\gamma\log p}}dv\end{aligned}$$

Then, with an application of a tail estimate, we find that

$$\int_{\alpha\sqrt{2\gamma\log p}} \exp\left(-\frac{v^2}{2}\right)\frac{dv}{\sqrt{2\gamma\log p}} \leq \frac{1}{2\alpha\gamma\log p}\exp\left(-\frac{2\alpha^2\gamma\log p}{2}\right) = \frac{p^{-\alpha^2\gamma}}{2\alpha\gamma\log p}.$$

□

Example A.2. • *For $\alpha = 1, \gamma = \frac{1}{2}$, the result reads*

$$\int_1^\infty p^{-\frac{u^2}{2}}du \leq \frac{1}{\sqrt{p}\log p}.$$

• *For $\alpha = 1, \gamma = 1$,*

$$\int_1^\infty p^{-u^2}du \leq \frac{1}{2p\log p}.$$

Lemma A.3. *In the situation of Example 3.2 b) the fundamental (matrix) solution $(\check{x}(t,s) : s \in [t_0, T], t \in [s, T])$ is (locally) Lipschitz continuous in both arguments.*

Proof. Lipschitz continuity in the first argument is clear as we already mentioned before, so it remains to show only that the fundamental solution is Lipschitz continuous in the second argument. We assume that for $q > 0$ we have that

$$\max_{i\in\{1,\dots,N\}}\sup_{t\in[t_0,T]}\left|\frac{d}{dt}A_i(t)\right| \leq q < \infty,$$
$$\sup_{t\in[t_0,T]}\sup_{u\in[-r,0]}\left|\frac{d}{dt}A(t,u)\right| \leq q < \infty.$$

It is a crucial point to note that

$$\left|\int_{-r}^0 \eta(t,u) - \eta(t',u)du\right| \leq qr(t-t') + Nq(t-t') \qquad \text{for all } u \in [-r,0], t, t' \in [0,T].$$

Let us fix some arbitrary $s \in [0,T], t \in [s,T]$ to simplify quantifications. In order to deduce an estimate on the resolvent for fixed first argument, an application of the resolvent equation (3.3.7) shows that for an appropriately small $\Delta > 0$ (such that $s - \Delta \geq t_0$) we find that

$$\begin{aligned} R(t,s) - R(t,s-\Delta) = &-(\eta(t,s-t) - \eta(t,s-t-\Delta)) \\ &+ \int_s^t R(t,u)(\eta(u,s-u) - \eta(u,s-u-\Delta))du \\ &- \int_{s-\Delta}^s R(t,u)(\eta(u,s-\Delta-u))du \qquad \text{for all } s \in [0,T], t \in [s,T]. \end{aligned}$$

Because of the boundedness of $R(t,u)$ in (3.3.8) and of η, there is a constant $C_1 > 0$ such that

$$\left|\int_{s-\Delta}^s R(t,u)(\eta(u,s-\Delta-u))du\right| \leq C_1\Delta.$$

We define the two mappings $d, D : \{(t,s) \in [t_0, T]^2 : s \leq t\} \to \mathbb{R}$ through

$$d(t,s) := |\eta(t, s-t) - \eta(t, s-t-\Delta)| \quad \text{and} \quad D(t,s) = \int_s^t d(u,s)du.$$

Then due to the representation of η in (3.3.17),

$$D(t,s) \leq \int_s^t \left| \int_{s-u}^0 A(u,v)dv - \int_{s-u-\Delta}^0 A(u,v)dv \right| + \left| \sum_{i=1}^N A_i(u)(\mathbb{1}_{\{s-u\leq r_i\}} - \mathbb{1}_{\{s-u-\Delta\leq r_i\}}) \right| du$$

If we let

$$A_+ := \sup_{t\in[t_0,T]} \sup_{v\in[-r,0]} |A(t,v)|, \tag{A.2.2}$$

we find that

$$\int_s^t \left| \int_{s-u}^0 A(u,v)dv - \int_{s-u-\Delta}^0 A(u,v)dv \right| du \leq A_+\Delta(t-s).$$

For the second part of $D(t,s)$ we introduce the following notation for an upper boundary of the jump height,

$$B_+ := \max_{i\in\{1,\dots,N\}} \sup_{t\in[t_0,T]} |A_i(t)|.$$

It is then helpful to realize that

$$|A_i(u)(\mathbb{1}_{\{s-u\leq r_i\}} - \mathbb{1}_{\{s-u-\Delta\leq r_i\}})| \leq \begin{cases} 0 & \text{for } s-u \leq -r_i, \\ B_+ & \text{for } s-u-\Delta \leq -r_i \leq s-u, \\ 0 & \text{for } s-u-\Delta > -r_i. \end{cases}$$

Therefore, we find that

$$\int_s^t \left| \sum_{i=1}^N A_i(u)(\mathbb{1}_{\{s-u\leq r_i\}} - \mathbb{1}_{\{s-u-\Delta\leq r_i\}}) \right| du \leq NB_+\Delta.$$

And summarizing what we have achieved,

$$D(t,s) \leq \Delta(A_+T + NB_+) \tag{A.2.3}$$

By carefully going through the argument, we realize that the same arguments work with few modifications also in case $\Delta < 0$. And finally, regarding the definition of the fundamental solution in (3.3.9), we find that

$$\check{x}(t,s) - \check{x}(t,s-\Delta) = \int_s^t R(u,s) - R(u,s-\Delta)du - \int_{s-\Delta}^s R(u,s-\Delta)du \tag{A.2.4}$$

$$\text{for all } s,t \in [t_0,T],\ s \leq t, \tag{A.2.5}$$

which implies that for every time horizon $T > t_0$, there is a constant $C = C(T)$ such that

$$|\check{x}(t,s) - \check{x}(t,s-\Delta)| \leq D(t,s) + \int_s^t c_R D(u,s) + C_1|\Delta| ds + c_R\Delta \leq C|\Delta|, \tag{A.2.6}$$

where c_R satisfies $|R(t,s)| \leq c_R$ over $[t_0, T]$ which is due to (3.3.8). □

A.3. Brownian First-Exit Distribution - Lower Tail Estimates

A.3.1. First Approach to Small-Ball Probabilities

In the book of Revuz and Yor [RY05], for $\tau := \inf\{t \geq 0 : W(t) \notin (-l, r)\}$ we find that

$$\mathbb{E} \exp\left(\frac{\gamma^2}{2}\tau\right) = \frac{\cos\left(\frac{1}{2}\gamma(r-l)\right)}{\cos\left(\frac{1}{2}\gamma(r+l)\right)} \quad \text{for all } \gamma \in \left[0, \frac{\pi}{l+r}\right). \tag{A.3.1}$$

In case $l = r$ with $\tau_r := \inf\{t \geq 0 : |W(t)| \geq r\}$ that simplifies to

$$\mathbb{E} \exp\left(\frac{\gamma^2}{2}\tau_r\right) = \frac{1}{\cos(\gamma r)} \quad \text{for all } \gamma \in \left[0, \frac{\pi}{2r}\right).$$

Therefore, we may deduce by means of the Markov inequality that

$$\mathbb{P}\left\{\sup_{s\in[0,T]} |W(s)| < \beta\right\} = \mathbb{P}\{\tau_\beta > T\} = \mathbb{P}\left\{e^{\frac{\gamma^2}{2}\tau_\beta} > e^{\frac{\gamma^2}{2}T}\right\} \leq \frac{e^{-\frac{\gamma^2}{2}T}}{\cos(\gamma\beta)} \quad \text{for all } \gamma \in \left[0, \frac{\pi}{2\beta}\right).$$

This estimate is only useful when T is at least of order β^2.

A.3.2. Small-Ball Estimates

Based on the result from [CT62] N. Berglund and B. Gentz provide through [BG06, Corollary C.2.2] in case $d = 1$ the following small-ball estimate for a Brownian Motion $(W(s))_{s\geq 0}$.

Corollary A.4. *For any $r > 0$,*

$$\mathbb{P}\left\{\sup_{s\in[0,1]} |W(s)| < r\right\} \leq \frac{4}{\pi} \exp\left(-\frac{\pi^2}{8r^2}\right).$$

Rescaling $r = \frac{\delta(1+ar)}{\sqrt{T}}$ yields the very result we desire when comparing SDDEs first-exit time behavior with the one from properly rescaled Brownian motion in Section 4.4:

$$\mathbb{P}\left\{\sup_{s\in[0,T]} \frac{|W(s)|}{1+ar} < \delta\right\} \leq \frac{4}{\pi} \exp\left(-\frac{\pi^2}{8\delta^2(1+ar)^2}T\right) \quad \text{for all } T > 0, \delta > 0. \tag{A.3.2}$$

B. SDDEs - Case Studies

This section provides the formal verification of the the concentration inequalities that were discussed in section 4.3. Let us briefly remember the form of the Fernique inequality in the regarding section:

$$\mathbb{P}\left\{\sup_{s\in[0,T]}|y(s)|\geq h\Big(\sqrt{\|\Gamma\|}+Q(p,T)\Big)\right\}\leq\frac{5}{2}p^{2n}e^{-\frac{h^2}{2}}\qquad\text{for } h>\sqrt{1+4\log p},$$

where

$$\frac{\|\Gamma\|}{\sigma}=\sup_{t\in[0,T]}\mathbb{E}\left[\left(\int_0^t\check{x}(t-u)e^{\mu u}dW(u)\right)^2\right]=\sup_{t\in[0,T]}\int_0^t\check{x}^2(t-u)e^{2\mu u}du$$

where $Q(T)=Q(p,T)\leq\mathcal{Q}_1+\mathcal{Q}_2$ with $\mathcal{Q}_1=\mathcal{Q}_1(p,T)$ and $\mathcal{Q}_2=\mathcal{Q}_2(p,T)$, defined as in (3.4.10),(3.4.11):

$$\mathcal{Q}_1(p,T):=\int_1^\infty\sqrt{\sup_{\substack{s,t\in[0,T],s<t,\\|t-s|\leq Tp^{-u^2}}}\int_0^s\sigma^2(v)\Big(\check{x}(t-v)-\check{x}(s-v)\Big)^2dv}\,du,$$

$$\mathcal{Q}_2(p,T):=\int_1^\infty\sqrt{\sup_{\substack{s,t\in[0,T],s<t,\\|t-s|\leq Tp^{-u^2}}}\int_s^t\sigma^2(v)\check{x}^2(t-v)dv}\,du.$$

B.1. Critical Regime

In addition to the special-case analysis of Theorem (4.18), a couple of scenarios of values of μ and κ have been discussed for the critical regime. We keep the notations of section (4.3), and in particular let κ be given as in (4.1.12). For the most part, the derived estimates are based upon the convergence of fundamental solutions.

Critical regime, non-neutralizing case ($a=b>0$, $\mu\notin\{0,-\kappa,-\frac{\kappa}{2}\}$). In the situation where $\mu\notin\{0,-\kappa,-\frac{\kappa}{2}\}$, we consider the SDDE (4.3.3) for $a=b>0$ and $\mu\in\mathbb{R}\backslash\{0\}$. Then for all $s,t\in[0,T]$, $s\leq t$,

$$\mathbb{E}\left[\left(\int_s^t\check{x}(t-u)e^{\mu u}dW(u)\right)^2\right]\leq\int_s^t e^{2\mu u}du\leq\frac{e^{2\mu s}}{2\mu}\left(e^{2\mu(t-s)}-1\right).\tag{B.1.1}$$

$$\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du = \int_0^s \left(\int_s^t (-a)\check{x}(v-u)+a\check{x}(v-u-r)dv\right)^2 e^{2\mu u}du$$
$$\le a^2\int_0^s \left(\int_s^t e^{-\kappa(v-u)}+e^{-\kappa(v-u-r)}dv\right)^2 e^{2\mu u}du$$
$$\le a^2(1+e^{\kappa r})^2\int_0^s \left(\int_s^t e^{-\kappa v}dv\right)^2 e^{2\mu u+2\kappa u}du$$
$$= a^2(1+e^{\kappa r})^2 e^{-2\kappa s}\left(\frac{1-e^{-\kappa(t-s)}}{\kappa}\right)^2\int_0^s e^{2(\kappa+\mu)u}du$$
$$= a^2(1+e^{\kappa r})^2 e^{-2\kappa s}(t-s)^2\frac{e^{2(\kappa+\mu)s}-1}{2(\kappa+\mu)}$$
$$= a^2(1+e^{\kappa r})^2 e^{-2\kappa s}\frac{e^{2(\kappa+\mu)s}-1}{2(\kappa+\mu)}(t-s)^2. \tag{B.1.2}$$

$$\frac{\|\Gamma\|}{\sigma^2} = \sup_{t\in[0,T]}\mathbb{E}\left[\left(\int_0^t \check{x}(t-u)e^{\mu u}dW_u\right)^2\right] \le \sup_{t\in[0,T]}\int_0^t\left(\frac{1}{1+ar}+e^{-\kappa(t-u)}\right)^2 e^{2\mu u}du$$
$$= \sup_{t\in[0,T]}\frac{1}{(1+ar)^2}\frac{e^{2\mu t}-1}{2\mu}+e^{-\kappa t}\frac{2}{(1+ar)}\frac{e^{(2\mu+\kappa)t}-1}{(2\mu+\kappa)}+e^{-2\kappa t}\frac{e^{2(\mu+\kappa)t}-1}{2(\mu+\kappa)}.$$

Vanishing noise ($a=b>0$, $\mu<0$, $\mu\notin\{0,-\kappa,-\frac{\kappa}{2}\}$).

$$\frac{\|\Gamma\|}{\sigma^2} \le \sup_{t\in[0,T]}\frac{1}{(1+ar)^2}\frac{e^{2\mu t}-1}{2\mu}+e^{-\kappa t}\frac{2}{(1+ar)}\frac{e^{(2\mu+\kappa)t}-1}{(2\mu+\kappa)}+e^{-2\kappa t}\frac{e^{2(\mu+\kappa)t}-1}{2(\mu+\kappa)}$$
$$= \sup_{t\in[0,T]}\frac{1}{(1+ar)^2}\frac{e^{2\mu t}-1}{2\mu}+\frac{2}{(1+ar)}\frac{e^{2\mu t}-e^{-\kappa t}}{(2\mu+\kappa)}+\frac{e^{2\mu t}-e^{-2\kappa t}}{2(\mu+\kappa)}$$
$$\le \frac{1}{2|\mu|(1+ar)^2}+\frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|}+\frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|\}}\right).$$

$$\frac{\|\Gamma\|}{\sigma^2} \le \int_0^T e^{2\mu u}du \le \frac{1}{|2\mu|}.$$

So, we introduce the notation

$$v_0^2 := \max\left\{\frac{1}{|2\mu|},\frac{1}{(1+ar)^2 2|\mu|}+\frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|}+\frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)\right\}. \tag{4.3.12}$$

With (B.1.1)

$$\mathbb{E}\left[\left(\int_s^t \check{x}(t-u)e^{\mu u}dW(u)\right)^2\right] \le e^{2\mu s}(t-s) \le t-s, \tag{B.1.3}$$

and starting from (B.1.2), we find out that

$$\begin{aligned}
\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du &\leq a^2(1+e^{\kappa r})^2 e^{-2\kappa s} \frac{\left(1-e^{-\kappa(t-s)}\right)^2}{\kappa^2} \frac{e^{2(\kappa+\mu)s}-1}{2(\kappa+\mu)} \\
&\leq a^2(1+e^{\kappa r})^2 \underbrace{\frac{e^{2\mu s}-e^{-2\kappa s}}{2(\kappa+\mu)}}_{\leq \left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)\frac{1}{2|\kappa+\mu|}} (t-s)^2 \\
&\leq \frac{a^2(1+e^{\kappa r})^2}{2|\kappa+\mu|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)(t-s)^2. \qquad \text{(B.1.4)}
\end{aligned}$$

Collecting the results,

$$\begin{aligned}
\frac{\mathcal{Q}_1}{\sigma} &\leq \int_1^\infty \sqrt{Tp^{-u^2}}\,du \leq \frac{\sqrt{T}}{\sqrt{p}\log p}, \\
\frac{\mathcal{Q}_2}{\sigma} &\leq \frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}}\sqrt{1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}}\,\frac{T}{p\log p}.
\end{aligned}$$

And therefore, with v_0 from (4.3.12),

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_0 + \frac{7}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}}\sqrt{1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}}\,\frac{T}{p\log p}.$$

Increasing noise ($a = b > 0$, $\mu > 0$, $\mu \notin \{0, -\kappa, -\frac{\kappa}{2}\}$).

$$\frac{\|\Gamma\|}{\sigma^2} = \sup_{t\in[0,T]} \int_0^t \check{x}^2(t-u)e^{2\mu u}du \leq \frac{e^{2\mu T}-1}{2\mu}.$$

$$\int_s^t \check{x}^2(t-u)e^{2\mu u}du \leq \int_s^t e^{2\mu u}du \leq e^{2\mu t}(t-s).$$

With (B.1.2), we obtain

$$\begin{aligned}
\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du &\leq a^2(1+e^{\kappa r})^2 e^{-2\kappa s}\frac{e^{2(\kappa+\mu)s}-1}{2(\kappa+\mu)}(t-s)^2 \\
&= a^2(1+e^{\kappa r})^2 e^{2\mu s}\frac{1-e^{-2(\kappa+\mu)s}}{2(\kappa+\mu)}(t-s)^2 \\
&\leq \frac{a^2(1+e^{\kappa r})^2}{2(\kappa+\mu)} e^{2\mu T}(t-s)^2.
\end{aligned}$$

$$\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du \leq \int_0^s e^{2\mu u} du \leq \frac{e^{2\mu s}-1}{2\mu} \leq \frac{e^{2\mu T}}{2\mu}.$$

Collecting the results,

$$\frac{\mathcal{Q}_1}{\sigma} \le \int_1^\infty \sqrt{e^{2\mu T} T p^{-u^2}} du \le \frac{\sqrt{T} e^{\mu T}}{\sqrt{p}\log p},$$
$$\frac{\mathcal{Q}_2}{\sigma} \le \int_1^\infty \sqrt{\frac{a^2(1+e^{\kappa r})^2}{2(\kappa+\mu)} e^{2\mu T} T^2 p^{-2u^2}} du \le \frac{a(1+e^{\kappa r})}{\sqrt{2(\kappa+\mu)}} \frac{T e^{\mu T}}{2p\log p}.$$

And therefore,

$$\begin{aligned}\frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} &\le \frac{e^{\mu T}}{\sqrt{2\mu}} + \frac{7}{2}\frac{\sqrt{T}e^{\mu T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2(\kappa+\mu)}}\frac{Te^{\mu T}}{2p\log p} \\ &= \frac{e^{\mu T}}{\sqrt{2\mu}}\left(1 + \frac{7\sqrt{2\mu}}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}a(1+e^{\kappa r})\frac{\sqrt{2\mu}}{\sqrt{2(\kappa+\mu)}}\frac{T}{2p\log p}\right).\end{aligned}$$

Critical regime, vanishing noise, parameter-neutralization issues ($a = b > 0$, $\mu < 0$, $\mu \in \{0, -\kappa, -\frac{\kappa}{2}\}$). *Case #1:* $\mu + \kappa = 0$ We begin with the case $\mu + \kappa = 0$ and $2\mu + \kappa < 0$. For $\|\Gamma\|$ we find that

$$\begin{aligned}\frac{\|\Gamma\|}{\sigma^2} &= \sup_{t\in[0,T]} \mathbb{E}\left[\left(\int_0^t \check{x}(t-u)e^{\mu u} dW_u\right)^2\right] \le \sup_{t\in[0,T]} \int_0^t \left(\frac{1}{1+ar} + e^{-\kappa(t-u)}\right)^2 e^{2\mu u} du \\ &= \sup_{t\in[0,T]} \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|} + e^{-2\kappa t}\int_0^t e^{2(\kappa+\mu)u} du.\end{aligned}$$

where the first two terms have been carried over from the non-neutralizing vanishing-noise case, and for third term, we find that

$$e^{-2\kappa t}\int_0^t e^{2(\kappa+\mu)u} du = e^{-2\kappa t} t \le \frac{1}{2\kappa e},$$

Leading to the estimate $\|\Gamma\| \le v_1^2$, where

$$v_1^2 := \max\left\{\frac{1}{|2\mu|}, \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{1+ar}\frac{1-\frac{\kappa\wedge(2|\mu|)}{\kappa\vee(2|\mu|)}}{|(2\mu+\kappa)|} + \frac{1}{2\kappa e}\right\}.$$

Also here (B.1.3) yields

$$\mathbb{E}\left[\left(\int_s^t \check{x}(t-u)e^{\mu u} dW_u\right)^2\right] \le e^{2\mu s}(t-s) \le t-s,$$

For the integral term concerning $\mathcal{Q}_2$, we find an improvement in the estimate

$$\begin{aligned}
\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du &= \int_0^s \left(\int_s^t -a\check{x}(v-u)-a\check{x}(v-u-r)dv\right)^2 e^{2\mu u}du \\
&= a^2(1+e^{\kappa r})^2 \int_0^s \left(\int_s^t e^{-\kappa(v-u)}dv\right)^2 e^{2\mu u}du \\
&\leq a^2(1+e^{\kappa r})^2 \int_0^s \left(\int_s^t e^{-\kappa v}dv\right)^2 e^{2(\mu+\kappa)u}du \\
& a^2(1+e^{\kappa r})^2 \int_0^s e^{-2\kappa s}\left(\int_0^{t-s} e^{-\kappa v}dv\right)^2 du \\
&\leq a^2(1+e^{\kappa r})^2 \underbrace{se^{-2\kappa s}}_{\leq \frac{1}{2\kappa e}}(t-s)^2 \\
&\leq \frac{a^2(1+e^{\kappa r})^2}{2\kappa e}(t-s)^2.
\end{aligned}$$

Collecting the results,

$$\frac{\mathcal{Q}_1}{\sigma} \leq \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a(1+e^{\kappa r})}{\sqrt{2\kappa e}}\frac{T}{2p\log p}.$$

And therefore,

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_1 + \frac{7}{2}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{7}{2}\frac{a(1+e^{\kappa r})}{\sqrt{2\kappa e}}\frac{T}{2p\log p}.$$

Case #2: $2\mu+\kappa=0$. We continue with the case, where $\mu=-\frac{\kappa}{2}$, implying that $2\mu+\kappa<0$. For $\|\Gamma\|$, we find out that

$$\frac{\|\Gamma\|}{\sigma^2} \leq \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{1+ar}e^{-\kappa t}\int_0^t e^{\kappa u+2\mu u}du + \frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)$$

where the first and the third term were carried over from the non-neutralizing vanishing-noise case. For the middle term, we find that

$$\frac{2}{1+ar}e^{-\kappa t}\int_0^t e^{\kappa u+2\mu u}du \leq \frac{2e^{-\kappa t}t}{1+ar} \leq \frac{2}{(1+ar)\kappa e}.$$

That leads to the estimate

$$\|\Gamma\| \leq v_2^2,$$

where

$$v_2^2 := \max\left\{\frac{1}{|2\mu|}, \frac{1}{2|\mu|(1+ar)^2} + \frac{2}{(1+ar)\kappa e} + \frac{1}{2|\mu+\kappa|}\left(1-\frac{\kappa\wedge|\mu|}{\kappa\vee|\mu|}\right)\right\}.$$

Estimate (B.1.3) still holds, i.e.

$$\mathbb{E}\left[\left(\int_s^t \check{x}(t-u)e^{\mu u}dW(u)\right)^2\right] \leq e^{2\mu s}(t-s) \leq t-s,$$

And also can we take over from (B.1.4)

$$\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du \leq \frac{a^2(1+e^{\kappa r})^2}{2|\kappa+\mu|} \left(1 - \frac{\kappa \wedge |\mu|}{\kappa \vee |\mu|}\right) (t-s)^2.$$

Collecting the results,

$$\frac{\mathcal{Q}_1}{\sigma} \leq \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}} \sqrt{1 - \frac{\kappa \wedge |\mu|}{\kappa \vee |\mu|} \frac{T}{2p\log p}}.$$

So that

$$\frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} \leq v_2 + \frac{7}{2}\left(\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a(1+e^{\kappa r})}{\sqrt{2|\kappa+\mu|}} \sqrt{1 - \frac{\kappa \wedge |\mu|}{\kappa \vee |\mu|} \frac{T}{2p\log p}}\right).$$

B.2. Stable Regime

As *stable regime* we consider the SDDE in case where $a > b > 0$.

- $\exists \lambda > 0 : \tilde{a} = a - \lambda = be^{\lambda r} = \tilde{b}$ which implies $\check{x}(t) \leq e^{-\lambda t}$ for all $t \in [-r, \infty)$,
- $\exists \lambda > 0, \tilde{\kappa} > 0 : \check{x}(t) \leq \left(\frac{1}{1+\tilde{a}r} + e^{-\tilde{\kappa}t}\right) e^{-\lambda t}$.

Stable regime, non-neutralizing case ($a > b > 0$, $-\lambda \neq \mu$). Here, we start with some preparations for the cases where $-\lambda \neq \mu$.

$$\begin{aligned} \frac{\|\Gamma\|}{\sigma^2} &= \sup_{t\in[0,T]} \mathbb{E}\left[\left(\int_0^t \check{x}(t-u)e^{\mu u} dW_u\right)^2\right] \\ &\leq \sup_{t\in[0,T]} \int_0^t e^{-2\lambda(t-u)} e^{2\mu u} du \\ &\leq \sup_{t\in[0,T]} e^{-2\lambda t} \frac{e^{2(\lambda+\mu)t} - 1}{2(\lambda+\mu)} \leq \sup_{t\in[0,T]} \frac{e^{2\mu t} - e^{-2\lambda t}}{2(\lambda+\mu)} \end{aligned} \tag{B.2.1}$$

$$\begin{aligned} \frac{\|\Gamma\|}{\sigma^2} &\leq \sup_{t\in[0,T]} \int_0^t \left(\frac{1}{1+\tilde{a}r} + e^{-\tilde{\kappa}(t-u)}\right) e^{-\lambda(t-u)} e^{2\mu u} \\ &= \sup_{t\in[0,T]} \frac{e^{-2\lambda t}}{(1+\tilde{a}r)^2} \int_0^t e^{2\lambda u} e^{2\mu u} du + \frac{2e^{-\tilde{\kappa}t - 2\lambda t}}{1+\tilde{a}r} \int_0^t e^{\tilde{\kappa}u + 2\lambda u + 2\mu u} du \\ &\quad + e^{-2\tilde{\kappa}t - 2\lambda t} \int_0^t e^{2\tilde{\kappa}u + 2\lambda u + 2\mu u} du \\ &= \sup_{t\in[0,T]} \frac{e^{-2\lambda t}}{(1+\tilde{a}r)^2} \frac{e^{2(\lambda+\mu)t} - 1}{2(\lambda+\mu)} + \frac{2e^{-\tilde{\kappa}t-2\lambda t}}{1+\tilde{a}r} \frac{e^{(\tilde{\kappa}+2\lambda+2\mu)t} - 1}{\tilde{\kappa}+2\lambda+2\mu} \\ &\quad + e^{-2(\tilde{\kappa}+\lambda)t} \frac{e^{2(\tilde{\kappa}+\lambda+\mu)t} - 1}{2(\tilde{\kappa}+\lambda+\mu)} \\ &= \sup_{t\in[0,T]} \frac{1}{(1+\tilde{a}r)^2} \frac{e^{2\mu t} - e^{-2\lambda t}}{2(\lambda+\mu)} + \frac{2}{1+\tilde{a}r} \frac{e^{2\mu t} - e^{-\tilde{\kappa}t-2\lambda t}}{\tilde{\kappa}+2\lambda+2\mu} + \frac{e^{2\mu t} - e^{-2(\tilde{\kappa}+\lambda)t}}{2(\tilde{\kappa}+\lambda+\mu)}. \end{aligned} \tag{B.2.2}$$

$$\begin{aligned}\int_s^t \check{x}^2(t-u)e^{2\mu u}du &\le \int_s^t e^{-2\lambda(t-u)}e^{2\mu u} \le e^{-2\lambda t}e^{2(\lambda+\mu)s}\frac{e^{2(\lambda+\mu)(t-s)}-1}{2(\lambda+\mu)} \\ &= e^{2\mu s}\frac{e^{2\mu(t-s)}-e^{-2\lambda(t-s)}}{2(\lambda+\mu)}. \qquad \text{(B.2.3)}\end{aligned}$$

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du &\le \int_0^s \left(\int_s^t -a\check{x}(v-u)+b\check{x}(v-r-u)dv\right)^2 e^{2\mu u}du \\ &\le \int_0^s \left(\int_s^t ae^{-\lambda(v-u)}+be^{-\lambda(v-r-u)}dv\right)^2 e^{2\mu u}du \\ &= (a+be^{\lambda r})^2\int_0^s \left(\int_s^t e^{-\lambda(v-u)}dv\right)^2 e^{2\mu u}du \\ &= (a+be^{\lambda r})^2\int_0^s \left(e^{-\lambda s}\int_0^{t-s} e^{-\lambda(v-u)}dv\right)^2 e^{2\mu u}du \\ &\le (a+be^{\lambda r})^2\int_0^s e^{-2\lambda s}(t-s)^2 e^{2(\lambda+\mu)u}du \\ &= (a+be^{\lambda r})^2\int_0^s e^{-2\lambda s}e^{2(\lambda+\mu)u}du(t-s)^2 \\ &= (a+be^{\lambda r})^2 e^{-2\lambda s}\int_0^s e^{2(\lambda+\mu)u}du(t-s)^2 \\ &= (a+be^{\lambda r})^2 e^{-2\lambda s}\frac{e^{2(\lambda+\mu)s}-1}{2(\lambda+\mu)}(t-s)^2. \qquad \text{(B.2.4)}\end{aligned}$$

Which can also be written as

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du &\le (a+be^{\lambda r})^2 e^{-2\lambda s}\frac{e^{2(\lambda+\mu)s}-1}{2(\lambda+\mu)}(t-s)^2 \\ &= (a+be^{\lambda r})^2\frac{e^{2\mu s}-e^{-2\lambda s}}{2(\lambda+\mu)}(t-s)^2. \qquad \text{(B.2.5)}\end{aligned}$$

Stable regime, white noise. ($a > b > 0$, $\mu = 0$). From (B.2.1), we deduce that

$$\frac{\|\Gamma\|}{\sigma^2} \le \frac{1}{2\lambda}.$$

From (B.2.2), we find that

$$\begin{aligned}\frac{\|\Gamma\|}{\sigma^2} &\le \sup_{t\in[0,T]} \frac{1}{(1+\tilde{a}r)^2}\frac{1-e^{-2\lambda t}}{2\lambda} + \frac{2}{1+\tilde{a}r}\frac{1-e^{-\tilde{\kappa}t-2\lambda t}}{\tilde{\kappa}+2\lambda} + \frac{1-e^{-2(\tilde{\kappa}+\lambda)t}}{2(\tilde{\kappa}+\lambda)} \\ &\le \frac{1}{(1+\tilde{a}r)^2}\frac{1}{2\lambda} + \frac{2}{1+\tilde{a}r}\frac{1}{\tilde{\kappa}+2\lambda} + \frac{1}{2(\tilde{\kappa}+\lambda)}.\end{aligned}$$

Hence,

$$\frac{\|\Gamma\|}{\sigma^2} \le v_0^2 \quad \text{where} \quad v_0^2 := \min\left\{\frac{1}{2\lambda}, \frac{1}{(1+\tilde{a}r)^2}\frac{1}{2\lambda} + \frac{2}{1+\tilde{a}r}\frac{1}{\tilde{\kappa}+2\lambda} + \frac{1}{2(\tilde{\kappa}+\lambda)}\right\}.$$

With (B.2.3), we find that

$$\int_s^t \check{x}^2(t-u)du \leq \frac{1-e^{-2\lambda(t-s)}}{2\lambda} \leq t-s.$$

Then, starting from (B.2.5)

$$\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 du \leq (a+be^{\lambda r})^2 \frac{1-e^{-2\lambda s}}{2\lambda}(t-s)^2 \leq \frac{(a+be^{-\lambda r})^2}{2\lambda}(t-s)^2.$$

$$\frac{\mathcal{Q}_1}{\sigma} \leq \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a+be^{-\lambda r}}{\sqrt{2\lambda}} \frac{T}{2p\log p}.$$

Collecting the results,

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_0 + \frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a+be^{-\lambda r}}{\sqrt{2\lambda}} \frac{T}{2p\log p}.$$

Stable regime, vanishing noise, no parameter-cancellation issues ($a > b > 0$, $\mu < 0$, $\mu \notin \{-\lambda, -\lambda - \frac{\tilde{\kappa}}{2}, -\lambda - \tilde{\kappa}\}$). From (B.2.1) we get

$$\frac{\|\Gamma\|}{\sigma^2} \leq \sup_{t\in[0,T]} \frac{e^{2\mu t}-e^{-2\lambda t}}{2(\lambda+\mu)} \leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|}. \tag{B.2.6}$$

We know from (B.2.2) that

$$\begin{aligned}\frac{\|\Gamma\|}{\sigma^2} &\leq \sup_{t\in[0,T]} \frac{1}{(1+\tilde{a}r)^2}\frac{e^{2\mu t}-e^{-2\lambda t}}{2(\lambda+\mu)} + \frac{2}{1+\tilde{a}r}\frac{e^{2\mu t}-e^{-\tilde{\kappa}t-2\lambda t}}{\tilde{\kappa}+2\lambda+2\mu} + \frac{e^{2\mu t}-e^{-2(\tilde{\kappa}+\lambda)t}}{2(\tilde{\kappa}+\lambda+\mu)} \\ &\leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2} + \frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde{\kappa}+2\lambda)}{(2|\mu|)\vee(\tilde{\kappa}+2\lambda)}\right)}{|\tilde{\kappa}+2\lambda+2\mu|(1+\tilde{a}r)} + \frac{1-\frac{|\mu|\wedge(\tilde{\kappa}+\lambda)}{|\mu|\vee(\tilde{\kappa}+\lambda)}}{2|\tilde{\kappa}+\lambda+\mu|} =: v_1^2. \end{aligned} \tag{B.2.7}$$

From (B.2.3), we have that

$$\begin{aligned}\int_s^t \check{x}^2(t-u)e^{2\mu u}du &\leq e^{2\mu s}\frac{e^{2\mu(t-s)}-e^{-2\lambda(t-s)}}{2(\lambda+\mu)} \qquad \text{(B.2.8)}\\ &= \underbrace{e^{2\mu s}e^{-2(|\mu|\wedge\lambda)(t-s)}}_{\leq 1}\frac{1-e^{-2(|\mu|\vee\lambda-|\mu|\wedge\lambda)(t-s)}}{2|\lambda+\mu|} \\ &\leq \frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}(t-s). \qquad \text{(B.2.9)}\end{aligned}$$

From (B.2.5), we get that

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2e^{2\mu u}du &\leq (a+be^{\lambda r})^2e^{-2\lambda s}\frac{e^{2(\lambda+\mu)s}-1}{2(\lambda+\mu)}(t-s)^2 \\ &= (a+be^{\lambda r})^2\frac{e^{2\mu s}-e^{-2\lambda s}}{2(\lambda+\mu)}(t-s)^2 \\ &\leq \frac{(a+be^{\lambda r})^2}{2|\lambda+\mu|}\left(1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}\right)(t-s)^2. \qquad \text{(B.2.10)}\end{aligned}$$

Collecting the results, we find that

$$\frac{\mathcal{Q}_1}{\sigma} \le \sqrt{\frac{(|\mu| \vee \lambda) - (|\mu| \wedge \lambda)}{|\lambda + \mu|}} \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \le \frac{a + be^{\lambda r}}{\sqrt{2|\lambda + \mu|}} \sqrt{1 - \frac{|\mu| \wedge \lambda}{|\mu| \vee \lambda}} \frac{T}{2p \log p}.$$

And so,

$$\begin{aligned} \frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} &\le v_1 + \sqrt{\frac{(|\mu| \vee \lambda) - (|\mu| \wedge \lambda)}{|\lambda + \mu|}} \frac{\sqrt{T}}{\sqrt{p}\log p} \\ &\quad + \frac{a + be^{\lambda r}}{\sqrt{2|\lambda + \mu|}} \sqrt{1 - \frac{|\mu| \wedge \lambda}{|\mu| \vee \lambda}} \frac{T}{2p \log p}. \end{aligned}$$

Stable regime, increasing noise ($a > b > 0$, $\mu > 0$). From (B.2.1), we obtain

$$\frac{\|\Gamma\|}{\sigma^2} \le \sup_{t \in [0,T]} \frac{e^{2\mu t} - e^{-2\lambda t}}{2(\lambda + \mu)} \le \frac{e^{2\mu T}}{2(\lambda + \mu)}.$$

And (B.2.2) yields

$$\begin{aligned} \frac{\|\Gamma\|}{\sigma^2} &\le \sup_{t \in [0,T]} \frac{1}{(1 + \tilde{a}r)^2} \frac{e^{2\mu t} - e^{-2\lambda t}}{2(\lambda + \mu)} + \frac{2}{1 + \tilde{a}r} \frac{e^{2\mu t} - e^{-\tilde{\kappa}t - 2\lambda t}}{\tilde{\kappa} + 2\lambda + 2\mu} + \frac{e^{2\mu t} - e^{-2(\tilde{\kappa} + \lambda)t}}{2(\tilde{\kappa} + \lambda + \mu)} \\ &\le \frac{e^{2\mu T}}{2(\lambda + \mu)(1 + \tilde{a}r)^2} + \frac{2e^{2\mu T}}{(1 + \tilde{a}r)(\tilde{\kappa} + 2\lambda + 2\mu)} + \frac{e^{2\mu T}}{2(\tilde{\kappa} + \lambda + \mu)}. \end{aligned}$$

Then,

$$\frac{\|\Gamma\|}{\sigma^2} \le v_0^2 e^{2\mu T}, \quad \text{where} \quad v_0^2 := \frac{1}{2(\lambda + \mu)(1 + \tilde{a}r)^2} + \frac{2}{(1 + \tilde{a}r)(\tilde{\kappa} + 2\lambda + 2\mu)} + \frac{1}{2(\tilde{\kappa} + \lambda + \mu)}.$$

From (B.2.3), we have that

$$\int_s^t \check{x}^2(t - u) e^{2\mu u} du \le e^{2\mu s} \frac{e^{2\mu(t-s)} - e^{-2\lambda(t-s)}}{2(\lambda + \mu)} = e^{2\mu t} \frac{1 - e^{-2(\lambda + \mu)(t-s)}}{2(\lambda + \mu)} \le e^{2\mu T}(t - s). \tag{B.2.11}$$

With the help of (B.2.5), we obtain

$$\begin{aligned} \int_0^s (\check{x}(t - u) - \check{x}(s - u))^2 e^{2\mu u} du &\le \left(a + be^{\lambda r}\right)^2 \frac{e^{2\mu s} - e^{-2\lambda s}}{2(\lambda + \mu)} (t - s)^2 \\ &\le e^{2\mu T} \frac{\left(a + be^{\lambda r}\right)^2}{2(\lambda + \mu)} (t - s)^2. \end{aligned}$$

Collecting the results,

$$\frac{\mathcal{Q}_1}{\sigma} \le e^{\mu T} \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \le \frac{a + be^{\lambda r}}{\sqrt{2(\mu + \lambda)}} e^{\mu T} \frac{T}{2p \log p}.$$

And therefore,

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \le v_0 e^{\mu T}\left(1+\frac{7\sqrt{2(\lambda+\mu)}}{2v_0}\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{7(a+be^{\lambda r})}{2v_0}\frac{T}{2p\log p}\right).$$

Stable regime, neutralizing-parameter issues ($a > b > 0$, $\mu \in \{-\lambda, -\lambda-\frac{\tilde\kappa}{2}, -\lambda-\tilde\kappa)$. *Case #1:* $\mu+\lambda=0$. We start with the vanishing-noise case $-\lambda=\mu$. For $\|\Gamma\|$, we may carry over part of the computation for the second and third arising term from (B.2.2). In particular, we know that $2|\mu| < 2\lambda+\tilde\kappa$.

$$\begin{aligned}
\frac{\|\Gamma\|}{\sigma^2} &= \sup_{t\in[0,T]} \frac{e^{-2\lambda t}}{(1+\tilde a r)^2}\int_0^t e^{2(\lambda+\mu)u}du + \frac{2e^{-\tilde\kappa t-2\lambda t}}{1+\tilde a r}\int_0^t e^{\tilde\kappa u+2\lambda u+2\mu u}du \\
&\quad + e^{-2\tilde\kappa t-2\lambda t}\int_0^t e^{2\tilde\kappa u+2\lambda u+2\mu u}du \\
&= \sup_{t\in[0,T]} \frac{e^{-2\lambda t}}{(1+\tilde a r)^2}t + \frac{2e^{-\tilde\kappa t-2\lambda t}}{1+\tilde a r}\frac{e^{(\tilde\kappa+2\lambda+2\mu)t}-1}{\tilde\kappa+2\lambda+2\mu} + e^{-2(\tilde\kappa+\lambda)t}\frac{e^{2(\tilde\kappa+\lambda+\mu)t}-1}{2(\tilde\kappa+\lambda+\mu)} \\
&= \sup_{t\in[0,T]} \frac{1}{(1+\tilde a r)^2 2\lambda e} + \frac{2}{1+\tilde a r}\frac{e^{2\mu t}-e^{-\tilde\kappa t-2\lambda t}}{\tilde\kappa+2\lambda+2\mu} + \frac{e^{2\mu t}-e^{-2(\tilde\kappa+\lambda)t}}{2(\tilde\kappa+\lambda+\mu)} \\
&= \sup_{t\in[0,T]} \frac{1}{(1+\tilde a r)^2 2\lambda e} + \frac{2}{1+\tilde a r}\frac{\left(1-\frac{(2|\mu|)\wedge(\tilde\kappa+2\lambda)}{(2|\mu|)\vee(\tilde\kappa+2\lambda)}\right)}{\tilde\kappa+2\lambda+2\mu} + \frac{\left(1-\frac{|\mu|\wedge(\tilde\kappa+\lambda)}{|\mu|\vee(\tilde\kappa+\lambda)}\right)}{2(\tilde\kappa+\lambda+\mu)}. \\
&= \frac{1}{(1+\tilde a r)^2 2\lambda e} + \frac{2}{1+\tilde a r}\frac{\left(1-\frac{2|\mu|}{\tilde\kappa+2\lambda}\right)}{\tilde\kappa+2\lambda+2\mu} + \frac{\left(1-\frac{|\mu|}{\tilde\kappa+\lambda}\right)}{2(\tilde\kappa+\lambda+\mu)}.
\end{aligned}$$

$$\begin{aligned}
\frac{\|\Gamma\|}{\sigma^2} = \sup_{t\in[0,T]}\int_0^t \check x^2(t-u)e^{2\mu u}du &\le \sup_{t\in[0,T]}\int_0^t e^{-2\lambda(t-u)}e^{2\mu u}du \\
&\le \sup_{t\in[0,T]}\int_0^t e^{-2\lambda t}du = \sup_{t\in[0,T]} te^{-2\lambda t} \le \frac{1}{2\lambda e}.
\end{aligned}$$

Then,

$$\frac{\|\Gamma\|}{\sigma^2} \le v_0^2 \quad \text{where} \quad v_0^2 := \max\left\{\frac{1}{2\lambda e}, \frac{1}{(1+\tilde a r)^2 2\lambda e} + \frac{2}{1+\tilde a r}\frac{\left(1-\frac{2|\mu|}{\tilde\kappa+2\lambda}\right)}{\tilde\kappa+2\lambda+2\mu} + \frac{\left(1-\frac{|\mu|}{\tilde\kappa+\lambda}\right)}{2(\tilde\kappa+\lambda+\mu)}\right\}.$$

$$\int_s^t \check x^2(t-u)e^{2\mu u}du \le \int_s^t e^{-2\lambda(t-u)}e^{2\mu u}du = \int_s^t e^{-2\lambda t}du = e^{-2\lambda t}(t-s) \le (t-s)$$

$$\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du \leq \int_0^t \left(\int_s^t -a\check{x}(v-u)+b\check{x}(v-u-r)dv\right)^2 e^{2\mu u}du$$
$$\leq \int_0^t \left(\int_s^t ae^{-\lambda(v-u)}+be^{-\lambda(v-u-r)}dv\right)^2 e^{2\mu u}du$$
$$= \left(a+be^{\lambda r}\right)^2 \int_0^t e^{-2\lambda v}e^{2\mu u}du \left(\int_s^t e^{-\lambda u}dv\right)^2$$
$$= \left(a+be^{\lambda r}\right)^2 e^{-2\lambda s} t(t-s)^2 \leq \frac{\left(a+be^{\lambda r}\right)^2}{2\lambda e}(t-s)^2.$$

Collecting the results

$$\frac{\mathcal{Q}_1}{\sigma} \leq \frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a+be^{\lambda r}}{\sqrt{2\lambda e}}\frac{T}{2p\log p}.$$

And, therefore,

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_0 + \frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a+be^{\lambda r}}{\sqrt{2\lambda e}}\frac{T}{2p\log p}.$$

Case #2: $\tilde{\kappa}+2\mu+2\lambda=0$. In this case $\mu=-\lambda-\frac{\tilde{\kappa}}{2}$ so that this parameter-neutralization occurs as a special case in the vanishing-noise case. That implies that we take over most of the estimates from the prior case and work out new estimates when it is necessary. In particular $\lambda<|\mu|$ and $|\mu+\lambda|=\frac{\tilde{\kappa}}{2}$. To begin with, we may keep estimate (B.2.6)

$$\frac{\|\Gamma\|}{\sigma^2} \leq \sup_{t\in[0,T]} \frac{e^{2\mu t}-e^{-2\lambda t}}{2(\lambda+\mu)} \leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|} = \frac{1-\frac{\lambda}{|\mu|}}{\tilde{\kappa}}.$$

From (B.2.2) we may take over the first and the third term, and achieve that

$$\frac{\|\Gamma\|}{\sigma^2} \leq \sup_{t\in[0,T]} \int_0^t \left(\frac{1}{1+\tilde{a}r}+e^{-\tilde{\kappa}(t-u)}\right)e^{-\lambda(t-u)}e^{2\mu u}du$$
$$= \sup_{t\in[0,T]} \frac{e^{-2\lambda t}}{(1+\tilde{a}r)^2}\int_0^t e^{2\lambda u}e^{2\mu u}du + \frac{2e^{-\tilde{\kappa}t-2\lambda t}}{1+\tilde{a}r}\int_0^t e^{\tilde{\kappa}u+2\lambda u+2\mu u}du$$
$$+ e^{-2\tilde{\kappa}t-2\lambda t}\int_0^t e^{2\tilde{\kappa}u+2\lambda u+2\mu u}du$$
$$\leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2} + \frac{2e^{-\tilde{\kappa}t-2\lambda t}}{1+\tilde{a}r}t + \frac{1-\frac{|\mu|\wedge(\tilde{\kappa}+\lambda)}{|\mu|\vee(\tilde{\kappa}+\lambda)}}{2|\tilde{\kappa}+\lambda+\mu|} \leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2}$$
$$+ \frac{2}{(1+\tilde{a}r)(\tilde{\kappa}+2\lambda)e} + \frac{1-\frac{|\mu|\wedge(\tilde{\kappa}+\lambda)}{|\mu|\vee(\tilde{\kappa}+\lambda)}}{2(\tilde{\kappa}+\lambda+\mu)}$$

And so,

$$\frac{\|\Gamma\|}{\sigma^2} \leq v_0^2 \quad \text{where} \quad v_0^2 := \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2} + \frac{2}{(1+\tilde{a}r)(\tilde{\kappa}+2\lambda)e} + \frac{1-\frac{|\mu|\wedge(\tilde{\kappa}+\lambda)}{|\mu|\vee(\tilde{\kappa}+\lambda)}}{2(\tilde{\kappa}+\lambda+\mu)}.$$

From (B.2.9)

$$\int_s^t \check{x}^2(t-u)e^{2\mu u}du \leq \frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}(t-s).$$

From (B.2.10)

$$\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du \leq \frac{(a+be^{\lambda r})^2}{2|\lambda+\mu|}\left(1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}\right)(t-s)^2.$$

$$\frac{\mathcal{Q}_1}{\sigma} \leq \sqrt{\frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}}\frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a+be^{\lambda r}}{\sqrt{2|\lambda+\mu|}}\sqrt{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}\frac{T}{2p\log p}.$$

And so,

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq v_0 + \sqrt{\frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{|\lambda+\mu|}}\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a+be^{\lambda r}}{\sqrt{2|\lambda+\mu|}}\sqrt{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}\frac{T}{2p\log p}.$$

Case #3: $\tilde{\kappa}+\mu+\lambda=0$. In this case, $\tilde{\kappa}+2\mu+2\lambda<0$, $\mu+\lambda<0$ and .... So we may again take over several of the results from the vanishing-noise regime.

Estimate (B.2.6) is preserved, besides $|\lambda+\mu|=\frac{\tilde{\kappa}}{2}$. Hence,

$$\frac{\|\Gamma\|}{\sigma^2} \leq \sup_{t\in[0,T]}\frac{e^{2\mu t}-e^{-2\lambda t}}{2(\lambda+\mu)} \leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|} = \frac{1-\frac{\lambda}{|\mu|}}{2\tilde{\kappa}}.$$

From (B.2.2), we may carry over the estimates for the first two terms from (B.2.7)

$$\begin{aligned}
\frac{\|\Gamma\|}{\sigma^2} &\leq \sup_{t\in[0,T]}\int_0^t \left(\frac{1}{1+\tilde{a}r}+e^{-\tilde{\kappa}(t-u)}\right)e^{-\lambda(t-u)}e^{2\mu u} \\
&= \sup_{t\in[0,T]}\frac{e^{-2\lambda t}}{(1+\tilde{a}r)^2}\int_0^t e^{2\lambda u}e^{2\mu u}du + \frac{2e^{-\tilde{\kappa}t-2\lambda t}}{1+\tilde{a}r}\int_0^t e^{\tilde{\kappa}u+2\lambda u+2\mu u}du \\
&\quad + e^{-2\tilde{\kappa}t-2\lambda t}\int_0^t e^{2\tilde{\kappa}u+2\lambda u+2\mu u}du \\
&\leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2} + \frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde{\kappa}+2\lambda)}{(2|\mu|)\vee(\tilde{\kappa}+2\lambda)}\right)}{|\tilde{\kappa}+2\lambda+2\mu|(1+\tilde{a}r)} + \sup_{t\in[0,T]} e^{-(\tilde{\kappa}+2\lambda)t}t \\
&\leq \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2|\lambda+\mu|(1+\tilde{a}r)^2} + \frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde{\kappa}+2\lambda)}{(2|\mu|)\vee(\tilde{\kappa}+2\lambda)}\right)}{|\tilde{\kappa}+2\lambda+2\mu|(1+\tilde{a}r)} + \frac{1}{(\tilde{\kappa}+2\lambda)e}.
\end{aligned}$$

That motivates the notation

$$\frac{\|\Gamma\|}{\sigma^2} \leq v_0^2 \quad \text{with} \quad v_0^2 := \max\left\{\frac{1-\frac{\lambda}{|\mu|}}{2\tilde{\kappa}}, \frac{1-\frac{|\mu|\wedge\lambda}{|\mu|\vee\lambda}}{2\tilde{\kappa}(1+\tilde{a}r)^2} + \frac{2\left(1-\frac{(2|\mu|)\wedge(\tilde{\kappa}+2\lambda)}{(2|\mu|)\vee(\tilde{\kappa}+2\lambda)}\right)}{\tilde{\kappa}(1+\tilde{a}r)} + \frac{1}{(\tilde{\kappa}+2\lambda)e}\right\}.$$

From (B.2.9)

$$\int_s^t \check{x}^2(t-u)^2 e^{2\mu u}du \leq \frac{(|\mu|\vee\lambda)-(|\mu|\wedge\lambda)}{\tilde{\kappa}}(t-s).$$

From (B.2.10)

$$\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du \le \frac{(a+be^{\lambda r})^2}{2\tilde{\kappa}} \left(1 - \frac{|\mu| \wedge \lambda}{|\mu| \vee \lambda}\right) (t-s)^2.$$

That leads to

$$\frac{\mathcal{Q}_1}{\sigma} = \sqrt{\frac{(|\mu| \vee \lambda) - (|\mu| \wedge \lambda)}{\tilde{\kappa}}} \frac{\sqrt{T}}{\sqrt{p}\log p},$$
$$\frac{\mathcal{Q}_2}{\sigma} = \sqrt{\frac{(a+be^{\lambda r})^2}{2\tilde{\kappa}} \left(1 - \frac{|\mu| \wedge \lambda}{|\mu| \vee \lambda}\right)} \frac{T}{p\log(p)}.$$

And so,

$$\frac{\sqrt{\|\Gamma\|} + Q(T)}{\sigma} \le v_0 + \sqrt{\frac{(|\mu| \vee \lambda) - (|\mu| \wedge \lambda)}{\tilde{\kappa}}} \frac{\sqrt{T}}{\sqrt{p}\log p} + \sqrt{\frac{(a+be^{\lambda r})^2}{2\tilde{\kappa}} \left(1 - \frac{|\mu| \wedge \lambda}{|\mu| \vee \lambda}\right)} \frac{T}{p\log(p)}.$$

B.3. Instable Regime

Instable regime, non-neutralizing case ($0 < a < b$, $\mu \neq \lambda$). As long as $\mu \neq \lambda$,

$$\begin{aligned}\int_s^t \check{x}^2(t-u) e^{2\mu u} du &\le \int_s^t e^{2\lambda(t-u)} e^{2\mu u} du \\ &= e^{2\lambda t} \int_s^t e^{2(\mu-\lambda)u} du \\ &= e^{2\lambda t} e^{2(\mu-\lambda)s} \frac{e^{2(\mu-\lambda)(t-s)} - 1}{2(\mu-\lambda)},\end{aligned} \tag{B.3.1}$$

$$\begin{aligned}\int_0^s (\check{x}(t-u) - \check{x}(s-u))^2 e^{2\mu u} du &= \int_0^s \left(\int_s^t -a\check{x}(u-v) + b\check{x}(v-u-r) dv\right)^2 e^{2\mu u} du \\ &\le \int_0^s \left(\int_s^t ae^{\lambda(v-u)} + be^{\lambda(v-u-r)} dv\right)^2 e^{2\mu u} du \\ &= \int_0^s (a+be^{-\lambda r})^2 \left(\int_s^t e^{\lambda v} dv\right)^2 e^{2(\mu-\lambda)u} du \\ &= \frac{(a+be^{-\lambda r})^2}{2(\mu-\lambda)} \left(e^{2(\mu-\lambda)s} - 1\right) e^{2\lambda t} (t-s)^2.\end{aligned} \tag{B.3.2}$$

Instable regime, vanishing noise ($0 < a < b$, $\mu < 0$).

$$\int_0^t \check{x}^2(t-u) e^{2\mu} du \le \int_0^t \left(\frac{1}{1+\tilde{a}r} + e^{-\kappa(t-u)}\right) e^{2\lambda(t-u)} e^{2\mu u} du = \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3,$$

where

$$\mathcal{I}_1 := \frac{1}{(1+\tilde{a}r)^2} e^{2\lambda t} \frac{1}{2|\mu-\lambda|} \left(1 - e^{2(\mu-\lambda)t}\right) \tag{B.3.3}$$

$$\mathcal{I}_2 := \frac{2}{1+\tilde{a}r} e^{(-\kappa+2\lambda)t} \int_0^t e^{(\kappa-2\lambda+2\mu)u} du \tag{B.3.4}$$

$$\mathcal{I}_3 := e^{(-2\kappa+2\lambda)t} \int_0^t e^{2(\kappa-\lambda+\mu)} du \tag{B.3.5}$$

Case #1: $\kappa - \lambda + \mu < 0$. For the term $\mathcal{I}_2$ from (B.3.4), we find

$$\mathcal{I}_2 = \frac{2}{1+\tilde{a}r} e^{(-\kappa+2\lambda)t} \frac{1}{|\kappa-2\lambda+2\mu|} \underbrace{\left(1 - e^{(\kappa-2\lambda+2\mu)t}\right)}_{\leq 1} \leq \frac{2}{1+\tilde{a}r} e^{-\kappa t} e^{2\lambda t} \frac{1}{|\kappa-2\lambda+2\mu|}. \tag{B.3.6}$$

For term $\mathcal{I}_3$ from (B.3.5)

$$\mathcal{I}_3 = e^{(-2\kappa+2\lambda)t} \frac{1}{2|\kappa-\lambda+\mu|} \underbrace{\left(1 - e^{2(\kappa-\lambda+\mu)t}\right)}_{\leq 1} \leq e^{2\lambda t} \frac{e^{-2\kappa t}}{2|\kappa-\lambda+\mu|}. \tag{B.3.7}$$

Case #2: $\kappa - 2\lambda + 2\mu < 0, \kappa - \lambda + \mu > 0$. Estimate (B.3.6) holds for term $\mathcal{I}_2$ from (B.3.4), and for term $\mathcal{I}_3$ from (B.3.5) we find

$$\begin{aligned} \mathcal{I}_3 &= e^{(-2\kappa+2\lambda)t} \frac{1}{2(\kappa-\lambda+\mu)} \left(e^{2(\kappa-\lambda+\mu)t} - 1\right) \\ &= e^{2\lambda t} \frac{1}{2(\kappa-\lambda+\mu)} \underbrace{\left(e^{-2\lambda t} - e^{-2(\kappa-\mu)t}\right)}_{\leq 1} e^{2\mu t} \leq e^{2\lambda t} e^{2\mu t} \frac{1}{2(\kappa-\lambda+\mu)}. \end{aligned} \tag{B.3.8}$$

Case #3: $\kappa - 2\lambda + 2\mu > 0$. For term $\mathcal{I}_2$ from (B.3.4), we get

$$\begin{aligned} \mathcal{I}_2 &= \frac{2e^{2\lambda t}}{1+\tilde{a}r} e^{-\kappa t} \frac{1}{\kappa-2\lambda+2\mu} \left(e^{(\kappa-2\lambda+2\mu)t} - 1\right) \\ &= \frac{2e^{2\lambda t}}{1+\tilde{a}r} \frac{1}{\kappa-2\lambda+2\mu} \underbrace{\left(e^{-2\lambda t} - e^{(-\kappa-2\mu)t}\right)}_{|\cdot|\leq 1} e^{2\mu t} \\ &\leq \frac{2e^{2\lambda t}}{1+\tilde{a}r} \frac{e^{2\mu t}}{\kappa-2\lambda+2\mu}. \end{aligned} \tag{B.3.9}$$

And for the term $\mathcal{I}_3$ from (B.3.5) we may take over the estimate (B.3.8).

Case #4: $\kappa = 2\lambda - 2\mu$. Meaning that $\kappa - 2\lambda + 2\mu = 0$ and $\kappa - \lambda + \mu > 0$. Then we find for $\mathcal{I}_2$ from (B.3.4)

$$\begin{aligned} \mathcal{I}_2 = \frac{2}{1+\tilde{a}r} e^{(-\kappa+2\lambda)t} \int_0^t \underbrace{e^{(\kappa-2\lambda+2\mu)u}}_{=1} du = \frac{2}{1+\tilde{a}r} e^{2\lambda t} e^{-\kappa t} t = \frac{2}{1+\tilde{a}r} e^{2\lambda t} e^{2\mu t} \underbrace{e^{(-\kappa-2\mu)t} t}_{\leq \frac{1}{(-\kappa+2\mu)e}} \\ = \frac{2}{1+\tilde{a}r} \frac{1}{(-\kappa+2\mu)e} e^{2\lambda t} e^{2\mu t}. \end{aligned}$$

For term $\mathcal{I}_3$ from (B.3.5), we can use estimate (B.3.7), i.e.

$$\mathcal{I}_3 = e^{(-2\kappa+2\lambda)t}\int_0^t e^{2(\kappa-\lambda+\mu)u}du \le e^{2\lambda t}\frac{e^{-2\kappa t}}{2|\kappa-\lambda+\mu|}.$$

Case #5: $\kappa = \lambda - \mu$. Here, $\kappa-\lambda+\mu = 0$ and $\kappa - 2\lambda + \mu < 0$. Then for term $\mathcal{I}_2$ from(B.3.4), we use estimate (B.3.6), i.e.

$$\mathcal{I}_2 \le \frac{2}{1+\tilde{a}r}e^{-\kappa t}e^{2\lambda t}\frac{1}{|\kappa-2\lambda+2\mu|}.$$

And for term $\mathcal{I}_3$ from (B.3.5), we have that for arbitrary $\nu \in (0,\kappa)$ that

$$\mathcal{I}_3 = e^{(-2\kappa+2\lambda)t}\int_0^t \underbrace{e^{2(\kappa-\lambda+\mu)u}}_{=1}du = e^{2\lambda t}e^{-2(\kappa-\nu)t}e^{-2\nu t}t \le e^{2\lambda t}e^{-2(\kappa-\nu)t}\frac{1}{2\nu e}.$$

Alltogether, we find that for aritrarily fixed $\nu \in (0,\kappa)$,

$$\frac{\|\Gamma\|}{\sigma^2} \le \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 \le \frac{e^{2\lambda T}}{(1+\tilde{a}r)^2}\left(1+\mathcal{O}\left(\frac{e^{-\min\{\kappa-\nu,2|\mu|\}t}}{\nu}\right)\right)$$

From (B.3.1),

$$\begin{aligned}\int_s^t \check{x}^2(t-u)e^{2\mu u}du &\le e^{2\lambda t}e^{2(\mu-\lambda)s}\frac{e^{2(\mu-\lambda)(t-s)}}{2(\mu-\lambda)} \le e^{2\lambda t}e^{(\mu-\lambda)s}(t-s) = e^{2\lambda(t-s)}e^{2\mu s}(t-s)\\ &\le e^{2\lambda T}(t-s).\end{aligned}$$

And with (B.3.2)

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2e^{2\mu u}du &\le \frac{(a+be^{-\lambda r})^2}{2(\mu-\lambda)}\left(e^{2(\mu-\lambda)s}-1\right)e^{2\lambda s}(t-s)^2\\ &\le \frac{(a+be^{-\lambda r})^2}{2(\mu-\lambda)}e^{2\lambda T}(t-s)^2,\end{aligned}$$

$$\frac{\mathcal{Q}_1}{\sigma} \le e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \le e^{\lambda T}\frac{a+be^{-\lambda r}}{\sqrt{2(\mu-\lambda)}}\cdot\frac{T}{2p\log p}.$$

Collecting the results,

$$\begin{aligned}\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \le \frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde{a}r)}&\left(1+\mathcal{O}\left(\frac{e^{-\min\{\kappa-\nu,2|\mu|\}T}}{\nu}\right)\right.\\ &\left.+\sqrt{2\lambda}(1+\tilde{a}r)\left(\frac{\sqrt{T}}{\sqrt{p}\log p}+\frac{a+be^{\lambda r}}{\sqrt{2(\mu-\lambda)}}\frac{T}{2p\log p}\right)\right).\end{aligned}$$

Instable regime, increasing noise $(0 < a < b,\ \mu > 0)$.

$$\int_0^t \check{x}^2(t-u)e^{2\mu}du = \int_0^t\left(\frac{1}{1+\tilde{a}r}+e^{-\kappa(t-u)}\right)^2 e^{2\lambda(t-u)}e^{2\mu u}du \le \mathcal{I}_1+\mathcal{I}_2+\mathcal{I}_3$$

where

$$\mathcal{I}_1 := \frac{e^{2\lambda t}}{(1+\tilde{a}r)^2} \int_0^t e^{2(\mu-\lambda)u} du, \tag{B.3.10}$$

$$\mathcal{I}_2 := \frac{2e^{2(\lambda-\kappa)t}}{1+\tilde{a}r} \int_0^t e^{(-2\lambda-\kappa+2\mu)u} du \tag{B.3.11}$$

$$\mathcal{I}_3 := e^{(2\lambda-2\kappa)t} \int_0^t e^{(-2\lambda+2\kappa+2\mu)} du. \tag{B.3.12}$$

Instable regime, weakly increasing noise ($0 < a < b$, $\mu > 0$, $\lambda > \mu$). *Case #1:* $\lambda > \mu$, $\kappa - 2\lambda + 2\mu \neq 0$, $\kappa - \lambda + \mu \neq 0$.

$$\mathcal{I}_1 = \frac{1}{(1+\tilde{a}r)^2} \frac{e^{2\lambda t} - e^{2\mu t}}{2(\mu - \lambda)} \leq \frac{e^{2\lambda t}}{2(\lambda-\mu)(1+\tilde{a}r)^2}.$$

$$\mathcal{I}_2 = \frac{2e^{(2\lambda-\kappa)t}}{1+\tilde{a}r} \frac{e^{(\kappa-2\lambda+2\mu)t} - 1}{\kappa - 2\lambda + 2\mu} = \frac{2e^{2\lambda t}}{1+\tilde{a}r} \frac{e^{-2(\lambda-\mu)t} - e^{-\kappa t}}{\kappa - 2\lambda + 2\mu} \leq \frac{2e^{2\lambda t}}{1+\tilde{a}r} \frac{e^{-\min\{2(\lambda-\mu),\kappa\}t}}{|2\lambda - 2\mu - \kappa|}.$$

$$\mathcal{I}_3 = e^{2(\lambda-\kappa)t} \frac{e^{2(\kappa-\lambda+\mu)t} - 1}{2(\kappa - \lambda + \mu)} = e^{2\lambda t} \frac{e^{-2(\lambda-\mu)t} - e^{-2\kappa t}}{2(\kappa - \lambda + \mu)} \leq e^{2\lambda t} \frac{e^{-2\min\{\lambda-\mu,\kappa\}}}{2|\lambda - \mu - \kappa|}.$$

Case #2: $\lambda > \mu$, $\kappa - 2\lambda + 2\mu = 0 \Leftrightarrow \mu = \lambda - \frac{\kappa}{2}$, and implying that $\kappa - \lambda + \mu > 0$. We may keep terms $\mathcal{I}_1$ and $\mathcal{I}_3$, i.e.

$$\mathcal{I}_1 \leq \frac{e^{2\lambda t}}{2(\lambda-\mu)(1+\tilde{a}r)^2} \quad \text{and} \quad \mathcal{I}_3 \leq e^{2\lambda t} \frac{e^{-2\min\{\lambda-\mu,\kappa\}}}{2|\lambda - \mu - \kappa|}.$$

And for the remaining $\mathcal{I}_2$, we find that for arbitrary $\nu \in (0, \kappa)$

$$\mathcal{I}_2 = \frac{2e^{(2\lambda-\kappa)t}}{1+\tilde{a}r} \int_0^t \underbrace{e^{(-2\lambda+\kappa-2\mu)u}}_{=1} du = \frac{2e^{2\lambda t}}{1+\tilde{a}r} e^{-2(\kappa-\nu)t} \frac{1}{2\nu e}.$$

Therefore,

$$\frac{\|\Gamma\|}{\sigma^2} \leq e^{2\lambda t} \left(\frac{1}{(1+\tilde{a}r)^2} \frac{1}{2(\lambda-\mu)} + \frac{e^{-2\min\{\lambda-\mu,\kappa\}t}}{2(\kappa-\lambda+\mu)} + \frac{2}{1+\tilde{a}r} \frac{e^{-2(\kappa-\nu)t}}{2\nu e} \right).$$

Case #3 : $\kappa - \lambda + \mu = 0 \Leftrightarrow \mu = \lambda - \kappa$. Also implying $\kappa - 2\lambda + 2\mu < 0$.

We may keep the terms

$$\mathcal{I}_1 \leq \frac{e^{2\lambda t}}{2(\lambda-\mu)(1+\tilde{a}r)^2} \quad \text{and} \quad \mathcal{I}_2 \leq \frac{2e^{2\lambda t}}{1+\tilde{a}r} \frac{e^{-\min\{2(\lambda-\mu),\kappa\}t}}{\kappa - 2\lambda + 2\mu}.$$

And for the remaining $\mathcal{I}_3$ we find for every $\nu \in (0, \kappa)$ that

$$\mathcal{I}_3 = e^{(2\lambda-2\kappa)t} \int_0^t \underbrace{e^{2(-\lambda+\kappa+\mu)u}}_{=1} du = e^{2\lambda t} e^{-2(\kappa-\nu)t} e^{-2\nu t} t \leq e^{2\lambda t} \frac{e^{-2(\kappa-\nu)t}}{2\nu e}.$$

And hence for arbitrary $\nu \in (0, \kappa)$,

$$\begin{aligned}\frac{\|\Gamma\|}{\sigma^2} &\le \frac{e^{2\lambda t}}{2(\lambda-\mu)(1+\tilde{a}r)^2}\Bigg(1+2(1+\tilde{a}r)\frac{2(\lambda-\mu)e^{-\min\{2(\lambda-\mu),\kappa\}t}}{|2\lambda-2\mu-\kappa|} \\ &\qquad\qquad + \frac{2(\lambda-\mu)(1+\tilde{a}r)^2}{2\nu e}e^{-2(\kappa-\nu)t}\Bigg) \\ &= \frac{e^{2\lambda t}}{2(\lambda-\mu)(1+\tilde{a}r)^2}\left(1+\mathcal{O}\left(\frac{e^{-\min\{2(\lambda-\mu),\kappa-\nu\}t}}{\nu}\right)\right).\end{aligned}$$

This case, $\mu < \lambda$, we still may take over (B.3.2) to receive

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du &\le \frac{(a+be^{-\lambda r})^2}{2(\lambda-\mu)}\underbrace{\left(1-e^{2(\mu-\lambda)s}\right)}_{\le 1}e^{2\lambda s}(t-s)^2 \\ &\le \frac{(a+b^{-\lambda r})^2}{2(\mu-\lambda)}e^{\lambda s}(t-s)^2 \\ &\le \frac{(a+be^{-\lambda r})^2}{2(\mu-\lambda)}e^{2\lambda T}(t-s)^2. \qquad \text{(B.3.13)}\end{aligned}$$

And with the estimate (B.3.1), we get that

$$\begin{aligned}\int_s^t \check{x}^2(t-u)e^{2\mu u}du &\le e^{2\lambda t}e^{(\mu-\lambda)s}\frac{e^{2(\mu-\lambda)(t-s)}-1}{2(\mu-\lambda)} \le e^{2\lambda(t-s)}e^{2\mu s}(t-s) \\ &\le e^{2\lambda t}(t-s). \qquad \text{(B.3.14)}\end{aligned}$$

Therefore,

$$\frac{\mathcal{Q}_1}{\sigma} \le e^{\lambda T}\frac{\sqrt{T}}{\log(p)\sqrt{p}} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \le \frac{a+be^{-\lambda r}}{\sqrt{2(\mu-\lambda)}}e^{\lambda T}\frac{T}{2p\log p}.$$

Collecting the results, we receive that

$$\begin{aligned}\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \le \frac{e^{\lambda t}}{\sqrt{2(\lambda-\mu)}(1+\tilde{a}r)}\Bigg(1+\mathcal{O}\left(\frac{1}{\nu}e^{-\frac{\min\{2(\lambda-\mu),\kappa-\nu\}t}{2}}\right) \\ + \frac{\sqrt{T}}{\log(p)\sqrt{p}} + \frac{a+be^{-\lambda r}}{\sqrt{2(\mu-\lambda)}}e^{\lambda T}\frac{T}{2p\log p}\Bigg)\end{aligned}$$

Instable regime, strong increasing noise ($0 < a < b$, $\mu > \lambda > 0$). *Case #4:* $\lambda < \mu$.

$$\mathcal{I}_1 = \frac{1}{(1+\tilde{a}r)^2}\frac{e^{2\lambda t}-e^{2\mu t}}{2(\mu-\lambda)} \le \frac{e^{2\mu t}}{(1+\tilde{a}r)^2}\frac{1}{2(\mu-\lambda)},$$

$$\begin{aligned}\mathcal{I}_2 = \frac{2e^{(2\lambda-\kappa)t}}{1+\tilde{a}r}\frac{e^{(2\mu-2\lambda+\kappa)t}-1}{\kappa-2\lambda+2\mu} &= \frac{2e^{2\mu t}}{1+\tilde{a}r}e^{2\lambda t}e^{-\kappa t}\frac{e^{\kappa t}e^{-2\lambda t}-e^{-2\mu t}}{2\mu-2\lambda+\kappa} \\ &\le \frac{2e^{2\mu t}}{1+\tilde{a}r}\frac{1-e^{2(\mu-\lambda)t-\kappa t}}{2\mu-2\lambda+\kappa} \\ &\le \frac{2e^{2\mu t}}{1+\tilde{a}r}\frac{1}{2\mu-2\lambda+\kappa}.\end{aligned}$$

$$\mathcal{I}_3 = e^{2(\lambda-\kappa)t}\frac{e^{2(\kappa-\lambda+\mu)t}-1}{2(\kappa-\lambda+\mu)} = e^{2\mu t}\frac{1-e^{-2(\mu-\lambda+\kappa)}}{2(\mu-\lambda+\kappa)} \leq \frac{e^{2\mu t}}{2(\mu-\lambda+\kappa)}.$$

So, we receive

$$\frac{\|\Gamma\|}{\sigma^2} \leq e^{2\mu T}v_0^2, \text{ where}$$
$$v_0^2 := \frac{1}{2(\mu-\lambda)(1+\tilde{a}r)^2} + \frac{2}{(2\mu-2\lambda+\kappa)(1+\tilde{a}r)} + \frac{1}{2(\mu-\lambda+\kappa)}.$$

In this case, we may take over the estimates (B.3.13) and (B.3.14),

$$\int_s^t \check{x}^2(t-u)e^{2\mu u}du \leq e^{2\lambda t}(t-s).$$

$$\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du \leq \frac{(a+be^{-\lambda r})^2}{2(\mu-\lambda)}e^{2\lambda T}(t-s)^2.$$

Therefore,

$$\frac{\mathcal{Q}_1}{\sigma} \leq e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \leq \frac{a+be^{-\lambda r}}{\sqrt{2(\mu-\lambda)}}e^{\lambda T}\frac{T}{2p\log p}.$$

Collecting the results yields

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma} \leq e^{\lambda T}v_0\left(1+\frac{\sqrt{T}}{v_0\sqrt{p}\log p}+\frac{a+be^{-\lambda r}}{v_0\sqrt{2(\mu-\lambda)}}\frac{T}{2p\log p}\right).$$

Instable regime, critical noise $(0 < a < b,\ \mu = \lambda)$. *Case #5:* $\lambda = \mu$.

$$\mathcal{I}_1 = \frac{e^{2\lambda t}}{(1+\tilde{a}r)^2}\int_0^t \underbrace{e^{2(\mu-\lambda)u}}_{=1}du = \frac{e^{2\lambda t}}{(1+\tilde{a}r)^2}t.$$

$$\begin{aligned}\mathcal{I}_2 &= \frac{2e^{2(\lambda-\kappa)t}}{1+\tilde{a}r}\int_0^t e^{(-2\lambda+\kappa+2\mu)u}du = \frac{2e^{2(\lambda-\kappa)t}}{1+\tilde{a}r}\int_0^t e^{\kappa u}du \\ &= \frac{2e^{2(\lambda-\kappa)t}}{(1+\tilde{a}r)}\left(e^{\kappa t}-1\right) \leq \frac{2e^{2\lambda t}}{(1+\tilde{a}r)\kappa}.\end{aligned}$$

$$\mathcal{I}_3 = e^{(2\lambda-2\kappa)t}\int_0^t e^{2(-\lambda+\mu+\kappa)u}du = \frac{e^{2(\lambda-\kappa)t}}{2\kappa}\left(e^{2\kappa t}-1\right) \leq \frac{e^{2\lambda t}}{2\kappa}.$$

Hence,

$$\frac{\|\Gamma\|}{\sigma^2} \leq e^{2\lambda T}\left(\frac{T}{(1+\tilde{a}r)^2}+\frac{1}{(1+\tilde{a}r)\kappa}+\frac{1}{2\kappa}\right) = \frac{e^{2\lambda T}T}{(1+\tilde{a}r)^2}\left(1+\frac{1+\tilde{a}r}{T\kappa}+\frac{(1+\tilde{a}r)^2}{2\kappa t}\right).$$

From (B.3.1), we compute that

$$\begin{aligned}\int_0^s (\check{x}(t-u)-\check{x}(s-u))^2 e^{2\mu u}du &\le \int_0^s (a+be^{-\lambda r})\left(\int_s^t e^{\lambda v}dv\right)e^{2(\mu-\lambda)u}du\\ &= s(a+be^{-\lambda r})^2 e^{2\lambda T}(t-s)^2\\ &= Te^{2\lambda T}(a+be^{-\lambda r})^2(t-s)^2.\end{aligned}$$

And, from (B.3.2), we get

$$\int_s^t \check{x}^2(t-u)e^{2\mu u}du \le \int_s^t e^{2\lambda(t-u)}e^{2\mu u}du = \int_s^t e^{2\lambda t}du \le e^{2\lambda T}(t-s).$$

Therefore,

$$\frac{\mathcal{Q}_1}{\sigma} \le e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} \le e^{\lambda T}(a+be^{-\lambda r})\frac{T^{\frac{3}{2}}}{2p\log p}.$$

Collecting the results,

$$\begin{aligned}&\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma}\\ &\le \frac{\sqrt{T}e^{\lambda T}}{1+\tilde{a}r}\left(1+\sqrt{\frac{1+\tilde{a}r}{T\kappa}}+\frac{1+\tilde{a}r}{\sqrt{2\kappa T}}+\frac{1+\tilde{a}r}{\sqrt{p}\log p}+(1+\tilde{a}r)(a+be^{-\lambda r})\frac{T}{2p\log p}\right).\end{aligned}$$

Instable regime, white noise ($0<a<b$, $\mu=0$). *Case #1:* $\kappa \notin \{\lambda, 2\lambda\}$.

$$\begin{aligned}\int_0^t \check{x}^2(t-u)du &\le \int_0^t \left(\frac{1}{1+\tilde{a}r}+e^{-\kappa(t-u)}\right)^2 e^{2\lambda(t-u)}\\ &= \int_0^t \frac{e^{2\lambda(t-u)}}{(1+\tilde{a}r)^2}du + 2\int_0^t \frac{e^{-\kappa(t-u)}}{1+\tilde{a}r}e^{2\lambda(t-u)}du + \int_0^t e^{2(\lambda-\kappa)(t-u)}du\\ &= \frac{e^{2\lambda t}-1}{2\lambda(1+\tilde{a}r)^2}+\frac{2}{1+\tilde{a}r}\int_0^t e^{(2\lambda-\kappa)(t-u)}du+\frac{e^{2(\lambda-\kappa)t}-1}{2(\lambda-\kappa)}\\ &= \frac{e^{2\lambda t}-1}{2\lambda(1+\tilde{a}r)^2}+\frac{2\left(e^{(2\lambda-\kappa)t}-1\right)}{(1+\tilde{a}r)(2\lambda-\kappa)}+\frac{e^{2(\lambda-\kappa)t}-1}{2(\lambda-\kappa)} \qquad \text{(B.3.15)}\end{aligned}$$

Case #2: $\kappa \in (0,\lambda)$. Starting from (B.3.15)

$$\begin{aligned}\int_0^t \check{x}^2(t-u)du &\le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}+\frac{2e^{(2\lambda-\kappa)t}}{(1+\tilde{a}r)(2\lambda-\kappa)}+\frac{e^{2(\lambda-\kappa)t}}{2(\lambda-\kappa)}\\ &= \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}\left(1+\frac{4\lambda(1+\tilde{a}r)}{2\lambda-\kappa}e^{-\kappa t}+\frac{\lambda(1+\tilde{a}r)^2}{\lambda-\kappa}e^{-2\kappa t}\right).\end{aligned}$$

Case #3: $\kappa \in (\lambda, 2\lambda)$. Beginning from (B.3.15)

$$\begin{aligned}\int_0^t \check{x}^2(t-u)du &\le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}+\frac{2e^{(2\lambda-\kappa)t}}{(1+\tilde{a}r)(2\lambda-\kappa)}+e^{2\lambda t}\frac{e^{-2\kappa t}-e^{-2\lambda t}}{2(\lambda-\kappa)}\\ &\le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}+\frac{2e^{(2\lambda-\kappa)t}}{(1+\tilde{a}r)(2\lambda-\kappa)}+e^{2\lambda t}e^{-2\lambda t}\frac{1-e^{-2(\kappa-\lambda)t}}{2(\kappa-\lambda)}\\ &\le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}\left(1+\frac{2\lambda(1+\tilde{a}r)}{2\lambda-\kappa}e^{-\kappa t}+\frac{2\lambda(1+\tilde{a}r)^2}{2(\kappa-\lambda)}e^{-2\lambda t}\right).\end{aligned}$$

Case #4: $\kappa \in (2\lambda, \infty)$. As before, from (B.3.15) we obtain

$$\int_0^t \breve{x}^2(t-u)du \le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2} + \frac{2\left(1-e^{(\kappa-2\lambda)t}\right)}{(1+\tilde{a}r)(\kappa-2\lambda)} + \frac{1-e^{-2(\kappa-\lambda)t}}{2(\kappa-\lambda)}$$
$$\le \frac{e^{2\lambda t}}{2\lambda(1+\tilde{a}r)^2}\left(1+\frac{4\lambda(1+\tilde{a}r)}{(\kappa-2\lambda)}e^{-2\lambda t} + \frac{\lambda(1+\tilde{a}r)^2}{\kappa-\lambda}e^{-2\lambda t}\right).$$

In all four of the cases, we found out that

$$\frac{\|\Gamma\|}{\sigma^2} \le \frac{e^{2\lambda T}}{2\lambda(1+\tilde{a}r)^2}\left(1+\mathcal{O}\left(e^{-(\kappa\wedge(2\lambda))t}\right)\right).$$

From (B.3.2), we deduce

$$\int_0^s (\breve{x}(t-u)-\breve{x}(s-u))^2 du = \frac{(a+be^{-\lambda r})^2}{2(-\lambda)}\left(e^{2(\mu-\lambda)s}-1\right)e^{2\lambda t}(t-s)^2$$
$$\le \frac{(a+be^{\lambda r})^2}{2\lambda}e^{2\lambda T}(t-s)^2.$$

Starting with (B.3.1), we get

$$\int_s^t \breve{x}^2(t-u)du \le e^{2\lambda t}e^{2(-\lambda)s}\frac{e^{2(-\lambda)(t-s)}-1}{2(-\lambda)} \le e^{2\lambda(t-s)}(t-s) \le e^{2\lambda T}(t-s).$$

And therefore,

$$\frac{\mathcal{Q}_1}{\sigma} = e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p} \quad \text{and} \quad \frac{\mathcal{Q}_2}{\sigma} = e^{\lambda T}\frac{a+be^{\lambda}r}{\sqrt{2\lambda}}\frac{T}{2p\log p}.$$

Collecting the results, we find that

$$\frac{\sqrt{\|\Gamma\|}+Q(T)}{\sigma}$$
$$= \frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde{a}r)}\sqrt{\left(1+\mathcal{O}\left(e^{-(\kappa\wedge(2\lambda))t}\right)\right)} + e^{\lambda T}\frac{\sqrt{T}}{\sqrt{p}\log p} + e^{\lambda T}\frac{a+be^{\lambda r}}{\sqrt{2\lambda}}\frac{T}{2p\log p}$$
$$\le \frac{e^{\lambda T}}{\sqrt{2\lambda}(1+\tilde{a}r)}\left(1+\mathcal{O}\left(e^{-\frac{(\kappa\wedge(2\lambda))t}{2}}\right) + \sqrt{2\lambda}(1+\tilde{a}r)\left(\frac{\sqrt{T}}{\sqrt{p}\log p} + \frac{a+be^{\lambda r}}{\sqrt{2\lambda}}\frac{T}{2p\log p}\right)\right).$$

C. REFERENCES

References

[AB10] John A. D. Appleby and Evelyn Buckwar. A constructive comparison technique for determining the asymptotic behaviour of linear functional differential equations with unbounded delay. *Differ. Equ. Dyn. Syst.*, 18(3):271–301, 2010.

[AGR06] John A. D. Appleby, István Győri, and David W. Reynolds. On exact convergence rates for solutions of linear systems of Volterra difference equations. *J. Difference Equ. Appl.*, 12(12):1257–1275, 2006.

[AGR11] John A. D. Appleby, István Győri, and David W. Reynolds. History-dependent decay rates for a logistic equation with infinite delay. *Proc. Roy. Soc. Edinburgh Sect. A*, 141(1):23–44, 2011.

[AHMP07] Mercedes Arriojas, Yaozhong Hu, Salah-Eldin Mohammed, and Gyula Pap. A delayed Black and Scholes formula. *Stoch. Anal. Appl.*, 25(2):471–492, 2007.

[AI05] V. Anh and A. Inoue. Financial markets with memory. I. Dynamic models. *Stoch. Anal. Appl.*, 23(2):275–300, 2005.

[AIK05] V. Anh, A. Inoue, and Y. Kasahara. Financial markets with memory. II. Innovation processes and expected utility maximization. *Stoch. Anal. Appl.*, 23(2):301–328, 2005.

[AKM05] A. Ashoorioon, A. Kempf, and R. B. Mann. Minimum length cutoff in inflation and uniqueness of the action (11 pages). *PHYSICAL REVIEW -SERIES D-, 2005, 71, 2, 023503*, 2005.

[AMW10] John A. D. Appleby, Xuerong Mao, and Huizhong Wu. On the almost sure running maxima of solutions of affine stochastic functional differential equations. *SIAM J. Math. Anal.*, 42(2):646–678, 2010.

[AP15] John A. D. Appleby and Denis D. Patterson. Subexponential growth rates in functional differential equations. *Discrete Contin. Dyn. Syst.*, (Dynamical systems, differential equations and applications. 10th AIMS Conference. Suppl.):56–65, 2015.

[AP17] John A. D. Appleby and Denis D. Patterson. Hartman-Wintner growth results for sublinear functional differential equations. *Electron. J. Differential Equations*, pages Paper No. 21, 45, 2017.

[ARS13] John A. D. Appleby, Markus Riedle, and Catherine Swords. Bubbles and crashes in a Black-Scholes model with delay. *Finance Stoch.*, 17(1):1–30, 2013.

[AT07] Robert J. Adler and Jonathan E. Taylor. *Random fields and geometry.* Springer monographs in mathematics. Springer, New York, NY, 2007.

[Bau96] Heinz Bauer. *Probability theory*, volume 23 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1996. Translated from the fourth (1991) German edition by Robert B. Burckel and revised by the author.

[BB00] Christopher T. H. Baker and Evelyn Buckwar. Numerical analysis of explicit one-step methods for stochastic delay differential equations. *LMS J. Comput. Math.*, 3:315–335, 2000.

[BC94] Jacques Bélair and Sue Ann Campbell. Stability and bifurcations of equilibria in a multiple-delayed differential equation. *SIAM J. Appl. Math.*, 54(5):1402–1424, 1994.

[BDS01] Rodrigo Banuelos, R. Dante DeBlassie, and Robert Smits. The first exit time of planar brownian motion from the interior of a parabola. *The Annals of Probability, 2001, 29, 2, 882*, 2001.

[Ber64] Simeon M. Berman. Limit theorems for the maximum term in stationary sequences. *Annals of Mathematical Statistics, 1964, 35, 502*, 1964.

[Ber71a] Simeon M. Berman. Excursions above high levels for stationary gaussian processes. *Pacific Journal of Mathematics, 1971, 36, 63*, 1971.

[Ber71b] Simeon M. Berman. Maxima and high level excursions of stationary gaussian processes. *Transactions of the American Mathematical Society, 1971, 160, 65*, 1971.

[Ber72a] Simeon M. Berman. A class of limiting distributions of high level excursions of Gaussian processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 21:121–134, 1972.

[Ber72b] Simeon M. Berman. Maximum and high level excursion of a gaussian process with stationary increments. *Annals of Mathematical Statistics, 1972, 43, 1247*, 1972.

[BG06] Nils Berglund and Barbara Gentz. *Noise-induced phenomena in slow-fast dynamical systems*. Probability and its applications. Springer, New York [u.a.], 2006.

[BKMS08] Evelyn Buckwar, Rachel Kuske, Salah-Eldin Mohammed, and Tony Shardlow. Weak convergence of the Euler scheme for stochastic differential delay equations. *LMS J. Comput. Math.*, 11:60–99, 2008.

[BM16] Mouffak Benchohra and Kadda Mazouz. Existence and uniqueness results for nonlinear implicit fractional differential equations with delay in Fréchet spaces. *Comm. Appl. Nonlinear Anal.*, 23(4):48–59, 2016.

[BMS01] Steve Blythe, Xuerong Mao, and Anita Shah. Razumikhin-type theorems on stability of stochastic neural networks with delays. *Stochastic Anal. Appl.*, 19(1):85–101, 2001.

[Buc04] Evelyn Buckwar. The Θ-Maruyama scheme for stochastic functional differential equations with distributed memory term. *Monte Carlo Methods Appl.*, 10(3-4):235–244, 2004.

[Buc06] Evelyn Buckwar. One-step approximations for stochastic functional differential equations. *Appl. Numer. Math.*, 56(5):667–681, 2006.

[BY15] Jianhai Bao and Chenggui Yuan. Large deviations for neutral functional SDEs with jumps. *Stochastics*, 87(1):48–70, 2015.

[Cai03] Jean-Michel Caillol. Some applications of the Lambert W function to classical statistical mechanics. *J. Phys. A*, 36(42):10431–10442, 2003.

[CGHJ93] Robert M. Corless, Gaston H. Gonnet, D. E. G. Hare, and David J. Jeffrey. Lambert's w function in maple. *http://www.apmaths.uwo.ca/ rcorless/papers/LambertW/mtnw.ps.Z*, 1993.

[Cra07] Steven R. Cranmer. New views of the solar wind with the lambert w function. 2004-06-07.

[CT62] Z. Ciesielski and S. J. Taylor. First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path. *Trans. Amer. Math. Soc.*, 103:434–450, 1962.

[CYB05] Sue Ann Campbell, Yuan Yuan, and Sharene D. Bungay. Equivariant Hopf bifurcation in a ring of identical cells with delayed coupling. *Nonlinearity*, 18(6):2827–2846, 2005.

[Del65] Jean Delporte. Fonctions aléatoires presque sûrement continues sur un intervalle fermé. *Ann. Inst. H. Poincaré Sect. B*, 1:111–215, 1964/1965.

[DK92] A. D. Drozdov and V. B. Kolmanovskiĭ. Stochastic stability of viscoelastic bars. *Stochastic Anal. Appl.*, 10(3):265–276, 1992.

[DKW56] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.*, 27:642–669, 1956.

[DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of mathematics and its applications ; 152*. Cambridge Univ. Press, Cambridge, 2. ed. edition, 2014.

[DvGVLW95] Odo Diekmann, Stephan A. van Gils, Sjoerd M. Verduyn Lunel, and Hans-Otto Walther. *Delay equations*, volume 110 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[DZ92] Amir Dembo and Ofer Zeitouni. *Large deviations techniques*. Jones and Bartlett books in mathematics. Jones and Bartlett, Boston [u.a.], 1. [dr.] edition, 1992.

[Els99] S. A. Elsanousi. Lyapunov exponents of linear stochastic functional-differential equations with infinite memory. *Stochastics Stochastics Rep.*, 65(3-4):153–176, 1999.

[Ern09] Thomas Erneux. *Applied delay differential equations*, volume 3 of *Surveys and tutorials in the applied mathematical sciences ; 3*. Springer, New York, NY, 2009.

[Fen79] Neil Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979.

[Fer64] Xavier Fernique. Continuité des processus gaussiens. *CR Acad. Sci. Paris*, 258:6058–6060, 1964.

[Fer75] X. Fernique. Des résultats nouveaux sur les processus gaussiens. pages 318–335. Lecture Notes in Math., Vol. 465, 1975.

[Fer90] X. Fernique. Régularité de fonctions aléatoires gaussiennes a valeurs vectorielles. *The Annals of Probability, 1990, 18, 4, 1739*, 1990.

[FI05] Markus Fischer and Peter Imkeller. A two-state model for noise-induced resonance in bistable systems with delay. *Stoch. Dyn.*, 5(2):247–270, 2005.

[FN09] Neville J. Ford and Stewart J. Norton. Noise-induced changes to the behaviour of semi-implicit Euler methods for stochastic delay differential equations undergoing bifurcation. *J. Comput. Appl. Math.*, 229(2):462–470, 2009.

[Fre12] Mark I. Freidlin. *Random Perturbations of Dynamical Systems*. Die Grundlehren der mathematischen Wissenschaften, A Series of Comprehensive Studies in Mathematics ; 260. Springer Berlin Heidelberg, Berlin, Heidelberg, 3rd ed. 2012 edition, 2012.

[Gau15] Carl Friedrich (1777-1855) Gauss. Methodus nova integralium valores per approximationem inveniendi auctore carolo friderico gauss. *Bibliothèque nationale de France*, 1815.

[GFF17] Valery Y. Glizer, Emilia Fridman, and Yuri Feigin. A novel approach to exact slow-fast decomposition of linear singularly perturbed systems with small delays. *SIAM J. Control Optim.*, 55(1):236–274, 2017.

[Gir29] Albert (1595-1632) Girard. Invention nouvelle en l'algèbre , par albert girard,... 1629.

[Gra53] I. S. Gradšteĭn. Application of A. M. Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives. *Mat. Sbornik N. S.*, 32(74):263–286, 1953.

[Gum67] Emil Julius Gumbel. *Statistics of extremes*. Columbia Univ. Press, New York [u.a.], 4. print. edition, 1967.

[Hel01] D Helbing. Traffic and related self-driven many-particle systems. *REVIEWS OF MODERN PHYSICS, 2001, 73, 4, 1067*, 2001.

[HKGS82] F. A. Hopf, D. L. Kaplan, H. M. Gibbs, and R. L. Shoemaker. Bifurcations to chaos in optical bistability. *Phys. Rev. A*, 25:2172–2182, Apr 1982.

[HMY04] Yaozhong Hu, Salah-Eldin A. Mohammed, and Feng Yan. Discrete-time approximations of stochastic delay equations: the Milstein scheme. *Ann. Probab.*, 32(1A):265–314, 2004.

[HP14] Ferenc Hartung and Mihály Pituk, editors. *Recent advances in delay differential and difference equations*, volume 94 of *Springer Proceedings in Mathematics & Statistics*. Springer, Cham, 2014.

[HVL93] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to functional differential equations*, volume 99 of *Applied mathematical sciences ; 99*. Springer, New York [u.a.], 1993.

[IKS03] A. F. Ivanov, Y. I. Kazmerchuk, and A. V. Swishchuk. Theory, stochastic stability and applications of stochastic delay differential equations: a survey of results. *Differential Equations Dynam. Systems*, 11(1-2):55–115, 2003.

[Jac79] Jean Jacod. *Calcul stochastique et problèmes de martingales*, volume 714 : Series: Institut de Mathématiques, Université de Strasbourg of *Lecture notes in mathematics ; 714 : Series: Institut de Mathématiques, Université de Strasbourg*. Springer, Berlin [u.a.], 1979.

[KÖ4] Konrad Königsberger. *Analysis. 1.* Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, sixth edition, 2004.

[Kim16] Young-Ho Kim. Euler-Maruyama's approximate solutions to stochastic differential delay equation. *J. Nonlinear Convex Anal.*, 17(1):185–195, 2016.

[Kle14] Achim Klenke. *Probability theory*. Universitext. Springer, London, second edition, 2014. A comprehensive course.

[KM92] Uwe Küchler and Beatrice Mensch. Langevin's stochastic differential equation extended by a time-delayed term. *Stochastics Stochastics Rep.*, 40(1-2):23–42, 1992.

[KP07] Uwe KÃ$\frac{1}{4}$chler and Eckhard Platen. Time delay and noise explaining cyclical fluctuations in prices of commodities. *Research paper, 2007*, 2007.

[KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[KS12] P. E. Kloeden and T. Shardlow. The Milstein scheme for stochastic delay differential equations without using anticipative calculus. *Stoch. Anal. Appl.*, 30(2):181–202, 2012.

[Kue15] Christian Kuehn. *Multiple Time Scale Dynamics*. Applied Mathematical Sciences ; 191. Springer International Publishing, Cham, 2015.

[Lak11] Muthusamy Lakshmanan. *Dynamics of Nonlinear Time-Delay Systems*. Springer Series in Synergetics. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2011.

[Li99] W. V. Li. Small deviations for gaussian markov processes under the sup-norm. *Journal of Theoretical Probability*, 12(4):971–984, 1999.

[Li03] Wenbo V. Li. The first exit time of a brownian motion from an unbounded convex domain. *The Annals of Probability, 2003, 31, 2, 1078*, 2003.

[LODO91] R. Langevin, W. M. Oliva, and J. C. F. De Oliveira. Retarded functional-differential equations with white noise perturbations. *Annales de l'Institut Henri Poincaré. Physique Théorique, 1991, 55, 2, 671*, 1991.

[Lon10] André Longtin. Stochastic delay-differential equations. In *Complex time-delay systems*, Underst. Complex Syst., pages 177–195. Springer, Berlin, 2010.

[LS01] W.V. Li and Q.-M. Shao. Gaussian processes: Inequalities, small ball probabilities and applications. *Handbook of Statistics, 2001, 19, 533*, 2001.

[LS02] Mikhail Lifshits and Zhan Shi. The first exit time of brownian motion from a parabolic domain. *Bernoulli, 2002, 8, 6, 745*, 2002.

[LZ14] Kai Liu and Tusheng Zhang. A large deviation principle of retarded Ornstein-Uhlenbeck processes driven by Lévy noise. *Stoch. Anal. Appl.*, 32(5):889–910, 2014.

[Man06] Kenneth Manders. Algebra in roth, faulhaber, and descartes. *Historia Mathematica, 2006, 33, 2, 184*, 2006.

[Mao03] Xuerong Mao. Numerical solutions of stochastic functional differential equations. *LMS J. Comput. Math.*, 6:141–161, 2003.

[Mao07] Xuerong Mao. Exponential stability of equidistant Euler-Maruyama approximations of stochastic differential delay equations. *J. Comput. Appl. Math.*, 200(1):297–316, 2007.

[Mao08] Xuerong Mao. *Stochastic differential equations and their applications*. Horwood, Chichester, 2. ed. edition, 2008.

[Mar70] Michael Marcus. A bound for the distribution of the maximum of continuous gaussian processes. *Annals of Mathematical Statistics, 1970, 41, 305*, 1970.

[Mar72] M. B. Marcus. Upper bounds for the asymptotic maxima of continuous Gaussian processes. *Ann. Math. Statist.*, 43:522–533, 1972.

[Moh84] Salah-Eldin A. Mohammed. *Stochastic functional differential equations*, volume 99 of *Research notes in mathematics ; 99*. Pitman, Boston [u.a.], 1984.

[MS72] M. B. Marcus and L. A. Shepp. Sample behavior of gaussian processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 423–441, Berkeley, Calif., 1972. University of California Press.

[MS90] Salah Eldin A. Mohammed and Michael K. R. Scheutzow. Lyapunov exponents and stationary solutions for affine stochastic delay equations. *Stochastics Stochastics Rep.*, 29(2):259–283, 1990.

[MS96] Salah-Eldin A. Mohammed and Michael K. R. Scheutzow. Lyapunov exponents of linear stochastic functional differential equations driven by semimartingales. I. The multiplicative ergodic theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(1):69–105, 1996.

[MS97] Salah-Eldin A. Mohammed and Michael K. R. Scheutzow. Lyapunov exponents of linear stochastic functional-differential equations. II. Examples and case studies. *Ann. Probab.*, 25(3):1210–1240, 1997.

[MZ06] Salah-Eldin A. Mohammed and Tusheng Zhang. Large deviations for stochastic systems with memory. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):881–893, 2006.

[Nis67] Makiko Nisio. On the extreme values of gaussian processes. *Osaka Journal of Mathematics, 1967, 4, 313*, 1967.

[NV13] Hoi H. Nguyen and Van H. Vu. Small ball probability, inverse theorems, and applications. *Erdös centennial, 2013*, 2013.

[Pic67] James Pickands, III. Maxima of stationary Gaussian processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7:190–223, 1967.

[Pic69] James Pickands, III. Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.*, 145:75–86, 1969.

[Rol90] Michel Rolle. Traité d'algèbre ou principes généraux pour résoudre les questions de mathématique. 1690.

[Rou96] Marc R. Roussel. The use of delay differential equations in chemical kinetics. *The journal of physical chemistry, 1996, 100, 20, 8323*, 1996.

[RY05] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften ; 293.* Springer, Berlin [u.a.], corr. 3. print. of the 3. ed. edition, 2005.

[SBDM93] T. C. Scott, J. F. Babb, A. Dalgarno, and John D. Morgan. The calculation of exchange forces: General results and specific models. *Journal of Chemical Physics, 1993, 99, 4, 2841*, 1993.

[Sch84] M. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics and Stochastics Reports*, 12:41–80, 1984.

[Sch05] Michael Scheutzow. Exponential growth rates for stochastic delay differential equations. *Stoch. Dyn.*, 5(2):163–174, 2005.

[Sch13] Michael Scheutzow. Exponential growth rate for a singular linear stochastic delay differential equation. *Discrete Contin. Dyn. Syst. Ser. B*, 18(6):1683–1696, 2013.

[SMM06] Tony C. Scott, Robert Mann, and Roberto E. Martinez, II. General relativity and quantum mechanics: towards a generalization of the Lambert W function: a generalization of the Lambert W function. *Appl. Algebra Engrg. Comm. Comput.*, 17(1):41–47, 2006.

[SN10] Rifat Sipahi and Silviu-Iulian Niculescu. Deterministic time-delayed traffic flow models: a survey. *Complex time-delay systems, 2010*, 2010.

[SV06] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[Tih52] A. N. Tihonov. Systems of differential equations containing small parameters in the derivatives. *Mat. Sbornik N. S.*, 31(73):575–586, 1952.

[TKBM15] Antoine Tambue, Elisabeth Kemajou Brown, and Salah Mohammed. A stochastic delay model for pricing debt and equity: numerical techniques and applications. *Commun. Nonlinear Sci. Numer. Simul.*, 20(1):281–297, 2015.

[vC15] Jan ˇ Cermák. Stability conditions for linear delay difference equations: a survey and perspective. *Tatra Mt. Math. Publ.*, 63:1–29, 2015.

[vRS10] Max-K. von Renesse and Michael Scheutzow. Existence and uniqueness of solutions of stochastic functional differential equations. *Random Oper. Stoch. Equ.*, 18(3):267–284, 2010.

[Wat54] G. S. Watson. Extreme values in samples from m-dependent stationary stochastic processes. *Annals of Mathematical Statistics, 1954, 25, 798*, 1954.

[Wei82] K. Weierstrass. Neuer beweis des satzes, dass jede ganze rationale function einer veranderlichen dargestellt werden kann als ein product aus linearen functionen derselben veraderlichen. 1882.

[WYM17] Fuke Wu, George Yin, and Hongwei Mei. Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity. *J. Differential Equations*, 262(3):1226–1252, 2017.

[YB11] Yuan Yuan and Jacques Bélair. Stability and Hopf bifurcation analysis for functional differential equation with distributed delay. *SIAM J. Appl. Dyn. Syst.*, 10(2):551–581, 2011.

[ZAL+17] Xianmin Zhang, Praveen Agarwal, Zuohua Liu, Hui Peng, Fang You, and Yajun Zhu. Existence and uniqueness of solutions for stochastic differential equations of fractional-order with finite delays. *Adv. Difference Equ.*, page 2017:123, 2017.

[Zhe15] Yun Zheng. *Asset pricing based on stochastic delay differential equations*. ProQuest LLC, Ann Arbor, MI, 2015. Thesis (Ph.D.)–Iowa State University.