DOI: 10.1214/EJP.v9-232

Electronic Journal of Probability

Vol. 9 (2004), Paper no. 28, pages 862-886.

Journal URL
http://www.math.washington.edu/∼ejpecp/

On Some Degenerate Large Deviation Problems

Anatolii A. Puhalskii

Mathematics Department, University of Colorado at Denver,
Campus Box 170, P.O. Box 173364, Denver, CO 80203, U.S.A.
and
Institute for Problems in Information Transmission, Moscow, Russia

Email: puhalski@math.cudenver.edu

Abstract

This paper concerns the issue of obtaining the large deviation principle for solutions of stochastic equations with possibly degenerate coefficients. Specifically, we explore the potential of the methodology that consists in establishing exponential tightness and identifying the action functional via a maxingale problem. In the author's earlier work it has been demonstrated that certain convergence properties of the predictable characteristics of semimartingales ensure both that exponential tightness holds and that every large deviation accumulation point is a solution to a maxingale problem. The focus here is on the uniqueness for the maxingale problem. It is first shown that under certain continuity hypotheses existence and uniqueness of a solution to a maxingale problem of diffusion type are equivalent to Luzin weak existence and uniqueness, respectively, for the associated idempotent Ito equation. Consequently, if the idempotent equation has a unique Luzin weak solution, then the action functional is specified uniquely, so the large deviation principle follows. Two kinds of application are considered. Firstly, we obtain results on the logarithmic asymptotics of moderate deviations for stochastic equations with possibly degenerate diffusion coefficients which, as compared with earlier results, relax the growth conditions on the coefficients, permit certain non-Lipshitz-continuous coefficients, and allow the coefficients to depend on the entire past of the process and to be discontinuous functions of time. The other application concerns multiple-server queues with impatient customers.

Keywords and phrases: large deviations, moderate deviations, large deviation principle, diffusion processes, Ito equations, Freidlin-Wentzell theory, queues in heavy traffic

MSC 2000 subject classifications: primary 60F10, secondary 60J60, 60K25

Submitted to EJP on May 19, 2004. The final version accepted on December 14, 2004.

1 Introduction

There are certain difficulties in establishing the large deviation principle (LDP) for solutions of stochastic equations with degenerate coefficients. As a motivating example, let us consider the setting of a d-dimensional diffusion process with small noise

$$dX^\epsilon_t = b(X^\epsilon_t)\,dt + \sqrt{\epsilon}\,\sigma(X^\epsilon_t)\,dB^\epsilon_t,\; X^\epsilon_0 = x_0, \tag{1.1}$$

where $\epsilon \downarrow 0$, the $B^\epsilon = (B^\epsilon_t,\, t \in \mathbb{R}_+)$ are standard d-dimensional Wiener processes, functions $b(x)$ and $\sigma(x)$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, respectively, are bounded and continuous in x. If the diffusion matrix $c(x) = \sigma(x)\sigma(x)^T$, where ^{T} denotes transposition, is uniformly elliptic and the functions $b(x)$ and $\sigma(x)$ are, in addition, uniformly continuous, a seminal result by Freidlin and Wentzell [18] states that the processes $X^\epsilon = (X^\epsilon_t,\, t \in \mathbb{R}_+)$ obey the LDP in the metric space $\mathbb{C}$ of $\mathbb{R}^d$-valued continuous functions on $\mathbb{R}_+$ with compact-open topology. (Strictly speaking, Freidlin and Wentzell's original result concerns the space of continuous functions defined on the $[0,1]$ interval, however, the extension to $\mathbb{C}$ is routine.) The action functional for absolutely continuous functions $\mathbf{x} = (\mathbf{x}_t,\, t \in \mathbb{R}_+)$ with $\mathbf{x}_0 = x_0$ is given by

$$\mathbf{I}(\mathbf{x}) = \frac{1}{2}\int_0^\infty \big(\dot{\mathbf{x}}_t - b(\mathbf{x}_t)\big)\cdot c(\mathbf{x}_t)^{-1}\big(\dot{\mathbf{x}}_t - b(\mathbf{x}_t)\big)\,dt, \tag{1.2}$$

where $\dot{\mathbf{x}}$ denotes the derivative and $u \cdot v$ denotes the inner product of vectors u and v. For all other $\mathbf{x}$, $\mathbf{I}(\mathbf{x}) = \infty$. The proof employs the change-of-measure technique. If one retains some form of the ellipticity condition, this LDP can be generalised to the settings of diffusions with jumps, which was already done in Freidlin and Wentzell [18], and of continuous in x unbounded coefficients, depending on ϵ, time and the past history, see Dupuis and Ellis [11], Feng [14], Feng and Kurtz [15], Friedman [19], Liptser and Pukhalskii [27], Mikami [29], Narita [31], Stroock [40], and Wentzell [43]. Certain cases of a discontinuous drift coefficient can also be analysed, see, e.g., Boué, Dupuis and Ellis [3], Korostelëv and Leonov [24, 25].

Allowing $c(x)$ to be degenerate complicates matters significantly. The large deviation lower bound can no longer be obtained by a straightforward application of a change of measure. It is possible to prove, however, that if one requires, in addition, that $b(x)$ and $c(x)^{1/2}$ be Lipshitz continuous or locally Lipshitz continuous, then the above LDP holds without an ellipticity requirement on the diffusion matrix. The action functional is still given by (1.1), where the inverse of $c(x)$ has to be replaced with the pseudo-inverse, provided $\mathbf{x}$ is absolutely continuous with $\mathbf{x}_0 = x_0$ and $\dot{\mathbf{x}}_t - b(\mathbf{x}_t)$ belongs to the range of $c(\mathbf{x}_t)$ a.e., and $\mathbf{I}(\mathbf{x})$ is equal to infinity otherwise. This version of the Freidlin-Wentzell LDP has been established under various additional assumptions and various techniques have been used.

Cutland [6] invokes the apparatus of infinitesimal analysis and is able to tackle non-time-homogeneous diffusions with bounded Lipshitz-continuous-in-the-space-variable coefficients depending on the past. The other results are confined to the time-homogeneous Markov setting. Azencott [1] considers the case of locally Lipshitz continuous b and continuously differentiable σ. No growth conditions are assumed, so diffusions with explosion are also allowed. Baldi and Chaleyat-Maurel [2], Dembo and Zeitouni [10], and Feng [14] analyse the case of diffusion processes with bounded Lipshitz-continuous coefficients. The first two articles use discretization considerations, while Feng [14] applies methods of nonlinear semigroup convergence. Azencott [1], Baldi and Chaleyat-Maurel [2], and Feng [14] also allow the coefficients to depend on ϵ. (Baldi and Chaleyat-Maurel [2] actually state the LDP for locally Lipshitz-continuous coefficients with no growth conditions, requiring instead that the equation $\dot{g}_t = b(g_t) + \sigma(g_t)\,\dot{f}_t,\;\; g_0 = x_0$, have a solution for every function f with a locally square-integrable derivative. However, the authors do not provide enough substantiation for the extension from the case of bounded Lipshitz continuous coefficients to the case of locally Lipshitz continuous coefficients.) In recent work Feng and Kurtz [15] improve on the results of Feng [14] and derive the LDP for diffusions with jumps having Lipshitz-continuous coefficients satisfying the linear growth condition. Along with discretization, a tool often used for tackling the case of degenerate diffusions is the perturbation approach dating back to Varadhan [42]. For the setting of (1.1), the approach consists in introducing a small nonsingular diffusion term and taking a limit in the lower bound for the perturbed equation as the perturbation's magnitude tends to zero. Dupuis and Ellis [11] combine the perturbation approach with discretization methods and representation formulas. Their results cover the case of bounded Lipshitz continuous coefficients. Liptser [26] applies the perturbation approach to study diffusions with averaging, which have Lipshitz continuous coefficients. De Acosta [9], who considers diffusions with jumps and bounded Lipshitz continuous coefficients, invokes both the perturbation approach and discretization arguments as well as a Banach-space version of the method of stochastic exponentials, cf., Puhalskii [33, 35, 36]. An interesting example of an equation with a non-Lipshitz-continuous degenerate diffusion coefficient is considered by Dawson and Feng [7, 8], Feng [16], and Feng and Xiong [17], who are motivated by the study of Fleming-Viot processes. The authors obtain the LDP for a particular form of equation (1.1) with the (i,j) entry of $c(x)$ given by $x_i(\delta_{ij} - x_j)$, where $x = (x_1, \ldots, x_d)^T$, the entries being in $[0,1]$, and δ_{ij} is Kronecker's delta. To sum up, we are aware of no results for equation (1.1) that would simultaneously allow the coefficients to be degenerate, to grow at linear rate, to be time-dependent, and to be discontinuous in the time variable. Nor are there, to our knowledge, any results available that combine the hypotheses that the coefficients depend on the past history, are possibly degenerate and unbounded.

In Puhalskii [32, 35, 36], we introduced and developed an approach to establishing the LDP for semimartingales, which, in analogy with the martingale problem method in the weak convergence theory, reduces the task to establishing exponential tightness and identifying the action functional via *a maxingale problem*. The establishing of exponential tightness is often routine. In particular, certain convergence conditions on the predictable characteristics ensure both that exponential tightness holds and that every large deviation accumulation point is a solution to the maxingale problem. The difficult part, as in the weak convergence theory, is proving uniqueness of a solution to the maxingale problem. The uniqueness is known to hold if certain nondegeneracy conditions are met, Puhalskii [36, Section 2.8]. The motivation for this article has been to incorporate degenerate maxingale problems of diffusion type.

Attaining this objective puts us in a position to derive results on the LDP for a range of large deviation settings that lead to a maxingale problem of diffusion type. In particular, the gaps mentioned above can be filled. For equation (1.1) we are able to tackle the case of possibly degenerate coefficients, which are locally Lipshitz-continuous in the space variable, satisfy the linear-growth condition and depend on the entire past of the process. Besides, the time-non-homogeneous Markov setting of the coefficients measurably depending on time is incorporated. In addition, the coefficients may depend on ϵ and grow slightly faster than linearly, say, at rate $|x|\log(1+|x|)$. For time-homogeneous non-degenerate diffusions the latter possibility has been known since the work of Narita [31]. If the drift is directed towards the origin, then no restrictions are needed on the growth rate of the drift coefficient, which is consistent with what is known for stochastic differential equations, see, e.g., Ethier and Kurtz [12]. Slightly non-Lipshitz behaviour of the coefficients is also permitted, e.g., in the one-dimensional setting one can take $\sigma(x) = x\log|x|$. We are not able, though, to obtain by our methods the results of Dawson and Feng [7, 8], Feng [16], and Feng and Xiong [17].

In fact, our results concern a more general setting than that of (1.1) by allowing "compensated"

Poisson jumps. The jump sizes are assumed to be large as compared with time periods between the jumps, so in the end the jump part also contributes to "the limiting c" in (1.2). It is thus a form of moderate deviation asymptotics. Moreover, we also obtain results on another type of moderate deviation asymptotics, where "the limiting b" is a linear function. For nondegenerate c both types of asymptotics have been analysed by Wentzell [43]. No results for possibly degenerate c are known to us. (The results of de Acosta [9], Dupuis and Ellis [11], and Feng and Kurtz [15] concern large deviation asymptotics when the time spans between the jumps are of the same order of magnitude as the jump sizes.) Finally, we derive moderate deviation asymptotics in the form of the LDP for the queueing model of a multiple-server system with impatient customers.

Before explaining how the outlined results are obtained, we recall the main steps of the maxingale problem approach by continuing with equation (1.1). This should help the interested reader to follow the developments below. We still assume that $b(x)$ and $c(x)$ are continuous and bounded functions. Let $\mathbf{E}$ denote expectation. Given $\lambda \in \mathbb{R}^d$, the stochastic processes $\big((\exp\big(\lambda \cdot X^\epsilon_t - \int_0^t \lambda \cdot b(X^\epsilon_s)\,ds - \epsilon/2 \int_0^t \lambda \cdot c(X^\epsilon_s)\lambda\,ds\big)\big),\ t \in \mathbb{R}_+\big)$ are martingales relative to the associated filtrations, so on taking λ/ϵ as λ we have that for an arbitrary bounded function $f(\mathbf{x})$ on $\mathbb{C}$ that is measurable with respect to the σ-algebra on $\mathbb{C}$ generated by the projections $\mathbf{x} \to \mathbf{x}_r$, where $r \le s \le t$, the following equality holds

$$\mathbf{E}\Big(\exp\big(\frac{1}{\epsilon}\,\lambda \cdot X^\epsilon_t - \frac{1}{\epsilon}\int_0^t \lambda \cdot b(X^\epsilon_q)\,dq - \frac{1}{2\epsilon}\int_0^t \lambda \cdot c(X^\epsilon_q)\lambda\,dq\big) f(X^\epsilon)^{1/\epsilon}\Big)$$
$$= \mathbf{E}\Big(\exp\big(\frac{1}{\epsilon}\,\lambda \cdot X^\epsilon_s - \frac{1}{\epsilon}\int_0^s \lambda \cdot b(X^\epsilon_q)\,dq - \frac{1}{2\epsilon}\int_0^s \lambda \cdot c(X^\epsilon_q)\lambda\,dq\big) f(X^\epsilon)^{1/\epsilon}\Big). \quad (1.3)$$

Boundedness of $b(x)$ and $c(x)$ implies the exponential uniform integrability condition that for $r \in \mathbb{R}_+$

$$\lim_{A\to\infty} \limsup_{\epsilon\to 0} \Big(\mathbf{E}\big(\exp(\frac{1}{\epsilon}\,\lambda \cdot X^\epsilon_r)\mathbf{1}(|X^\epsilon_r| > A)\big)\Big)^\epsilon = 0,$$

where $\mathbf{1}(\Gamma)$ denotes the indicator function of an event Γ. The boundedness also makes it straightforward to see that the net of laws of the X^ϵ is exponentially tight, so there exists a subnet obeying the LDP with some (tight) action functional.

Let us assume, in addition, that the function f is continuous. By a version of Varadhan's lemma (see, e.g., Dembo and Zeitouni [10]) on raising both sides of (1.3) to the power of ϵ and letting $\epsilon \downarrow 0$ we conclude that if $\mathbf{I}(\mathbf{x})$ is the action functional in the LDP for some subnet, then

$$\sup_{\mathbf{x}\in\mathbb{C}}\Big(\exp\big(\lambda \cdot \mathbf{x}_t - \int_0^t \lambda \cdot b(\mathbf{x}_q)\,dq - \frac{1}{2}\int_0^t \lambda \cdot c(\mathbf{x}_q)\lambda\,dq\big) f(\mathbf{x}) \exp(-\mathbf{I}(\mathbf{x}))\Big)$$
$$= \sup_{\mathbf{x}\in\mathbb{C}}\Big(\exp\big(\lambda \cdot \mathbf{x}_s - \int_0^s \lambda \cdot b(\mathbf{x}_q)\,dq - \frac{1}{2}\int_0^s \lambda \cdot c(\mathbf{x}_q)\lambda\,dq\big) f(\mathbf{x}) \exp(-\mathbf{I}(\mathbf{x}))\Big). \quad (1.4)$$

If we treat the set function $\mathbf{\Pi}(\Gamma) = \sup_{\mathbf{x}\in\Gamma} \exp(-\mathbf{I}(\mathbf{x}))$, $\Gamma \subset \mathbb{C}$, as a maxitive, or idempotent, analogue of probability, then equality (1.4) represents a maxitive analogue of one of characterisations of martingales, so we call the function $\big((\exp\big(\lambda\cdot\mathbf{x}_t - \int_0^t \lambda\cdot b(\mathbf{x}_s)\,ds - 1/2\int_0^t \lambda\cdot c(\mathbf{x}_s)\lambda\,ds\big)\big),\ t \in \mathbb{R}_+,\ \mathbf{x} \in \mathbb{C}\big)$ *an exponential maxingale*. The problem of identifying $\mathbf{I}$ from (1.4) is accordingly referred to as a maxingale problem. If $\mathbf{I}$ is specified uniquely by (1.4), then the LDP holds with action functional

$\mathbf{I}$. In fact, one can show that (1.4) is equivalent to an idempotent analogue of the definition of a martingale: for arbitrary $\hat{\mathbf{x}} \in \mathbb{C}$

$$\sup_{\substack{\mathbf{x}\in\mathbb{C}:\\ \mathbf{x}_q=\hat{\mathbf{x}}_q,\, q\le s}} \bigl(\lambda\cdot\mathbf{x}_t - \int_0^t \lambda\cdot b(\mathbf{x}_q)\,dq - \frac{1}{2}\int_0^t \lambda\cdot c(\mathbf{x}_q)\lambda\,dq - \mathbf{I}(\mathbf{x})\bigr)$$
$$= \lambda\cdot\hat{\mathbf{x}}_s - \int_0^s \lambda\cdot b(\hat{\mathbf{x}}_q)\,dq - \frac{1}{2}\int_0^s \lambda\cdot c(\hat{\mathbf{x}}_q)\lambda\,dq - \mathbf{I}(\hat{\mathbf{x}}). \quad (1.5)$$

We note similarity with the dynamic programming equation. It was shown in Puhalskii [36] that in (1.5) it is possible to replace constants λ with functions $\lambda(s, \mathbf{x})$ from a certain class. If the diffusion matrix $c(x)$ is nondegenerate, this class is large enough to identify $\mathbf{I}$. In the present article we build on that result but adopt a "probabilistic" rather than an analytical approach to the uniqueness problem.

Let us recall that the key element in obtaining results on weak convergence to stochastic processes with possibly degenerate coefficients is the fact that solutions to the limit martingale problems are weak solutions to the associated stochastic equations, see, e.g., Ikeda and Watanabe [22], Jacod and Shiryaev [23], Liptser and Shiryaev [28], Stroock and Varadhan [41]. In a similar vein, one can associate idempotent differential equations with maxingale problems and show that weak solutions to these equations are solutions to the maxingale problems. For instance, Theorem 2.8.9 in Puhalskii [36] establishes conditions under which a Luzin weak solution to the idempotent Ito equation

$$X_t = x_0 + \int_0^t b(X_s)\,ds + \int_0^t c(X_s)^{1/2}\,\dot{W}_s\,ds,\ t\in\mathbb{R}_+, \quad (1.6)$$

where $W = (W_t,\, t \in \mathbb{R}_+)$ is an idempotent Wiener process, is a solution to problem (1.4). (The terminology of idempotent probability theory is recalled in Section 2.) As for the converse, it has been known to hold for nondegenerate maxingale problems, Puhalskii [36, Theorems 2.6.30 and 2.8.21], and for a certain degenerate maxingale problem associated with idempotent Poisson processes, Puhalskii [38]. In this paper, the result that solutions to maxingale problems are weak solutions to the corresponding idempotent equations is extended to maxingale problems of the diffusion type (1.4) with possibly degenerate matrices c. Thus, in complete analogy with the situation in weak convergence theory, the issue of uniqueness for maxingale problems of diffusion type is equivalent to that of Luzin weak uniqueness for idempotent Ito differential equations. Interestingly, the proof draws on the perturbation idea, which is so useful in the stochastic setting, however, the perturbation is introduced in the idempotent equation (1.6) rather than in a stochastic process (of course, there are no stochastic processes to speak of !).

With the equivalence between maxingale problems of diffusion type and idempotent Ito differential equations in hand, we again follow in the footsteps of the stochastic setting to deduce that pathwise uniqueness for idempotent Ito equations implies uniqueness for the associated maxingale problems. Equation (1.6) is in fact an ordinary differential equation of a certain kind. The uniqueness issue for ordinary differential equations is well understood, e.g., Lipshitz continuity of the coefficients suffices. This enables us to obtain results on uniqueness for degenerate maxingale problems of diffusion type. In particular, as in the theory of ordinary differential equations, it is possible to relax the requirements of Lipshitz continuity and of linear growth.

The paper is structured as follows. Section 2 contains a background on idempotent probability. In Section 3 existence and uniqueness issues for diffusion maxingale problems are analysed. Section

4 is concerned with large deviation applications.

2 Technical preliminaries

This section reviews necessary results from idempotent probability theory. More details can be found in Puhalskii [36]. We let $\| \ \|$ denote the operator norm of a matrix, let $\vee$ denote the maximum, and let $\wedge$ denote the minimum.

Let Υ be a set. A function $\mathbf{\Pi}$ from the power set of Υ to $[0,1]$ is called an idempotent probability if $\mathbf{\Pi}(\Gamma) = \sup_{\upsilon\in\Gamma} \mathbf{\Pi}(\{\upsilon\})$ for $\Gamma \subset \Upsilon$ and $\mathbf{\Pi}(\Upsilon) = 1$. If, in addition, Υ is a metric space and the sets $\{\upsilon \in \Upsilon : \mathbf{\Pi}(\{\upsilon\}) \ge a\}$ are compact for all $a \in (0,1]$, then $\mathbf{\Pi}$ is called a deviability. The definition implies that $\mathbf{\Pi}$ is a deviability if and only if $\mathbf{I}(\upsilon) = -\log \mathbf{\Pi}(\{\upsilon\})$ is an action functional. We denote $\mathbf{\Pi}(\upsilon) = \mathbf{\Pi}(\{\upsilon\})$ and assume throughout this section that $\mathbf{\Pi}$ is an idempotent probability on Υ.

A property $\mathcal{P}(\upsilon)$, $\upsilon \in \Upsilon$, pertaining to the elements of Υ is said to hold $\mathbf{\Pi}$-a.e. if $\mathbf{\Pi}(\mathcal{P}(\upsilon)$ does not hold$) = 0$. A τ-algebra $\mathcal{A}$ on Υ is defined as a collection of subsets of Υ that are unions of the elements of some partitioning of Υ into disjoint sets. We call the elements of the partitioning the atoms of $\mathcal{A}$ and denote as $[\upsilon]$ the atom containing υ. τ-algebras are closed under set-differences, arbitrary unions and intersections of sets. The power set of Υ is referred to as the discrete τ-algebra. A τ-algebra $\mathcal{A}$ is called complete (or $\mathbf{\Pi}$-complete, or complete with respect to $\mathbf{\Pi}$ if the idempotent probability needs to be specified) if each one-point set $\{\upsilon\}$ with $\mathbf{\Pi}(\upsilon) = 0$ is an atom of $\mathcal{A}$; the completion (or the $\mathbf{\Pi}$-completion, or the completion with respect to $\mathbf{\Pi}$) of a τ-algebra $\mathcal{A}$ is defined as the τ-algebra obtained by taking as atoms the elements of Υ of idempotent probability 0 and the maximal subsets of the atoms of $\mathcal{A}$ that do not contain elements of idempotent probability 0; the completion of a τ-algebra is a complete τ-algebra. If Υ' is also a set equipped with idempotent probability $\mathbf{\Pi}'$ and τ-algebra $\mathcal{A}'$, then the product idempotent probability $\mathbf{\Pi} \times \mathbf{\Pi}'$ on $\Upsilon \times \Upsilon'$ is defined by $(\mathbf{\Pi} \times \mathbf{\Pi}')(\upsilon, \upsilon') = \mathbf{\Pi}(\upsilon)\mathbf{\Pi}'(\upsilon')$ for $(\upsilon, \upsilon') \in \Upsilon \times \Upsilon'$, the product τ-algebra $\mathcal{A} \otimes \mathcal{A}'$ is defined as having the atoms $[\upsilon] \times [\upsilon']$, where $\upsilon \in \Upsilon$ and $\upsilon' \in \Upsilon'$.

A function f from a set Υ equipped with idempotent probability $\mathbf{\Pi}$ to a set Υ' is called an idempotent variable. If Υ and Υ' are equipped with τ-algebras $\mathcal{A}$ and $\mathcal{A}'$, respectively, the idempotent variable f is said to be $\mathcal{A}/\mathcal{A}'$-measurable, or simply measurable if the τ-algebras are understood, provided $f^{-1}([\upsilon']) \in \mathcal{A}$ for all $\upsilon' \in \Upsilon'$. We say that f is $\mathcal{A}$-measurable if it is measurable for the discrete τ-algebra on Υ', i.e., $\{\upsilon \in \Upsilon : f(\upsilon) = \upsilon'\} \in \mathcal{A}$ for all $\upsilon' \in \Upsilon'$. The τ-algebra on Υ generated by f is defined by the atoms $\{\upsilon \in \Upsilon : f(\upsilon) = \upsilon'\}$, $\upsilon' \in \Upsilon'$.

As in probability theory, we routinely omit the argument υ in the notation for an idempotent variable. The idempotent distribution of an idempotent variable f is defined as the set function $\mathbf{\Pi} \circ f^{-1}(\Gamma) = \mathbf{\Pi}(f \in \Gamma)$, $\Gamma \subset \Upsilon'$; it is also called the image of $\mathbf{\Pi}$ under f. If Υ is a metric space, $\mathbf{\Pi}$ is a deviability on Υ, and f is a mapping from Υ to a metric space Υ', which is continuous when restricted to the sets $\{\upsilon \in \Upsilon : \mathbf{\Pi}(\upsilon) \ge a\}$ for $a \in (0,1]$, then $\mathbf{\Pi} \circ f^{-1}$ is a deviability on Υ'. More generally, f is said to be Luzin if $\mathbf{\Pi} \circ f^{-1}$ is a deviability on Υ'.

Subsets A and A' of Υ are said to be independent if $\mathbf{\Pi}(A \cap A') = \mathbf{\Pi}(A)\mathbf{\Pi}(A')$; τ-algebras $\mathcal{A}$ and $\mathcal{A}'$ are said to be independent if sets A and A' are independent for arbitrary $A \in \mathcal{A}$ and $A' \in \mathcal{A}'$; Υ'-valued idempotent variables f and f' are said to be independent if $\mathbf{\Pi}(f = \upsilon',\ f' = \upsilon'') = \mathbf{\Pi}(f = \upsilon')\mathbf{\Pi}(f' = \upsilon'')$ for all $\upsilon', \upsilon'' \in \Upsilon'$. An idempotent variable f and a τ-algebra $\mathcal{A}$ are said to be independent (or f to be independent of $\mathcal{A}$) if the τ-algebra generated by f and $\mathcal{A}$ are independent. If f is $\mathbb{R}_+$-valued, the idempotent expectation of f is defined by $\mathbf{S}f = \sup_{\upsilon\in\Upsilon} f(\upsilon)\mathbf{\Pi}(\upsilon)$, it is also denoted as $\mathbf{S}_{\mathbf{\Pi}} f$ if the reference idempotent probability needs to be indicated. The following analogue of the Markov inequality holds: $\mathbf{\Pi}(f \ge a) \le \mathbf{S}f/a$, where $a > 0$. If $\mathbb{R}_+$-valued idempotent variables f and f' are independent, then $\mathbf{S}(ff') = \mathbf{S}f\,\mathbf{S}f'$.

An $\mathbb{R}_+$-valued idempotent variable f is said to be maximable if $\lim_{b\to\infty} \mathbf{S}(f\mathbf{1}(f > b)) = 0$. A collection f_α of $\mathbb{R}_+$-valued idempotent variables is called uniformly maximable if $\lim_{b\to\infty} \sup_\alpha \mathbf{S}(f_\alpha \mathbf{1}(f_\alpha > b)) = 0$. The conditional idempotent expectation of an $\mathbb{R}_+$-valued idempotent variable f given a τ-algebra $\mathcal{A}$ is defined by

$$\mathbf{S}(f|\mathcal{A})(\upsilon) = \begin{cases} \sup\limits_{\upsilon' \in [\upsilon]} f(\upsilon') \dfrac{\mathbf{\Pi}(\upsilon')}{\mathbf{\Pi}([\upsilon])} & \text{if } \mathbf{\Pi}([\upsilon]) > 0, \\ f'(\upsilon) & \text{if } \mathbf{\Pi}([\upsilon]) = 0, \end{cases}$$

where $f'(\upsilon)$ is an $\mathbb{R}_+$-valued idempotent variable constant on the atoms of $\mathcal{A}$. Conditional idempotent expectation is thus specified uniquely $\mathbf{\Pi}$-a.e. It has many of the properties of conditional expectation, in particular, $\mathbf{S}(f|\mathcal{A})$ is $\mathcal{A}$-measurable, if f is $\mathcal{A}$-measurable then $\mathbf{S}(f|\mathcal{A}) = f$ $\mathbf{\Pi}$-a.e., and if f and $\mathcal{A}$ are independent then $\mathbf{S}(f|\mathcal{A}) = \mathbf{S}f$ $\mathbf{\Pi}$-a.e.

A flow of τ-algebras, or a τ-flow, on Υ is defined as a collection $\mathbf{A} = (\mathcal{A}_t,\, t \in \mathbb{R}_+)$ of τ-algebras on Υ such that $\mathcal{A}_s \subset \mathcal{A}_t$ for $s \le t$; the latter condition is equivalent to the atoms of $\mathcal{A}_s$ being unions of the atoms of $\mathcal{A}_t$. A τ-flow is called complete if it consists of complete τ-algebras, the completion of a τ-flow is obtained by completing its τ-algebras. An idempotent variable $\kappa : \Upsilon \to \mathbb{R}_+$ is called an idempotent $\mathbf{A}$-stopping time, or a stopping time relative to $\mathbf{A}$, if $\{\upsilon :\, \kappa(\upsilon) = t\} \in \mathcal{A}_t$ for $t \in \mathbb{R}_+$. Given a τ-flow $\mathbf{A}$ and an idempotent $\mathbf{A}$-stopping time κ, we define $\mathcal{A}_\kappa$ as the τ-algebra with atoms $[\upsilon]_{\mathcal{A}_{\kappa(\upsilon)}}$.

A collection $X = (X_t,\, t \in \mathbb{R}_+)$ of Υ'-valued idempotent variables on Υ is called an Υ'-valued idempotent process. The functions $(X_t(\upsilon),\, t \in \mathbb{R}_+)$ for various $\upsilon \in \Upsilon$ are called paths (or trajectories) of X. The idempotent process X is said to be $\mathbf{A}$-adapted if the X_t are $\mathcal{A}_t$-measurable for $t \in \mathbb{R}_+$. The τ-flow generated by X is defined as consisting of the τ-algebras generated by the X_t. It is the minimal τ-flow, to which X is adapted. If the idempotent process X is $\mathbf{A}$-adapted and $\mathbb{R}$-valued with unbounded above continuous paths, then $\kappa = \inf\{t \in \mathbb{R}_+ :\, X_t \ge a\}$, where $a \in \mathbb{R}$, is an idempotent $\mathbf{A}$-stopping time.

An $\mathbf{A}$-adapted $\mathbb{R}_+$-valued idempotent process $M = (M_t,\, t \in \mathbb{R}_+)$ is said to be an $\mathbf{A}$-exponential maxingale, or an exponential maxingale relative to $\mathbf{A}$, if the M_t are maximable and $\mathbf{S}(M_t|\mathcal{A}_s) = M_s$ $\mathbf{\Pi}$-a.e. for $s \le t$. If, in addition, the collection M_t, $t \in \mathbb{R}_+$, is uniformly maximable, then M is said to be a uniformly maximable exponential maxingale. An $\mathbf{A}$-adapted $\mathbb{R}_+$-valued idempotent process $M = (M_t,\, t \in \mathbb{R}_+)$ is called an $\mathbf{A}$-local exponential maxingale, or a local exponential maxingale relative to $\mathbf{A}$, if there exists a sequence τ_n of idempotent $\mathbf{A}$-stopping times such that $\tau_n \uparrow \infty$ as $n \to \infty$ and the stopped idempotent processes $(M_{t\wedge\tau_n},\, t \in \mathbb{R}_+)$ are uniformly maximable $\mathbf{A}$-exponential maxingales; the sequence τ_n is referred to as a localising sequence for M.

Lemma 2.1. *Let $(M_t,\, t \in \mathbb{R}_+)$ and $(M'_t,\, t \in \mathbb{R}_+)$ be local exponential maxingales on $(\Upsilon, \mathbf{\Pi})$ and $(\Upsilon', \mathbf{\Pi}')$, respectively, relative to respective τ-flows $(\mathcal{A}_t,\, t \in \mathbb{R}_+)$ and $(\mathcal{A}'_t,\, t \in \mathbb{R}_+)$. Then $(M_t M'_t,\, t \in \mathbb{R}_+)$ is a local exponential maxingale on $(\Upsilon \times \Upsilon', \mathbf{\Pi} \times \mathbf{\Pi}')$ relative to the τ-flow $(\mathcal{A}_t \otimes \mathcal{A}'_t,\, t \in \mathbb{R}_+)$.*

Proof. Let τ_n and τ'_n be localising sequences for M and M', respectively. By Puhalskii [36, Lemma 1.6.28], $\mathbf{S}_{\mathbf{\Pi}\times\mathbf{\Pi}'}(M_{t\wedge\tau_n} M'_{t\wedge\tau'_n} | \mathcal{A}_s \otimes \mathcal{A}'_s) = \mathbf{S}_{\mathbf{\Pi}}(M_{t\wedge\tau_n}|\mathcal{A}_s)\mathbf{S}_{\mathbf{\Pi}'}(M'_{t\wedge\tau'_n}|\mathcal{A}'_s)$ $\mathbf{\Pi} \times \mathbf{\Pi}'$-a.e. for $s \le t$. Uniform maximability of $(M_{t\wedge\tau_n} M'_{t\wedge\tau'_n},\, t \in \mathbb{R}_+)$ under $\mathbf{\Pi} \times \mathbf{\Pi}'$ is obvious. □

Let $\langle M \rangle = (\langle M \rangle_t,\, t \in \mathbb{R}_+)$ be an $\mathbb{R}^{d\times d}$-valued continuous-path $\mathbf{A}$-adapted idempotent process such that $\langle M \rangle_0 = 0$ and $\langle M \rangle_t - \langle M \rangle_s$ for $0 \le s \le t$ are positive semi-definite symmetric $d \times d$ matrices. An $\mathbb{R}^d$-valued continuous-path $\mathbf{A}$-adapted idempotent process $M = (M_t,\, t \in \mathbb{R}_+)$ is called a local maxingale with quadratic characteristic $\langle M \rangle$ relative to $\mathbf{A}$ if the idempotent processes $(\exp(\lambda \cdot M_t - \lambda \cdot \langle M \rangle_t \lambda/2), t \in \mathbb{R}_+)$ are $\mathbf{A}$-local exponential maxingales for all $\lambda \in \mathbb{R}^d$. The next lemma presents a version of Lenglart-Rebolledo's and Doob's inequalities, see, e.g., Liptser and Shiryaev [28].

Lemma 2.2. *If M is a local maxingale with quadratic characteristic $\langle M\rangle$ relative to $\mathbf{A}$ and $M_0 = 0$, then for arbitrary $a > 0$, $b > 0$, and $t > 0$*

$$\mathbf{\Pi}(\sup_{s\in[0,t]} |M_s| \geq a) \leq \exp(b-a) \vee \mathbf{\Pi}(\|\langle M\rangle_t\| \geq 2b).$$

If, in addition, the $\langle M\rangle_s$, $s \in [0,t]$, do not depend on $\upsilon \in \Upsilon$, then

$$\mathbf{\Pi}(\sup_{s\in[0,t]} |M_s| \geq a) \leq \exp\bigl(-a + \frac{1}{2}\|\langle M\rangle_t\|\bigr).$$

Proof. Let $Y_t(\lambda) = \exp(\lambda\cdot M_t - \lambda\cdot\langle M\rangle_t\lambda/2)$. Since the idempotent process $(Y_t(\lambda), t \in \mathbb{R}_+)$ is an $\mathbf{A}$-local exponential maxingale for arbitrary $\lambda \in \mathbb{R}^d$, so $\mathbf{S}Y_t(\lambda) \leq 1$ ("Fatou's lemma", Puhalskii [36, Theorem 1.4.19]) and one can write

$$\begin{aligned}\mathbf{\Pi}(\sup_{s\in[0,t]} |M_s| \geq a) &= \sup_{s\in[0,t]}\sup_{\substack{\lambda\in\mathbb{R}^d:\\|\lambda|=1}} \mathbf{\Pi}(\lambda\cdot M_s \geq a) = \sup_{s\in[0,t]}\sup_{\substack{\lambda\in\mathbb{R}^d:\\|\lambda|=1}} \mathbf{\Pi}\Bigl(Y_s(\lambda) \geq \exp\bigl(a - \frac{1}{2}\lambda\cdot\langle M\rangle_s\lambda\bigr)\Bigr)\\ &\leq \mathbf{\Pi}(\|\langle M\rangle_t\| \geq 2b) \vee \sup_{s\in[0,t]}\sup_{\substack{\lambda\in\mathbb{R}^d:\\|\lambda|=1}} \mathbf{\Pi}\bigl(Y_s(\lambda) \geq \exp(a-b)\bigr)\\ &\leq \mathbf{\Pi}(\|\langle M\rangle_t\| \geq 2b) \vee \sup_{s\in[0,t]}\sup_{\substack{\lambda\in\mathbb{R}^d:\\|\lambda|=1}} \mathbf{S}Y_s(\lambda)\exp(b-a) \leq \mathbf{\Pi}(\|\langle M\rangle_t\| \geq 2b) \vee \exp(b-a).\end{aligned}$$

If the $\langle M\rangle_s$ do not depend on $\upsilon \in \Upsilon$, then

$$\mathbf{\Pi}\Bigl(Y_s(\lambda) \geq \exp\bigl(a - \frac{1}{2}\lambda\cdot\langle M\rangle_s\lambda\bigr)\Bigr) \leq \mathbf{S}Y_s(\lambda)\exp\bigl(-a + \frac{1}{2}\lambda\cdot\langle M\rangle_s\lambda\bigr) \leq \exp\bigl(-a + \frac{1}{2}\|\langle M\rangle_s\|\bigr).$$

□

Let us recall that $\mathbb{C}$ denotes the metric space of continuous $\mathbb{R}^d$-valued functions on $\mathbb{R}_+$ with compact-open topology. The canonical τ-flow on $\mathbb{C}$ is defined as the τ-flow $\mathbf{C} = (\mathcal{C}_t,\, t \in \mathbb{R}_+)$ with the $\mathcal{C}_t$ having as atoms the sets $p_t^{-1}\mathbf{x}$, $\mathbf{x} \in \mathbb{C}$, where the $p_t : \mathbb{C} \to \mathbb{C}$ are defined by $(p_t\mathbf{x})_s = \mathbf{x}_{s\wedge t}$, $s \in \mathbb{R}_+$. The canonical idempotent process is defined by $X_t(\mathbf{x}) = \mathbf{x}_t$.

Given an $\mathbb{R}$-valued function $G = (G_t(\lambda;\mathbf{x}),\, t \in \mathbb{R}_+,\, \mathbf{x} \in \mathbb{C},\, \lambda \in \mathbb{R}^d)$, where $G_t(\lambda;\mathbf{x})$ is $\mathcal{C}_t$-measurable in $\mathbf{x}$, we say that a deviability $\mathbf{\Pi}$ on $\mathbb{C}$ solves the maxingale problem (x_0, G), where $x_0 \in \mathbb{R}^d$, if $X_0 = x_0$ $\mathbf{\Pi}$-a.e. and $(\exp(\lambda\cdot X_t - G_t(\lambda;X)),\, t \in \mathbb{R}_+)$ is a $\mathbf{C}$-local exponential maxingale under $\mathbf{\Pi}$ for arbitrary $\lambda \in \mathbb{R}^d$, where $X = (X_t,\, t \in \mathbb{R}_+)$ is the canonical idempotent process on $\mathbb{C}$. In particular, if X is a local maxingale with quadratic characteristic $\langle M\rangle = (\langle M\rangle_t(\mathbf{x}),\, t \in \mathbb{R}_+, \mathbf{x} \in \mathbb{C})$ under $\mathbf{\Pi}$ and $X_0 = x_0$ $\mathbf{\Pi}$-a.e., then $\mathbf{\Pi}$ solves the maxingale problem (x_0, G) for $G_t(\lambda;\mathbf{x}) = \lambda\cdot\langle M\rangle_t(\mathbf{x})\lambda/2$.

Wiener idempotent probability (or Wiener deviability) is a deviability on $\mathbb{C}$ defined by

$$\mathbf{\Pi}^W(\mathbf{x}) = \begin{cases}\exp\Bigl(-\dfrac{1}{2}\displaystyle\int_0^\infty |\dot{\mathbf{x}}_t|^2\,dt\Bigr) & \text{if } \mathbf{x} \text{ is absolutely continuous and } \mathbf{x}_0 = 0,\\ 0 & \text{otherwise.}\end{cases}$$

A standard idempotent Wiener process is defined as an idempotent process with idempotent distribution $\mathbf{\Pi}^W$. Thus, it has $\mathbf{\Pi}$-a.e. absolutely continuous paths. The definition implies that the canonical idempotent process on $\mathbb{C}$ is idempotent Wiener under $\mathbf{\Pi}^W$. The component processes of an $\mathbb{R}^d$-valued standard idempotent Wiener process are independent one-dimensional standard idempotent Wiener processes.

If $W = (W_t,\, t \in \mathbb{R}_+)$ is an $\mathbb{R}^d$-valued standard idempotent Wiener process on $(\Upsilon, \mathbf{\Pi})$, then the idempotent process $\big(\exp(\lambda \cdot W_t - |\lambda|^2 t/2),\, t \in \mathbb{R}_+\big)$ is an exponential maxingale relative to the flow $\mathbf{A}^W = (\mathcal{A}_t^W,\, t \in \mathbb{R}_+)$, where the $\mathcal{A}_t^W$ are the τ-algebras generated by W_s, $s \le t$. An $\mathbb{R}^d$-valued continuous-path idempotent process W is said to be standard idempotent Wiener relative to a τ-flow $\mathbf{A}$ if $W_0 = 0$ and the idempotent process $\big(\exp(\lambda \cdot W_t - |\lambda|^2 t/2),\, t \in \mathbb{R}_+\big)$ is an $\mathbf{A}$-exponential maxingale for arbitrary $\lambda \in \mathbb{R}^d$. If an idempotent process is idempotent Wiener relative to a τ-flow, then it is idempotent Wiener.

Let W be an $\mathbb{R}^d$-valued standard idempotent Wiener process and $(\sigma_t(\upsilon),\, t \in \mathbb{R}_+)$ be an $\mathbb{R}^{d\times d}$-valued idempotent process, which is Lebesgue measurable in t for all υ and is such that $\sigma_s(\upsilon)\sigma_s(\upsilon)^T$ is locally integrable for all υ. We define the idempotent Ito integral $(\sigma \diamond W)_t$ by

$$(\sigma \diamond W)_t(\upsilon) = \begin{cases} \int\limits_0^t \sigma_s(\upsilon)\dot{W}_s(\upsilon)\, ds & \text{if } \mathbf{\Pi}(\upsilon) > 0, \\ Y(\upsilon) & \text{otherwise}\ , \end{cases}$$

where $Y(\upsilon)$ is an $\mathbb{R}^d$-valued idempotent variable. The integral is thus specified uniquely $\mathbf{\Pi}$-a.e. The idempotent process $\sigma \diamond W = \big((\sigma \diamond W)_t,\, t \in \mathbb{R}_+\big)$ has $\mathbf{\Pi}$-a.e. continuous paths. If $(W_t,\, t \in \mathbb{R}_+)$ and $(\sigma_t,\, t \in \mathbb{R}_+)$ are adapted to a complete τ-flow $\mathbf{A}$, then $\sigma \diamond W$ is $\mathbf{A}$-adapted. The next lemma is a version of Theorem 2.5.19 in Puhalskii [36]. Let tr denote the trace of a matrix.

Lemma 2.3. *Let $(\sigma_t(\mathbf{x}),\, t \in \mathbb{R}_+,\, \mathbf{x} \in \mathbb{C})$ be an $\mathbb{R}^{k\times m}$-valued function, which is continuous in $\mathbf{x}$, Lebesgue measurable in t, and is adapted to the canonical τ-flow on $\mathbb{C}$. Let $(X_t,\, t \in \mathbb{R}_+)$ be a continuous-path $\mathbb{R}^d$-valued idempotent process, which is adapted to a complete τ-flow $\mathbf{A}$ on Υ and whose idempotent distribution is a deviability on $\mathbb{C}$. Let $(W_t,\, t \in \mathbb{R}_+)$ be an $\mathbb{R}^m$-valued standard idempotent Wiener process relative to $\mathbf{A}$. If, in addition, $\int_0^t \operatorname{tr}\sigma_s(\mathbf{x})\sigma_s(\mathbf{x})^T\, ds$ is a well-defined continuous function of $\mathbf{x} \in \mathbb{C}$ for every $t \in \mathbb{R}_+$, then $\sigma \diamond W$ is a local maxingale relative to $\mathbf{A}$ with quadratic characteristic $\big(\int_0^t \sigma_s(X)\sigma_s(X)^T\, ds,\, t \in \mathbb{R}_+\big)$.*

Proof. Continuity of $\sigma_t(\mathbf{x})$ and $\int_0^t \operatorname{tr}\sigma_s(\mathbf{x})\sigma_s(\mathbf{x})^T\, ds$ in $\mathbf{x}$ implies that $\int_0^t \|\sigma_s(\mathbf{x})\sigma_s(\mathbf{x})^T\|\, ds$ is continuous in $\mathbf{x}$, so the functions $\int_0^t \|\sigma_s(\mathbf{x})\sigma_s(\mathbf{x})^T\|\, \mathbf{1}\big(\|\sigma_s(\mathbf{x})\sigma_s(\mathbf{x})^T\| \ge A\big)\, ds$ are upper semicontinuous in $\mathbf{x}$ for all A. By Dini's theorem, they converge to 0 as $A \to \infty$ uniformly on compacts, which checks the hypotheses of Theorem 2.5.19 in Puhalskii [36]. $\square$

Below, we use $\int_0^t \sigma_s \dot{W}_s\, ds$ to denote $(\sigma \diamond W)_t$. Let $\sigma_t(\mathbf{x})$, $\mathbf{x} \in \mathbb{C}$, $t \in \mathbb{R}_+$, and $b_t(\mathbf{x})$, $\mathbf{x} \in \mathbb{C}$, $t \in \mathbb{R}_+$, be functions with values in $\mathbb{R}^{d\times d}$ and $\mathbb{R}^d$, respectively, which are Lebesgue-measurable in t and $\mathcal{C}_t$-measurable in $\mathbf{x}$. Let also the functions $b_t(\mathbf{x})$ and $\sigma_t(\mathbf{x})\sigma_t(\mathbf{x})^T$ be locally integrable in t. Let W be an $\mathbb{R}^d$-valued standard idempotent Wiener process on an idempotent probability space $(\Upsilon, \mathbf{\Pi})$ relative to a complete τ-flow $\mathbf{A}$ and let $\overline{\mathcal{C}}_t^W$ for $t \in \mathbb{R}_+$ denote the completion of $\mathcal{C}_t$ with respect to the Wiener deviability on $\mathbb{C}$. We say that, given $x_0 \in \mathbb{R}^d$, an $\mathbb{R}^d$-valued idempotent process X on $(\Upsilon, \mathbf{\Pi})$ is a strong solution to the idempotent Ito equation

$$X_t = x_0 + \int\limits_0^t b_s(X)\, ds + \int\limits_0^t \sigma_s(X)\, \dot{W}_s\, ds,\ t \in \mathbb{R}_+, \tag{2.1}$$

if equality (2.1) holds $\mathbf{\Pi}$-a.e. and there exists a function $J : \mathbb{C} \to \mathbb{C}$, which is $\overline{\mathcal{C}}_t^W/\mathcal{C}_t$-measurable for every $t \in \mathbb{R}_+$, such that $X = J(W)$ $\mathbf{\Pi}$-a.e. As a consequence, X is $\mathbf{A}$-adapted. A strong solution is called Luzin if the function J is continuous when restricted to the sets $\{\mathbf{x} \in \mathbb{C} : \mathbf{\Pi}^W(\mathbf{x}) \geq a\}$ for $a \in (0, 1]$. We say that there exists a unique strong solution (respectively, Luzin strong solution) if any strong solution (respectively, Luzin strong solution) can be written as $X = J(W)$ $\mathbf{\Pi}$-a.e. for the same function J.

We say that idempotent probability $\mathbf{\Pi}_{x_0}$ on $\mathbb{C}$ is a weak solution of (2.1) if there exists an idempotent probability space $(\Upsilon, \mathbf{\Pi})$ with a complete τ-flow $\mathbf{A}$, on which are defined a continuous-path $\mathbb{R}^d$-valued $\mathbf{A}$-adapted idempotent process $X = (X_t,\, t \in \mathbb{R}_+)$ and a continuous-path $\mathbb{R}^d$-valued $\mathbf{A}$-adapted idempotent process $W = (W_t,\, t \in \mathbb{R}_+)$ such that W is standard idempotent Wiener relative to $\mathbf{A}$, X has idempotent distribution $\mathbf{\Pi}_{x_0}$, and (2.1) holds for all $t \in \mathbb{R}_+$ $\mathbf{\Pi}$-a.e. on Υ. If the joint idempotent distribution of (X, W) (and, hence, $\mathbf{\Pi}_{x_0}$) is, in addition, a deviability, then we call $\mathbf{\Pi}_{x_0}$ a Luzin weak solution. If a strong (respectively, Luzin strong) solution exists, then its idempotent distribution is a weak (respectively, Luzin weak) solution. Occasionally, we call the idempotent process X itself a weak solution (respectively, Luzin weak solution) so that, with a slight abuse of terminology, one can say that a (Luzin) strong solution is a (Luzin) weak solution. It is worth mentioning that weak solutions can always be realised on the canonical space $\mathbb{C} \times \mathbb{C}$, with $X = (X_t,\, t \in \mathbb{R}_+)$ and $W = (W_t,\, t \in \mathbb{R}_+)$ being the projection processes defined by $X_t(\mathbf{x}, \mathbf{x}') = \mathbf{x}_t$ and $W_t(\mathbf{x}, \mathbf{x}') = \mathbf{x}'_t$ for $(\mathbf{x}, \mathbf{x}') \in \mathbb{C} \times \mathbb{C}$. Then the joint idempotent distribution of (X, W) coincides with $\mathbf{\Pi}$ and $\mathbf{\Pi}_{x_0}(\mathbf{x}) = \sup_{\mathbf{x}' \in \mathbb{C}} \mathbf{\Pi}(\mathbf{x}, \mathbf{x}')$ for $\mathbf{x} \in \mathbb{C}$.

Pathwise uniqueness is said to hold for (2.1) if for every two weak solutions (X, W) and (X', W'), defined on a common idempotent probability space $(\Upsilon, \mathbf{\Pi})$ with the same τ-flow $\mathbf{A}$, we have $X = X'$ $\mathbf{\Pi}$-a.e. provided $W = W'$ $\mathbf{\Pi}$-a.e. Pathwise uniqueness holds if for every $\mathbf{w} \in \mathbb{C}$ such that $\mathbf{\Pi}^W(\mathbf{w}) > 0$ the initial value problem $\dot{\mathbf{x}}_t = b_t(\mathbf{x}) + \sigma_t(\mathbf{x})\, \dot{\mathbf{w}}_t,\ \mathbf{x}_0 = x_0$, has at most one solution $\mathbf{x}$. If pathwise uniqueness holds, then (2.1) has at most one weak solution so that if, in addition, there exists a (Luzin) weak solution, then it is unique, and there exists a unique (Luzin) strong solution.

As a consequence, we have the following result on strong existence and uniqueness. Let the functions $b_t(\mathbf{x})$ and $\sigma_t(\mathbf{x})$ be, in addition, locally Lipshitz continuous in $\mathbf{x}$, i.e., for every $a > 0$ there exists an $\mathbb{R}_+$-valued locally integrable in t function l_t^a, $t \in \mathbb{R}_+$, such that $|b_t(\mathbf{x}) - b_t(\mathbf{y})| \leq l_t^a \sup_{s \leq t} |\mathbf{x}_s - \mathbf{y}_s|$ and $\|\sigma_t(\mathbf{x}) - \sigma_t(\mathbf{y})\|^2 \leq l_t^a \sup_{s \leq t} |\mathbf{x}_s - \mathbf{y}_s|^2$ if $\sup_{s \leq t} |\mathbf{x}_s| \leq a$ and $\sup_{s \leq t} |\mathbf{y}_s| \leq a$. Let them also satisfy the linear-growth condition that there exists an $\mathbb{R}_+$-valued locally integrable function l_t, $t \in \mathbb{R}_+$, such that $|b_t(\mathbf{x})| \leq l_t(1 + \sup_{s \leq t} |\mathbf{x}_s|)$ and $\|\sigma_t(\mathbf{x})\|^2 \leq l_t(1 + \sup_{s \leq t} |\mathbf{x}_s|^2)$ for $\mathbf{x} \in \mathbb{C}$ and $t \in \mathbb{R}_+$. Then (2.1) has a unique weak solution, which is also a Luzin strong solution.

Let Υ be a metric space. A net $\mathbf{\Pi}_\phi$, $\phi \in \Phi$, where Φ is a directed set, of idempotent probabilities on Υ is said to converge weakly to idempotent probability $\mathbf{\Pi}$ on Υ if $\lim_{\phi \in \Phi} \mathbf{S}_{\mathbf{\Pi}_\phi} f = \mathbf{S}_{\mathbf{\Pi}} f$ for every $\mathbb{R}_+$-valued bounded and continuous function f on Υ; equivalently, $\limsup_{\phi \in \Phi} \mathbf{\Pi}_\phi(F) \leq \mathbf{\Pi}(F)$ for all closed sets $F \subset \Upsilon$ and $\liminf_{\phi \in \Phi} \mathbf{\Pi}_\phi(G) \geq \mathbf{\Pi}(G)$ for all open sets $G \subset \Upsilon$. A net of idempotent variables with values in a common metric space is said to converge in idempotent distribution if their idempotent distributions weakly converge. One has the continuous mapping principle for convergence in idempotent distribution: if a net X_ϕ, $\phi \in \Phi$, of idempotent variables with values in Υ converges in idempotent distribution to an idempotent variable X with values in Υ and f is a continuous function from Υ to a metric space Υ', then the net $f(X_\phi)$, $\phi \in \Phi$, converges in idempotent distribution to $f(X)$. A net $\mathbf{\Pi}_\phi$, $\phi \in \Phi$, of deviabilities on Υ is said to be tight if $\inf_{K \in \mathcal{K}} \limsup_{\phi \in \Phi} \mathbf{\Pi}_\phi(\Upsilon \setminus K) = 0$, where $\mathcal{K}$ denotes the collection of compact subsets of Υ. A tight net of deviabilities is weakly compact, i.e., it contains a subnet that converges weakly to a deviability (if $\mathbf{\Pi}_\phi$ is a sequence, then it contains a weakly convergent subsequence).

3 Maxingale problems of diffusion type and idempotent Ito equations

The purpose of this section is to establish equivalence between solutions of maxingale problems of diffusion type and Luzin weak solutions to the associated idempotent Ito equations. We also present results on existence and uniqueness for idempotent Ito equations.

As above, $\mathbb{C}$ denotes the metric space of continuous $\mathbb{R}^d$-valued functions $\mathbf{x} = (\mathbf{x}_t,\, t \in \mathbb{R}_+)$ endowed with compact-open topology, $\mathcal{C}_s$, for $s \in \mathbb{R}_+$, denotes the τ-algebra on $\mathbb{C}$ generated by the mappings $\mathbf{x} \to \mathbf{x}_t$, $t \in [0, s]$, and the canonical τ-flow on $\mathbb{C}$ is defined by $\mathbf{C} = (\mathcal{C}_t,\, t \in \mathbb{R}_+)$. We identify the product space $\hat{\mathbb{C}} = \mathbb{C} \times \mathbb{C}$ with the space of continuous $\mathbb{R}^d \times \mathbb{R}^d$-valued functions $(\mathbf{x}, \mathbf{x}') = \big((\mathbf{x}_t, \mathbf{x}'_t),\, t \in \mathbb{R}_+\big)$ with compact-open topology; the canonical τ-flow on $\hat{\mathbb{C}}$ is denoted by $\hat{\mathbf{C}} = (\hat{\mathcal{C}}_t,\, t \in \mathbb{R}_+)$; $\hat{\mathbf{C}}^{\hat{\mathbf{\Pi}}} = \big(\hat{\mathcal{C}}_t^{\hat{\mathbf{\Pi}}},\, t \in \mathbb{R}_+\big)$ denotes the completion of the τ-flow $\hat{\mathbf{C}}$ with respect to a deviability $\hat{\mathbf{\Pi}}$ on $\hat{\mathbb{C}}$. Besides, "a.e." refers to Lebesgue measure unless specified otherwise.

Let functions $b_s(\mathbf{x})$ and $c_s(\mathbf{x})$, where $s \in \mathbb{R}_+$ and $\mathbf{x} \in \mathbb{C}$, with values in $\mathbb{R}^d$ and the set of positive semi-definite $d \times d$-matrices, respectively, be $\mathcal{C}_s$-measurable and continuous in $\mathbf{x} \in \mathbb{C}$ for $s \in \mathbb{R}_+$, Lebesgue measurable in s for $\mathbf{x} \in \mathbb{C}$, and such that the integrals $\int_0^t |b_s(\mathbf{x})|\, ds$ and $\int_0^t \operatorname{tr} c_s(\mathbf{x})\, ds$ are finite and continuous in $\mathbf{x} \in \mathbb{C}$ for all $t \in \mathbb{R}_+$. We define the function $G = (G_t(\lambda; \mathbf{x}),\, \mathbf{x} \in \mathbb{C},\, t \in \mathbb{R}_+,\, \lambda \in \mathbb{R}^d)$ by

$$G_t(\lambda; \mathbf{x}) = \int_0^t \lambda \cdot b_s(\mathbf{x})\, ds + \frac{1}{2} \int_0^t \lambda \cdot c_s(\mathbf{x}) \lambda\, ds. \tag{3.1}$$

Let us introduce the idempotent Ito equation

$$X_t = x_0 + \int_0^t b_s(X)\, ds + \int_0^t c_s(X)^{1/2}\, \dot{W}_s\, ds,\; t \in \mathbb{R}_+, \tag{3.2}$$

where $x_0 \in \mathbb{R}^d$. The next result extends Puhalskii [36, Theorem 2.6.30] by showing that the condition of the non-degeneracy of the diffusion matrix assumed there can be omitted.

Theorem 3.1. *Deviability $\mathbf{\Pi}_{x_0}$ on $\mathbb{C}$ solves the maxingale problem (x_0, G) if and only if $\mathbf{\Pi}_{x_0}$ is a Luzin weak solution of idempotent Ito equation* (3.2)*.*

We preview the proof with a lemma in the theme of the continuous mapping principle.

Lemma 3.1. *Let idempotent variables X and X_ϕ, $\phi \in \Phi$, where Φ is a directed set, with values in a metric space E be defined on idempotent probability spaces $(\Upsilon, \mathbf{\Pi})$ and $(\Upsilon_\phi, \mathbf{\Pi}_\phi)$, $\phi \in \Phi$, respectively. Let f be a function from E to a metric space E'. If the X_ϕ converge in idempotent distribution to X as $\phi \in \Phi$ and there exists a sequence of closed sets $F_m \subset E$, $m \in \mathbb{N}$, such that $\lim_{m\to\infty} \limsup_{\phi\in\Phi} \mathbf{\Pi}(X_\phi \notin F_m) = 0$ and f is continuous when restricted to each F_m, then the $f(X_\phi)$ converge in idempotent distribution to $f(X)$.*

Proof. Let f_m denote a continuous function that coincides with f on F_m, which exists by the Tietze-Urysohn extension theorem. Then the $f_m(X_\phi)$ converge in idempotent distribution to $f_m(X)$ as $\phi \in \Phi$ by the continuous mapping principle,

$$\limsup_{m\to\infty} \mathbf{\Pi}(f_m(X) \neq f(X)) \le \limsup_{m\to\infty} \mathbf{\Pi}(X \notin F_m) \le \limsup_{m\to\infty} \liminf_{\phi\in\Phi} \mathbf{\Pi}(X_\phi \notin F_m) = 0,$$

and

$$\limsup_{m\to\infty} \limsup_{\phi\in\Phi} \mathbf{\Pi}(f_m(X_\phi) \neq f(X_\phi)) \le \limsup_{m\to\infty} \limsup_{\phi\in\Phi} \mathbf{\Pi}(X_\phi \notin F_m) = 0.$$

The claim follows. □

Proof of Theorem 3.1. Let (3.2) have a Luzin weak solution. Let the idempotent processes (X, W) be defined on (Υ, Π) with a complete τ-flow $\mathbf{A}$. By Lemma 2.3 the idempotent process $\big(\int_0^t c_s(X)^{1/2}\, \dot{W}_s\, ds,\ t \in \mathbb{R}_+\big)$ is an $\mathbf{A}$-local maxingale on $(\Upsilon, \mathbf{\Pi})$ with quadratic characteristic $\big(\int_0^t c_s(X)\, ds,\ t \in \mathbb{R}_+\big)$. Therefore, (3.2) implies that $\big(\exp(\lambda \cdot (X_t - x_0) - \int_0^t \lambda \cdot b_s(X)\, ds - \int_0^t \lambda \cdot c_s(X)\lambda\, ds/2),\ t \in \mathbb{R}_+\big)$ is an $\mathbf{A}$-local exponential maxingale on (Υ, Π). It is then a local exponential maxingale relative to the τ-flow generated by X. The sufficiency part of the theorem follows by the transitivity law for conditional idempotent expectations, Puhalskii [36, Lemma 1.6.27].

We prove necessity. By hypotheses, under $\mathbf{\Pi}_{x_0}$ the idempotent process $\big(\exp(\lambda \cdot (\mathbf{x}_t - x_0) - \int_0^t \lambda \cdot b_s(\mathbf{x})\, ds - \int_0^t \lambda \cdot c_s(\mathbf{x})\lambda/2\, ds),\ t \in \mathbb{R}_+\big)$ is a local exponential maxingale relative to $\mathbf{C}$. Let X and W denote the coordinate processes on $\hat{\mathbb{C}}$, i.e., $X_t(\mathbf{x}, \mathbf{x}') = \mathbf{x}_t$ and $W_t(\mathbf{x}, \mathbf{x}') = \mathbf{x}'_t$, and let $\hat{\mathbf{\Pi}} = \mathbf{\Pi}_{x_0} \times \mathbf{\Pi}^W$. By the maxingale property of an idempotent Wiener process and Lemma 2.1, the idempotent process $\big(\exp(\lambda \cdot (X_t - x_0 + \sqrt{\epsilon} W_t) - \int_0^t \lambda \cdot b_s(X)\, ds - \int_0^t \lambda \cdot (c_s(X) + \epsilon I)\lambda/2\, ds,\ t \in \mathbb{R}_+\big)$, where I is the $d \times d$ identity matrix and $\epsilon > 0$, is a local exponential maxingale on $(\hat{\mathbb{C}}, \hat{\mathbf{\Pi}})$ relative to the canonical τ-flow. Hence, by Theorem 2.5.22 in Puhalskii [36], on $(\hat{\mathbb{C}}, \hat{\mathbf{\Pi}})$ there exists a standard idempotent Wiener process W^ϵ relative to $\hat{\mathbf{C}}^{\hat{\mathbf{\Pi}}}$ such that

$$X_t + \sqrt{\epsilon} W_t = x_0 + \int_0^t b_s(X)\, ds + \int_0^t (c_s(X) + \epsilon\, I)^{1/2}\, \dot{W}^\epsilon_s\, ds,\ t \in \mathbb{R}_+. \tag{3.3}$$

Let $\mathbf{\Pi}^\epsilon$ be the deviability on $\hat{\mathbb{C}}$ that is the idempotent distribution of (X, W^ϵ) on $(\hat{\mathbb{C}}, \hat{\mathbf{\Pi}})$. Since the idempotent distributions of X and W^ϵ are $\mathbf{\Pi}_{x_0}$ and $\mathbf{\Pi}^W$, respectively, the net $\mathbf{\Pi}^\epsilon$ is tight as $\epsilon \to 0$. Let a subnet $\mathbf{\Pi}^{\epsilon'}$ of $\mathbf{\Pi}^\epsilon$ weakly converge to a deviability $\tilde{\mathbf{\Pi}}$ on $\hat{\mathbb{C}}$. Since $\tilde{\mathbf{\Pi}}$ has $\mathbf{\Pi}_{x_0}$ and $\mathbf{\Pi}^W$ as the coordinate projections, we can assume that the $(X, W^{\epsilon'})$ converge in idempotent distribution to the canonical idempotent process $(X, \tilde{W})$ on $(\hat{\mathbb{C}}, \tilde{\mathbf{\Pi}})$ and that $\tilde{W}$ is a standard idempotent Wiener process.

Let us show that $\tilde{W}$ is idempotent Wiener relative to $\hat{\mathbf{C}}^{\tilde{\mathbf{\Pi}}}$. It suffices to prove that $\tilde{W}_t - \tilde{W}_s$ is independent of $\hat{\mathcal{C}}_s$ on $(\hat{\mathbb{C}}, \tilde{\mathbf{\Pi}})$ for $s \le t$. Since W^ϵ is standard idempotent Wiener relative to $\hat{\mathbf{C}}^{\hat{\mathbf{\Pi}}}$ and X is $\hat{\mathbf{C}}^{\hat{\mathbf{\Pi}}}$-adapted, $W^\epsilon_t - W^\epsilon_s$ is independent of (X_r, W^ϵ_r) on $(\hat{\mathbb{C}}, \hat{\mathbf{\Pi}})$ for $r \le s$. By convergence in idempotent distribution of the $(X, W^{\epsilon'})$ to $(X, \tilde{W})$, $\tilde{W}_t - \tilde{W}_s$ is independent of $(X_r, \tilde{W}_r)$ on $(\hat{\mathbb{C}}, \tilde{\mathbf{\Pi}})$ for $r \le s$. The required property follows since $\hat{\mathcal{C}}_s$ is generated by $(X_r, \tilde{W}_r)$, $r \le s$.

We next prove that

$$X_t = x_0 + \int_0^t b_s(X)\, ds + \int_0^t c_s(X)^{1/2}\, \dot{\tilde{W}}_s\, ds,\ t \in \mathbb{R}_+,\ \tilde{\mathbf{\Pi}}\text{-a.e.} \tag{3.4}$$

We have by (3.3) for $T > 0$ and $\delta > 0$

$$\hat{\mathbf{\Pi}}\Big(\sup_{t \in [0,T]} |X_t - \big(x_0 + \int_0^t b_s(X)\, ds + \int_0^t c_s(X)^{1/2}\, \dot{W}^\epsilon_s\, ds\big)| > \delta\Big)$$
$$\le \hat{\mathbf{\Pi}}\Big(\sup_{t \in [0,T]} |\sqrt{\epsilon} W_t| > \frac{\delta}{2}\Big) \vee \hat{\mathbf{\Pi}}\Big(\sup_{t \in [0,T]} |\int_0^t \big(c_s(X) + \epsilon\, I)^{1/2} - c_s(X)^{1/2}\big) \dot{W}^\epsilon_s\, ds| > \frac{\delta}{2}\Big). \tag{3.5}$$

Since W is a local maxingale with quadratic characteristic $(t\, I,\ t \in \mathbb{R}_+)$, by Lemma 2.2 $\hat{\mathbf{\Pi}}\big(\sup_{t \in [0,T]} |\sqrt{\epsilon} W_t| > \delta/2\big) \le \exp\big(-\delta/(2\sqrt{\epsilon})\big) \exp(T/2)$, which tends to 0 as $\epsilon \to 0$.

Next, by Lemmas 2.2 and 2.3,

$$\hat{\mathbf{\Pi}}\Big(\sup_{t\in[0,T]}\Big|\int\limits_0^t\big((c_s(X)+\epsilon I)^{1/2}-c_s(X)^{1/2}\big)\dot{W}_s^\epsilon\,ds\Big|>\frac{\delta}{2}\Big)$$

$$\le e^{-\delta/(4\sqrt{\epsilon})}\vee\hat{\mathbf{\Pi}}\Big(\int\limits_0^T\big\|\big((c_s(X)+\epsilon I)^{1/2}-c_s(X)^{1/2}\big)^2\big\|\,ds\ge\frac{\delta\sqrt{\epsilon}}{2}\Big).$$

Since $\big\|\big((c_s(X)+\epsilon I)^{1/2}-c_s(X)^{1/2}\big)^2\big\|\le\epsilon$, the latter idempotent probability tends to 0 as $\epsilon\to0$, so the left-hand side tends to 0 as $\epsilon\to0$. Hence, the left-hand side of (3.5) tends to 0 as $\epsilon\to0$.

In addition, on noting that under the hypotheses the mapping $\mathbf{x}\to(c_s(\mathbf{x}),\,s\in[0,t])$ is continuous in $\mathbf{x}\in\mathbb{C}$ with respect to the norm $\int_0^t\|c_s(\mathbf{x})\|\,ds$, we have that the function $(\mathbf{x},\mathbf{y})\to\big(\int_0^t c_s(\mathbf{x})^{1/2}\,\dot{\mathbf{y}}_s\,ds,\,t\in\mathbb{R}_+\big)$ from $\hat{\mathbb{C}}$ to $\mathbb{C}$, which is defined arbitrarily when the integral is not well defined, is continuous when restricted to the sets $F_m=\{(\mathbf{x},\mathbf{y}):\,\int_0^\infty|\dot{\mathbf{y}}_s|^2\,ds\le m\}$ for $m\in\mathbb{N}$. Therefore, in view of Lemma 3.1

$$\liminf_{\epsilon'\to0}\hat{\mathbf{\Pi}}\Big(\sup_{t\in[0,T]}\Big|X_t-\Big(x_0+\int\limits_0^t b_s(X)\,ds+\int\limits_0^t c_s(X)^{1/2}\,\dot{W}_s^{\epsilon'}\,ds\Big)\Big|>\delta\Big)$$

$$\ge\tilde{\mathbf{\Pi}}\Big(\sup_{t\in[0,T]}\Big|X_t-\Big(x_0+\int\limits_0^t b_s(X)\,ds+\int\limits_0^t c_s(X)^{1/2}\,\dot{\tilde{W}}_s\,ds\Big)\Big|>\delta\Big).$$

Equality (3.4) follows. Since X has idempotent distribution $\mathbf{\Pi}_{x_0}$ and $\tilde{W}$ is a standard idempotent Wiener process relative to $\hat{\mathbf{C}}^{\tilde{\mathbf{\Pi}}}$ under $\tilde{\mathbf{\Pi}}$, $\mathbf{\Pi}_{x_0}$ is a Luzin weak solution of (3.4). □

As a consequence of the theorem, uniqueness holds for (x_0,G) if and only if equation (3.2) has a unique Luzin weak solution. We now provide sufficient conditions for the equation to have a Luzin weak solution, which extend the results of Puhalskii [36, 37]. Let us assume that for every $a>0$ there exists a locally integrable in s nonnegative function r_s^a, $s\in\mathbb{R}_+$, such that $|b_t(\mathbf{x})|+\|c_t(\mathbf{x})\|\le r_t^a$ for all $t\in\mathbb{R}_+$ and $\mathbf{x}\in\mathbb{C}$ satisfying the condition $\sup_{s\le t}|\mathbf{x}_s|\le a$. Let there exist a differentiable real-valued function $V(x)$, $x\in\mathbb{R}^d$, such that $|V(x)|\to\infty$ as $|x|\to\infty$, an increasing real-valued positive function $k(q)$, $q\in\mathbb{R}_+$, such that $k(q)\to\infty$ as $q\to\infty$ and $\int_1^\infty 1/k(q)\,dq=\infty$, and a locally integrable $\mathbb{R}_+$-valued function l_s, $s\in\mathbb{R}_+$, for which

$$\nabla V(\mathbf{x}_t)\cdot b_t(\mathbf{x})\le l_t\big(1+\sup_{s\le t}k(|V(\mathbf{x}_s)|)\big),\quad t\in\mathbb{R}_+,\,\mathbf{x}\in\mathbb{C},\tag{3.6a}$$

$$\nabla V(\mathbf{x}_t)\cdot c_t(\mathbf{x})\nabla V(\mathbf{x}_t)\le l_t\big(1+\sup_{s\le t}k(|V(\mathbf{x}_s)|)^2\big),\quad t\in\mathbb{R}_+,\,\mathbf{x}\in\mathbb{C},\tag{3.6b}$$

where ∇ denotes gradient.

Let $c_s(\mathbf{x})^\oplus$ denote the pseudo-inverse of $c_s(\mathbf{x})$ and let idempotent probability $\mathbf{\Pi}_{x_0}^*$ on $\mathbb{C}$ be defined by

$$\mathbf{\Pi}_{x_0}^*(\mathbf{x})=\exp\Big(-\frac{1}{2}\int\limits_0^\infty\big(\dot{\mathbf{x}}_s-b_s(\mathbf{x})\big)\cdot c_s(\mathbf{x})^\oplus\big(\dot{\mathbf{x}}_s-b_s(\mathbf{x})\big)\,ds\Big)\tag{3.7}$$

if $\mathbf{x}$ is absolutely continuous, $\mathbf{x}_0=x_0$, and $\dot{\mathbf{x}}_s-b_s(\mathbf{x})$ is in the range of $c_s(\mathbf{x})$ a.e., and $\mathbf{\Pi}_{x_0}^*(\mathbf{x})=0$ otherwise. The next lemma implies, in particular, that $\mathbf{\Pi}_{x_0}^*$ is a deviability.

Lemma 3.2. *Under the introduced hypotheses, $\mathbf{\Pi}^*_{x_0}$ is a Luzin weak solution of equation* (3.2).

Proof. Let $\sigma_t(\mathbf{x}) = c_t(\mathbf{x})^{1/2}$. We first prove that the ordinary differential equation

$$\dot{\mathbf{x}}_t = b_t(\mathbf{x}) + \sigma_t(\mathbf{x})\dot{\mathbf{w}}_t \text{ a.e.}, \quad \mathbf{x}_0 = x_0, \tag{3.8}$$

where $\mathbf{w}_t$ is an absolutely continuous function such that $\int_0^\infty |\dot{\mathbf{w}}_t|^2\, dt < \infty$, satisfies the extension condition, Puhalskii [36], i.e., given arbitrary $T > 0$, every absolutely continuous function $\hat{\mathbf{x}}_t$ that satisfies (3.8) on $[0, T]$ can be extended to a solution on $[T, \infty)$. Let, given $N \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{C}$, $\tau^N(\mathbf{x}) = \inf\{t \in \mathbb{R}_+ : |\mathbf{x}_t| + t \geq N\}$, $\tilde{p}_N\mathbf{x} = (\mathbf{x}_{t\wedge\tau^N(\mathbf{x})}, t \in \mathbb{R}_+)$, $b^N_t(\mathbf{x}) = b_t(\tilde{p}_N\mathbf{x})$ and $\sigma^N_t(\mathbf{x}) = \sigma_t(\tilde{p}_N\mathbf{x})$. By hypotheses, the functions $b^N_t(\mathbf{x})$ and $\sigma^N_t(\mathbf{x})$ are continuous in $\mathbf{x}$, and the functions $|b^N_t(\mathbf{x})|$ and $\|\sigma^N_t(\mathbf{x})\|^2$ are bounded by the locally integrable function r^N_t, so the equations

$$\dot{\mathbf{x}}_t = b^N_t(\mathbf{x}) + \sigma^N_t(\mathbf{x})\dot{\mathbf{w}}_t,\ t \geq T \text{ a.e.}, \text{ where } \mathbf{x}_t = \hat{\mathbf{x}}_t \text{ for } t \leq T,$$

have solutions $\mathbf{x}^N = (\mathbf{x}^N_t, t \in \mathbb{R}_+)$ (the latter follows, e.g., by the method of successive approximations). We have by (3.6a) and (3.6b) for $t \geq T$ a.e.

$$\begin{aligned}\frac{d}{dt}V(\mathbf{x}^N_t) &= \nabla V(\mathbf{x}^N_t)\cdot b_t(\mathbf{x}^N) + \nabla V(\mathbf{x}^N_t)\cdot\sigma_t(\mathbf{x}^N)\dot{\mathbf{w}}_t \\ &\leq \big(l_t + \sqrt{l_t}|\dot{\mathbf{w}}_t|\big)\big(1 + \sup_{r\leq T} k\big(|V(\hat{\mathbf{x}}_r)|\big)\big) + \big(l_t + \sqrt{l_t}|\dot{\mathbf{w}}_t|\big)\sup_{T\leq r\leq t} k\big(|V(\mathbf{x}^N_r)|\big).\end{aligned}$$

Let $q_0 \in \mathbb{R}_+$ be such that $k(q) \geq 1$ for $q \geq q_0$. Denoting $y^N_t = \int_T^t \big(1 + l_s + \sqrt{l_s}|\dot{\mathbf{w}}_s|\big)\sup_{T\leq r\leq s} k\big(|V(\mathbf{x}^N_r)|\big)\, ds$ and $z_t = q_0 + V(\hat{\mathbf{x}}_T) + \big(1 + \sup_{r\leq T} k\big(|V(\hat{\mathbf{x}}_r)|\big)\big)\int_T^t \big(l_s + \sqrt{l_s}|\dot{\mathbf{w}}_s|\big)\, ds$, we have that

$$\int\limits_{z_t}^{z_t+y^N_t} \frac{dq}{k(q)} = \int\limits_T^t \frac{dy^N_s}{k(z_t + y^N_s)} \leq \int\limits_T^t \big(1 + l_s + \sqrt{l_s}|\dot{\mathbf{w}}_s|\big)\, ds, \tag{3.9}$$

which implies that the y^N_t are bounded in N for a given t. Therefore, for suitable constants K_t, $t \geq T$,

$$\sup_{T\leq s\leq t} |\mathbf{x}^N_s| \leq K_t,\ N \in \mathbb{N},\ t \geq T. \tag{3.10}$$

Thus, if N is large enough, then $|\mathbf{x}^N_t| + t < N$ on $[T, T+1]$, $(\tilde{p}_N\mathbf{x}^N)_t = \mathbf{x}^N_t$ and, hence, $b_t(\tilde{p}_N\mathbf{x}^N) = b_t(\mathbf{x}^N)$ and $\sigma_t(\tilde{p}_N\mathbf{x}^N) = \sigma_t(\mathbf{x}^N)$ for $t \in [0, T+1]$. For these N the function $\mathbf{x} = \mathbf{x}^N$ satisfies equation (3.8) on $[T, T+1]$ and coincides with $\hat{\mathbf{x}}$ on $[0, T]$. Replacing T with $T+1$, we can extend $\mathbf{x}$ to a solution of (3.8) on $[T, T+2]$ and so forth.

Let idempotent probability $\hat{\mathbf{\Pi}}$ on $\hat{\mathbb{C}}$ be defined by $\hat{\mathbf{\Pi}}(\mathbf{x}, \mathbf{w}) = \mathbf{\Pi}^W(\mathbf{w})$ if $(\mathbf{x}, \mathbf{w})$ satisfies (3.8) and $\hat{\mathbf{\Pi}}(\mathbf{x}, \mathbf{w}) = 0$ otherwise. Then the first component process on $\hat{\mathbb{C}}$ has idempotent distribution $\mathbf{\Pi}^*_{x_0}$ under $\hat{\mathbf{\Pi}}$, the second component process has idempotent distribution $\mathbf{\Pi}^W$, and (3.2) holds $\hat{\mathbf{\Pi}}$-a.e. In addition, by the argument of the proof of part 1 of Lemma 2.6.17 in Puhalskii [36], the second component process is standard idempotent Wiener relative to $\hat{\mathbf{C}}^{\hat{\mathbf{\Pi}}}$. Thus, $\mathbf{\Pi}^*_{x_0}$ is a weak solution to (3.2). It is a Luzin weak solution if $\hat{\mathbf{\Pi}}$ is a deviability on $\hat{\mathbb{C}}$. Upper semi-continuity of $\hat{\mathbf{\Pi}}(\mathbf{x}, \mathbf{w})$ is established similarly to how it is done in the proof of Theorem 2.6.24 in Puhalskii [36]. To show that the sets $\{(\mathbf{x}, \mathbf{w}) : \hat{\mathbf{\Pi}}(\mathbf{x}, \mathbf{w}) \geq a\}$ are, moreover, compact for $a \in (0, 1]$, let us consider sequences $\mathbf{w}^{(n)} \in \mathbb{C}$, $n \in \mathbb{N}$, and $\mathbf{x}^{(n)} \in \mathbb{C}$, $n \in \mathbb{N}$, such that

$$\dot{\mathbf{x}}^{(n)}_t = b_t(\mathbf{x}^{(n)}) + \sigma_t(\mathbf{x}^{(n)})\dot{\mathbf{w}}^{(n)}_t \text{ a.e.}, \quad \mathbf{x}^{(n)}_0 = x_0, \tag{3.11}$$

$\mathbf{\Pi}^W(\mathbf{w}^{(n)}) \geq a$, and $\mathbf{w}^{(n)} \to \mathbf{w}$ as $n \to \infty$. Let us note that an analogue of inequality (3.9) also holds for $(\mathbf{x}^{(n)}, \mathbf{w}^{(n)})$, so in analogy with (3.10) $\limsup_{n\to\infty} \sup_{s\in[0,t]} |\mathbf{x}_s^{(n)}| < \infty$ for $t \in \mathbb{R}_+$. Also by (3.11) for $0 \leq s \leq t$

$$|\mathbf{x}_t^{(n)} - \mathbf{x}_s^{(n)}| \leq \int_s^t |b_r(\mathbf{x}^{(n)})| \, dr + \sqrt{\int_s^t \|c_r(\mathbf{x}^{(n)})\| \, dr} \sqrt{\int_s^t |\dot{\mathbf{w}}_r^{(n)}|^2 \, dr}.$$

An application of Arzelà–Ascoli's theorem completes the proof. □

As mentioned in Section 2, pathwise uniqueness holds for (3.2) under Lipshitz continuity conditions, so there exists a unique weak solution, which is in fact a Luzin strong solution. The following uniqueness result relaxes these requirements.

Corollary 3.1. *Let, in addition to the hypotheses of Lemma 3.2, for every $a > 0$ there exist a differentiable real-valued function $V^a(x)$, $x \in \mathbb{R}^d$, such that $V^a(x) \neq 0$ if $x \neq 0$, an increasing $\mathbb{R}_+$-valued function $k^a(q)$, $q \in \mathbb{R}_+$, such that $k^a(0) = 0$, $k^a(q) > 0$ for $q > 0$ and $\int_0^\epsilon dq/k^a(q) = \infty$ for $\epsilon > 0$, and a locally integrable $\mathbb{R}_+$-valued function l_t^a, $t \in \mathbb{R}_+$, for which the inequalities $\nabla V^a(\mathbf{x}_t - \mathbf{y}_t) \cdot \big(b_t(\mathbf{x}) - b_t(\mathbf{y})\big) \leq l_t^a \sup_{s\in[0,t]} k^a\big(|V^a(\mathbf{x}_s - \mathbf{y}_s)|\big)$ and $\nabla V^a(\mathbf{x}_t - \mathbf{y}_t) \cdot \big(c_t(\mathbf{x})^{1/2} - c_t(\mathbf{y})^{1/2}\big)^2 \nabla V^a(\mathbf{x}_t - \mathbf{y}_t) \leq l_t^a \sup_{s\in[0,t]} k^a\big(|V^a(\mathbf{x}_s - \mathbf{y}_s)|\big)^2$ hold when $\sup_{s\in[0,t]}|\mathbf{x}_s| \leq a$ and $\sup_{s\in[0,t]}|\mathbf{y}_s| \leq a$. Then (3.2) has a unique strong solution, which is a Luzin strong solution. This solution has idempotent distribution $\mathbf{\Pi}_{x_0}^*$, which is a unique weak solution and is a unique Luzin weak solution of (3.2).*

The proof is omitted since it is much the same as that of the next corollary, which permits non-Lipshitz behaviour for the Markov setting. The result is analogous to the one from the theory of ordinary differential equations, see, e.g., Coddington and Levinson [5], Hartman [21].

Corollary 3.2. *Let $b_s(\mathbf{x}) = \hat{b}_s(\mathbf{x}_s)$ and $c_s(\mathbf{x}) = \hat{c}_s(\mathbf{x}_s)$, where the functions $\hat{b}_s(x)$ and $\hat{c}_s(x)$ are Lebesgue-measurable in $s \in \mathbb{R}_+$, continuous in $x \in \mathbb{R}^d$, and such that the functions $\operatorname{ess\,sup}_{x\in\mathbb{R}^d:\,|x|\leq A}|\hat{b}_s(x)|$ and $\operatorname{ess\,sup}_{x\in\mathbb{R}^d:\,|x|\leq A}\|\hat{c}_s(x)\|$ are locally integrable in s for $A \in \mathbb{R}_+$. Let the following growth conditions be satisfied: $\nabla V(x) \cdot \hat{b}_t(x) \leq l_t\big(1 + k(|V(x)|)\big)$, $\nabla V(x) \cdot \hat{c}_t(x) \nabla V(x) \leq l_t\big(1 + k(|V(x)|)^2\big)$, where the functions $V(x)$, $k(q)$, and l_t are as in the hypotheses of Lemma 3.2. Let for every $a > 0$ there exist functions $V^a(x)$, $k^a(q)$, and l_t^a as in the statement of Corollary 3.1 such that the inequalities $\nabla V^a(x - y) \cdot (\hat{b}_t(x) - \hat{b}_t(y)) \leq l_t^a \, k^a\big(|V^a(x - y)|\big)$ and $\nabla V^a(x - y) \cdot \big(\hat{c}_t(x)^{1/2} - \hat{c}_t(y)^{1/2}\big)^2 \nabla V^a(x - y) \leq l_t^a \, k^a(|V^a(x - y)|)^2$ hold when $|x| \leq a$ and $|y| \leq a$. Then (3.2) has a unique strong solution, which is a Luzin strong solution. This solution has idempotent distribution $\mathbf{\Pi}_{x_0}^*$, which is a unique weak solution and is a unique Luzin weak solution of (3.2).*

Proof. By Lemma 3.2 under the hypotheses there exists a Luzin weak solution, so it suffices to prove pathwise uniqueness for (3.2). Let X and Y be two solutions. Given arbitrary $a > 0$ and $T > 0$, we define $\tau^a = \inf\{t \in \mathbb{R}_+ : |X_t| \vee |Y_t| \geq a\} \wedge T$. Then a.e. for $t \leq \tau^a$

$$\begin{aligned} \frac{d}{dt} V^a(X_t - Y_t) &= \nabla V^a(X_t - Y_t) \cdot (\hat{b}_t(X_t) - \hat{b}_t(Y_t)) + \nabla V^a(X_t - Y_t) \cdot (\hat{c}_t(X_t)^{1/2} - \hat{c}_t(Y_t)^{1/2}) \dot{W}_t \\ &\leq \big(l_t^a + \sqrt{l_t^a}|\dot{W}_t|\big) \, k^a\big(|V^a(X_t - Y_t)|\big). \end{aligned}$$

Let $\varphi_t = |V^a(X_t - Y_t)|$ and $\psi_t = 1 + l_t^a + \sqrt{l_t^a}|\dot{W}_t|$. Then ψ_t is positive, locally integrable, and $\varphi_t \leq \int_0^t \psi_s \, k^a(\varphi_s) \, ds$. Denoting as y_t the right-hand side of the latter inequality, we have that

$\int_0^t dy_s/k^a(\epsilon+y_s) \le \int_0^t \psi_s\,ds$ for arbitrary $\epsilon > 0$. Hence, the integral $\int_\epsilon^{\epsilon+y_t} dq/k^a(q) = \int_0^t dy_s/k^a(\epsilon+y_s)$ is bounded as $\epsilon \to 0$, so $y_t = 0$. We have thus proved that $X_t = Y_t$ for $t \le \tau^a$. Letting $a \to \infty$, we conclude that $X_t = Y_t$ for $t \le T$. □

Remark 3.1. The most common choice of functions V, V^a, k, and k^a is $V(x) = V^a(x) = |x|^2$ and $k(q) = k^a(q) = q$. One then obtains extensions of linear-growth and Lipshitz continuity conditions.

Remark 3.2. As an example of non-Lipshitz-continuous coefficients with super-linear growth at infinity, one can take $\hat{c}_t(x)^{1/2} = \Lambda_1 \text{diag}(|x_1 \log|x_1||, \ldots, |x_d \log|x_d||)\Lambda_2$ and $\hat{b}_t(x) = \sum_{i=1}^d v_i x_i \log|x_i|$, where $x = (x_1, \ldots, x_d)^T$, the v_i are elements of $\mathbb{R}^d$, and Λ_1 and Λ_2 are $d \times d$ matrices.

4 Applications to large deviations of stochastic equations

We now consider applications to large deviation convergence of solutions of stochastic equations to solutions of idempotent Ito equations with possibly degenerate coefficients. In the standard large deviation theory classification, this is the setting of moderate deviations. To recall the terminology, we say that a net $\mathbf{P}_\phi$, $\phi \in \Phi$, Φ being a directed set, of probability measures on the Borel σ-algebra of a metric space Υ large deviation (LD) converges as $\phi \in \Phi$ at rate k_ϕ, where $k_\phi \to \infty$ as $\phi \in \Phi$, to a deviability $\mathbf{\Pi}$ on Υ if $\lim_{\phi\in\Phi}\bigl(\int_\Upsilon f(\upsilon)^{k_\phi}\,d\mathbf{P}_\phi(\upsilon)\bigr)^{1/k_\phi} = \sup_{\upsilon\in\Upsilon} f(\upsilon)\mathbf{\Pi}(\{\upsilon\})$ for every continuous bounded $\mathbb{R}_+$-valued function f on Υ. As a matter of fact, the net $\mathbf{P}_\phi$, $\phi \in \Phi$, LD converges at rate k_ϕ to $\mathbf{\Pi}$ if and only if it obeys the LDP with action functional $\mathbf{I} = -\log \mathbf{\Pi}$ for scale k_ϕ, so LD convergence is another way of formulating LDP results, which we find more natural for our approach. Next, a net of random variables X_ϕ, $\phi \in \Phi$, with values in Υ is said to LD converge in distribution (denoted as $X_\phi \xrightarrow{ld} X$) at rate k_ϕ to a Luzin idempotent variable X with values in Υ if the net of laws of the X_ϕ LD converges at rate k_ϕ to the idempotent distribution of X. We also say that the net X_ϕ, $\phi \in \Phi$, converges to a point $\upsilon \in \Upsilon$ superexponentially in probability at rate k_ϕ and write $X_\phi \xrightarrow{\mathbf{P}_\phi^{1/k_\phi}} \upsilon$ as $\phi \in \Phi$ if $\lim_{\phi\in\Phi} \mathbf{P}_\phi(\rho(X_\phi, \upsilon) > \epsilon)^{1/k_\phi} = 0$ for all $\epsilon > 0$, where ρ is the metric on Υ. The theorems on LD convergence presented in this section refer to the Skorohod space $\mathbb{D}$ of right-continuous with left-hand limits $\mathbb{R}^d$-valued functions $\mathbf{x} = (\mathbf{x}_t,\, t \in \mathbb{R}_+)$ endowed with the Skorohod-Prohorov-Lindvall metric, see, e.g., Jacod and Shiryaev [23] for more detail. We denote as $\mathbf{x}_{s-}$ the left-hand limit of $\mathbf{x}$ at s.

The first application expands on Example 5.4.6 in Puhalskii [36]. Let $\mathcal{B}(E)$ denote the Borel σ-algebra of a metric space E and $\mathcal{P}(\mathbb{D})$ denote the predictable σ-algebra on $\mathbb{R}_+ \times \mathbb{D}$. Let $(U,\mathcal{U})$ be a measurable space and $m(d\vartheta)$ be a non-negative σ-finite measure on $(U,\mathcal{U})$. We assume as given functions $b_s^{\epsilon,\gamma}(\mathbf{x}), \sigma_s^{\epsilon,\gamma}(\mathbf{x})$, and $f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)$, where $\epsilon > 0$, $\gamma > 0$, $s \in \mathbb{R}_+$, $\mathbf{x} \in \mathbb{D}$, and $\vartheta \in U$, which are $\mathcal{P}(\mathbb{D})/\mathcal{B}(\mathbb{R}^d)$, $\mathcal{P}(\mathbb{D})/\mathcal{B}(\mathbb{R}^d \times \mathbb{R}^{d'})$, and $\mathcal{P}(\mathbb{D}) \otimes \mathcal{U}/\mathcal{B}(\mathbb{R}^d)$-measurable, respectively. Let there exist weak solutions to the stochastic equations

$$X_t^{\epsilon,\gamma} = x_0 + \int_0^t b_s^{\epsilon,\gamma}(X^{\epsilon,\gamma})\,ds + \sqrt{\epsilon}\int_0^t \sigma_s^{\epsilon,\gamma}(X^{\epsilon,\gamma})\,dB_s^\epsilon + \sqrt{\epsilon\gamma}\int_0^t\int_U f_s^{\epsilon,\gamma}(X^{\epsilon,\gamma},\vartheta)\left[\mathcal{N}^\gamma(ds,d\vartheta) - \gamma^{-1}ds\,m(d\vartheta)\right],$$

where $x_0 \in \mathbb{R}^d$, the $B^\epsilon = (B_s^\epsilon, s \in \mathbb{R}_+)$ are $\mathbb{R}^{d'}$-valued standard Wiener processes, and the $\mathcal{N}^\gamma = (\mathcal{N}^\gamma(ds,d\vartheta))$ are Poisson random measures on $(\mathbb{R}_+ \times U, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{U})$ with intensity measures $\gamma^{-1}\,ds\,m(d\vartheta)$. (The existence holds, e.g., if $b_s^{\epsilon,\gamma}(\mathbf{x})$ and $\sigma_s^{\epsilon,\gamma}(\mathbf{x})$ are continuous in $\mathbf{x} \in \mathbb{D}$,

$\lim_{\mathbf{y}\to\mathbf{x}}\int_U |f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)-f_s^{\epsilon,\gamma}(\mathbf{y},\vartheta)|^2 m(d\vartheta)=0$, $|b_s^{\epsilon,\gamma}(\mathbf{x})|^2+\|\sigma_s^{\epsilon,\gamma}(\mathbf{x})\|^2\le l_s^{\epsilon,\gamma}(1+\sup_{t\le s}|\mathbf{x}_t|^2)$, and $|f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)|\le h_s^{\epsilon,\gamma}(\vartheta)(1+\sup_{t\le s}|\mathbf{x}_t|)$, where $l_s^{\epsilon,\gamma}$ and $h_s^{\epsilon,\gamma}(\vartheta)$ are $\mathbb{R}_+$-valued and increasing in s, $h_s^{\epsilon,\gamma}(\vartheta)$ is $\mathcal{U}/\mathcal{B}(\mathbb{R}_+)$–measurable for every $s\in\mathbb{R}_+$, and $\int_U h_s^{\epsilon,\gamma}(\vartheta)^2\,m(d\vartheta)<\infty$, Gihman and Skorohod [20]. Another set of conditions is given by assuming that $b_s^{\epsilon,\gamma}(\mathbf{x})=\hat{b}_s^{\epsilon,\gamma}(\mathbf{x}_{s-})$, $\sigma_s^{\epsilon,\gamma}(\mathbf{x})=\hat{\sigma}_s^{\epsilon,\gamma}(\mathbf{x}_{s-})$, and $f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)=\hat{f}_s^{\epsilon,\gamma}(\mathbf{x}_{s-},\vartheta)$, where $\hat{b}_s^{\epsilon,\gamma}(x),\hat{\sigma}_s^{\epsilon,\gamma}(x)$, and $\hat{f}_s^{\epsilon,\gamma}(x,\vartheta)$ are $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R}^d)$, $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R}^d\times\mathbb{R}^{d'})$, and $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{U}/\mathcal{B}(\mathbb{R}^d)$-measurable, respectively, $\hat{b}_s^{\epsilon,\gamma}(x)$ and $\hat{\sigma}_s^{\epsilon,\gamma}(x)$ are continuous in $x\in\mathbb{R}^d$, $\lim_{y\to x}\int_U|\hat{f}_s^{\epsilon,\gamma}(y,\vartheta)-\hat{f}_s^{\epsilon,\gamma}(x,\vartheta)|^2 m(d\vartheta)=0$, $|\hat{b}_s^{\epsilon,\gamma}(x)|^2+\|\hat{\sigma}_s^{\epsilon,\gamma}(x)\|^2\le\hat{l}_s^{\epsilon,\gamma}(1+|x|^2)$, and $|\hat{f}_s^{\epsilon,\gamma}(x,\vartheta)|\le\hat{h}_s^{\epsilon,\gamma}(\vartheta)(1+|x|)$, where $\hat{l}_s^{\epsilon,\gamma}$ and $\hat{h}_s^{\epsilon,\gamma}(\vartheta)$ are $\mathbb{R}_+$-valued and increasing in s, $\hat{h}_s^{\epsilon,\gamma}(\vartheta)$ is $\mathcal{U}/\mathcal{B}(\mathbb{R}_+)$–measurable for every $s\in\mathbb{R}_+$, and $\int_U\hat{h}_s^{\epsilon,\gamma}(\vartheta)^2\,m(d\vartheta)<\infty$, Gihman and Skorohod [20], Ikeda and Watanabe [22], and Jacod and Shiryaev [23]. To ensure uniqueness one may require, in addition, that either the coefficients be Lipshitz-continuous or that they be bounded and a non-degeneracy condition hold, see Gihman and Skorohod [20], Ikeda and Watanabe [22], and Jacod and Shiryaev [23] for more detail. We also remark that for diffusions without jumps the linear-growth and (or) Lipshitz continuity conditions can be relaxed in analogy with the hypotheses of Lemma 3.2, Corollary 3.1, and Corollary 3.2, cf., Narita [30], Ethier and Kurtz [12], Ikeda and Watanabe [22], Fang and Zhang [13].)

Let us denote $\mathbf{x}_t^*=\sup_{s\le t}|\mathbf{x}_s|$ and introduce the following moment conditions on the jumps of the $X^{\epsilon,\gamma}$:

($\tilde{\text{P}}$) for some $p>2$

$$\limsup_{\substack{\epsilon\to0\\ \gamma\to0}}\ \sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A}\int_0^t\int_U|f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)|^p\,m(d\vartheta)\,ds<\infty,\ t\in\mathbb{R}_+,\ A\in\mathbb{R}_+,$$

($\widetilde{\text{SE}}$) for some $\beta\in(0,1]$ and $\alpha>0$

$$\limsup_{\substack{\epsilon\to0\\ \gamma\to0}}\ \sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A}\int_0^t\int_U\exp\bigl(\alpha|f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)|^\beta\bigr)\mathbf{1}\bigl(|f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)|\ge1\bigr)\,m(d\vartheta)\,ds<\infty,\ t\in\mathbb{R}_+,\ A\in\mathbb{R}_+.$$

It is assumed that the functions $b_s(\mathbf{x})$ and $c_s(\mathbf{x})$ as defined at the beginning of Section 3 are extended for $\mathbf{x}\in\mathbb{D}$ such that they are locally integrable in s, are adapted to the canonical τ-flow on $\mathbb{D}$, and are continuous in $\mathbf{x}$ at $\mathbf{x}\in\mathbb{C}$. (The canonical τ-flow on $\mathbb{D}$ is defined similarly to the one on $\mathbb{C}$ as consisting of τ-algebras generated by projection mappings.)

Let us denote $c_s^{\epsilon,\gamma}(\mathbf{x})=\sigma_s^{\epsilon,\gamma}(\mathbf{x})\sigma_s^{\epsilon,\gamma}(\mathbf{x})^T+\int_U f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)f_s^{\epsilon,\gamma}(\mathbf{x},\vartheta)^T\,m(d\vartheta)$.

Theorem 4.1. *Let equation* (3.2) *have a unique Luzin weak solution X. Let for arbitrary $t\in\mathbb{R}_+$ and $A\in\mathbb{R}_+$*

$$\int_0^t\operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_s^*\le A}|b_s(\mathbf{x})|\,ds<\infty,\qquad\int_0^t\operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_s^*\le A}\|c_s(\mathbf{x})\|\,ds<\infty.$$

Let, as $\epsilon\to0$ and $\gamma\to0$,

$$\sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A}\int_0^t|b_s^{\epsilon,\gamma}(\mathbf{x})-b_s(\mathbf{x})|\,ds\to0\quad\textit{and}\quad\sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A}\int_0^t\|c_s^{\epsilon,\gamma}(\mathbf{x})-c_s(\mathbf{x})\|\,ds\to0.$$

If, in addition, either condition ($\tilde{P}$) holds and $\epsilon\log(\gamma^{-1})\to\infty$, or condition ($\widetilde{SE}$) holds and $\epsilon^{2-\beta}/\gamma^\beta\to\infty$, then $X^{\epsilon,\gamma}\stackrel{ld}{\to}X$ at rate $1/\epsilon$.

Proof. The proof consists in checking the hypotheses of Theorem 5.2.12 in Puhalskii [36] and proceeds similarly to the proof of Lemma 8 in Puhalskii [37]. There are several groups of conditions to be checked.

The first group concerns the limit idempotent equation. For the setting at hand these are the continuity and local majoration conditions on $B_t(\mathbf{x}) = \int_0^t b_s(\mathbf{x})\,ds$ and $C_t(\mathbf{x}) = \int_0^t c_s(\mathbf{x})\,ds$, see p.375 of Puhalskii [36]. These conditions readily follow from the hypotheses. There is also condition (NE), which requires certain regularity properties of deviability $\mathbf{\Pi}^*_{x_0}$ from (3.7): not only does the function $\mathbf{\Pi}^*_{x_0}(\mathbf{x})$ have to be upper compact in $\mathbf{x}$ but also the sets $\cup_{s\in[0,t]}\{\mathbf{x}^*_s : \mathbf{\Pi}^*_{x_0,s}(\mathbf{x}) \geq a\}$ must be bounded for $a \in (0,1]$ and $t \in \mathbb{R}_+$, where in analogy with the definition of $\mathbf{\Pi}^*_{x_0}$

$$\mathbf{\Pi}^*_{x_0,s}(\mathbf{x}) = \exp\Big(-\frac{1}{2}\int_0^s \big(\dot{\mathbf{x}}_r - b_r(\mathbf{x})\big)\cdot c_r(\mathbf{x})^{\oplus}\big(\dot{\mathbf{x}}_r - b_r(\mathbf{x})\big)\,dr\Big)$$

if $\mathbf{x}$ is absolutely continuous, $\mathbf{x}_0 = x_0$, and $\dot{\mathbf{x}}_r - b_r(\mathbf{x})$ is in the range of $c_r(\mathbf{x})$ a.e. on $[0,s]$, and $\mathbf{\Pi}^*_{x_0,s}(\mathbf{x}) = 0$ otherwise. The stated property follows, as in the proof of Lemma 8 in Puhalskii [37], by the fact that $\mathbf{\Pi}^*_{x_0}$ is a deviability and the equation $\dot{\mathbf{x}}_t = b_t(\mathbf{x})$ satisfies the extension condition so that $\cup_{s\in[0,t]}\{\mathbf{x}^*_s : \mathbf{\Pi}^*_{x_0,s}(\mathbf{x}) \geq a\} = \cup_{s\in[0,t]}\{\mathbf{x}^*_s : \mathbf{\Pi}^*_{x_0}(\mathbf{x}) \geq a\}$.

The second group of conditions concerns the predictable measure of jumps of $X^{\epsilon,\gamma}$, which has the form

$$\nu^{\epsilon,\gamma}([0,t],\Gamma) = \gamma^{-1}\int_0^t\int_U \mathbf{1}\big(\sqrt{\epsilon\gamma}f^{\epsilon,\gamma}_s(X^{\epsilon,\gamma},\vartheta) \in \Gamma\big)\,m(d\vartheta)\,ds,\ t \in \mathbb{R}_+,\ \Gamma \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}).$$

These are conditions $(A)_{loc} + (a)_{loc}$ on p.377 of Puhalskii [36]. They require, roughly speaking, that $\nu^{\epsilon,\gamma}$ in the limit reside on sets Γ that are "close to the origin" and follow by the moment conditions $(\tilde{P})$ or $(\tilde{SE})$ as in the proof of Theorem 5.4.4 in Puhalskii [36].

There is also condition (0) requiring that the initial value $X^{\epsilon,\gamma}_0$ converge superexponentially in probability to x_0, which holds since $X^{\epsilon,\gamma}_0 = x_0$. The final group of conditions concerns convergence of the predictable characteristics of $X^{\epsilon,\gamma}$ to the characteristics of X. Under the moment conditions $(\tilde{P})$ or $(\tilde{SE})$ the measure $\nu^{\epsilon,\gamma}$ satisfies condition $(L_2)_{loc}$ on p.410 of Puhalskii [36]. Therefore, as in Theorem 5.3.5 and in view of the assertion of Lemma 5.3.2 in Puhalskii [36], it suffices to check conditions $(\sup B')_{loc}$ and $(C'_0)_{loc}$ on p.410 of Puhalskii [36]. These conditions feature the first and modified second characteristics of $X^{\epsilon,\gamma}$ without truncation, which are given by the respective equalities

$$B'^{\epsilon,\gamma}_t = \int_0^t b^{\epsilon,\gamma}_s(X^{\epsilon,\gamma})\,ds \quad\text{and}\quad \tilde{C}'^{\epsilon,\gamma}_t = \epsilon\int_0^t c^{\epsilon,\gamma}_s(X^{\epsilon,\gamma})\,ds.$$

It is required that $B'^{\epsilon,\gamma}_{t\wedge\tau^{\epsilon,\gamma}_N} - B_{t\wedge\tau^{\epsilon,\gamma}_N}(X^{\epsilon,\gamma})$ converge to 0 superexponentially in probability uniformly over t from bounded domains and $\epsilon^{-1}\tilde{C}'^{\epsilon,\gamma}_{t\wedge\tau^{\epsilon,\gamma}_N} - C_{t\wedge\tau^{\epsilon,\gamma}_N}(X^{\epsilon,\gamma})$ converge to 0 superexponentially in probability for all t, where $\tau^{\epsilon,\gamma}_N = \inf\{t \in \mathbb{R}_+ : |X^{\epsilon,\gamma}_t| + t \geq N\}$ and $N \in \mathbb{N}$ is arbitrary. These convergences follow by the convergence conditions in the hypotheses.

Thus, according to Theorem 5.2.12 in Puhalskii [36] the sequence of laws of the $X^{\epsilon,\gamma}$ is exponentially tight and every limit deviability solves the maxingale problem (x_0, G) with G from (3.1). According to Theorem 3.1 such a deviability is a weak solution (3.2), so is specified uniquely. □

Theorem 4.1 implies, in particular, that, if, in addition, the equation $\mathbf{x}_t = x_0 + \int_0^t b_s(\mathbf{x})\,ds$ has a unique solution, then under the hypotheses the $X^{\epsilon,\gamma}$ tend in probability to this solution. We now consider deviations of order $o(1)$ from such a solution. The limit proves to be an idempotent Ornstein-Uhlenbeck process.

Theorem 4.2. *Let the differential equation* $\mathbf{x}_t = x_0 + \int_0^t b_s(\mathbf{x})\,ds$ *have a solution* $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_t, t \in \mathbb{R}_+)$. *Let* $\epsilon \to 0$, $\gamma \to 0$, *and* $\eta \to 0$ *in such a way that* $\eta/\epsilon \to \infty$ *and that, for arbitrary* $t \in \mathbb{R}_+$ *and* $A \in \mathbb{R}_+$,

$$\sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A} \int_0^t \Big|\sqrt{\frac{\eta}{\epsilon}}\big(b_s^{\epsilon,\gamma}(\tilde{\mathbf{x}} + \sqrt{\frac{\epsilon}{\eta}}\mathbf{x}) - b_s(\tilde{\mathbf{x}})\big) - \tilde{b}_s\mathbf{x}_s\Big|\,ds \to 0,$$

$$\sup_{\mathbf{x}\in\mathbb{D}:\,\mathbf{x}_t^*\le A} \int_0^t \|c_s^{\epsilon,\gamma}(\tilde{\mathbf{x}} + \sqrt{\frac{\epsilon}{\eta}}\mathbf{x}) - c_s(\tilde{\mathbf{x}})\|\,ds \to 0,$$

where $\tilde{b}_s$, $s \in \mathbb{R}_+$, *is an* $\mathbb{R}^{d\times d}$*-valued function such that the function* $\|\tilde{b}_s\|$, $s \in \mathbb{R}_+$, *is locally integrable. If, in addition, either condition* $(\tilde{P})$ *holds and* $\eta\log(\gamma^{-1}) \to \infty$, *or condition* $(\widetilde{SE})$ *holds and* $\eta^{2-\beta}/\gamma^\beta \to \infty$, *then* $\tilde{X}^{\epsilon,\gamma,\eta} \xrightarrow{ld} \tilde{X}$ *at rate* $1/\eta$, *where the stochastic processes* $\tilde{X}^{\epsilon,\gamma,\eta} = (\tilde{X}_t^{\epsilon,\gamma,\eta}, t \in \mathbb{R}_+)$ *are defined by* $\tilde{X}_t^{\epsilon,\gamma,\eta} = \sqrt{\eta/\epsilon}\big(X_t^{\epsilon,\gamma} - \tilde{\mathbf{x}}_t\big)$ *and the idempotent process* $\tilde{X} = (\tilde{X}_t,\, t \in \mathbb{R}_+)$ *is the unique Luzin strong solution of the idempotent Ito equation* $\dot{X}_t = \tilde{b}_t X_t + c_t(\tilde{\mathbf{x}})^{1/2}\,\dot{W}_t$, $X_0 = 0$.

Proof. The proof proceeds along the lines of the proof of Theorem 4.1. The moment conditions imply condition $(L_2)_{loc}$, so one has to verify conditions (0), $(A)_{loc} + (a)_{loc}$, $(\sup B')_{loc}$, $(C_0')_{loc}$, and (NE). They follow by the argument of the proof of Theorem 4.1 if one notes that the predictable characteristics of $\tilde{X}^{\epsilon,\gamma,\eta}$ without truncation are of the form

$$B'^{\epsilon,\gamma,\eta}_t = \int_0^t \sqrt{\frac{\eta}{\epsilon}}\Big(b_s^{\epsilon,\gamma}\Big(\tilde{\mathbf{x}} + \sqrt{\frac{\epsilon}{\eta}}\tilde{X}^{\epsilon,\gamma,\eta}\Big) - b_s(\tilde{\mathbf{x}})\Big)ds,\; C_t^{\epsilon,\gamma,\eta} = \eta\int_0^t c_s^{\epsilon,\gamma}\Big(\tilde{\mathbf{x}} + \sqrt{\frac{\epsilon}{\eta}}\tilde{X}^{\epsilon,\gamma,\eta}\Big)\,ds,$$

$$\nu^{\epsilon,\gamma,\eta}([0,t],\Gamma) = \gamma^{-1}\int_0^t\int_U \mathbf{1}\Big(\sqrt{\gamma\eta}\,f_s^{\epsilon,\gamma}\Big(\tilde{\mathbf{x}} + \sqrt{\frac{\epsilon}{\eta}}\tilde{X}^{\epsilon,\gamma,\eta},\vartheta\Big) \in \Gamma\Big)m(d\vartheta)\,ds,\; \Gamma \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}).$$

□

Remark 4.1. If $b^{\epsilon,\gamma}(\mathbf{x}) = \hat{b}_s(\mathbf{x}_s)$, where $\hat{b}_s(x)$ is differentiable in x for almost all s and the derivative $\hat{b}_s'(x)$ is bounded on bounded domains, then $\tilde{\mathbf{x}}$ is specified uniquely and the first convergence condition in the hypotheses holds for $\tilde{b}_s = \hat{b}_s'(\tilde{\mathbf{x}}_s)$.

We now give an application to a queueing system with time and state dependent service rates. It concerns a one-dimensional setting, so each time the above notation is used, it is assumed that $d = 1$. We consider a sequence of multiple-server queues with exponential service times and exponential customer reneging. At time t the nth queueing system consists of K_t^n homogeneous servers in parallel with service rates μ_t^n. Arriving customers who find no available servers form a queue and are served in the order of arrival. If a server serving a customer at time t becomes unavailable due to a change of K_t^n, then this customer returns to the queue. The customers in the queue at time t independently leave the system at rate θ_t^n. The functions μ_t^n and θ_t^n are assumed to be locally integrable and the function K_t^n is assumed to be Lebesgue measurable with values in $\mathbb{R}_+ \cup \{+\infty\}$, where $K_t^n = \infty$ corresponds to the case of an infinite-server queue.

The probability objects associated with the nth system are defined on a complete probability space $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$. Let $A^n = (A^n_t,\, t \in \mathbb{R}_+)$ denote the exogenous arrival process in the nth queue. Let $L^{n,k} = (L^{n,k}_t,\, t \in \mathbb{R}_+)$, $k \in \mathbb{N}$, and $B^{n,k} = (B^{n,k}_t,\, t \in \mathbb{R}_+)$, $k \in \mathbb{N}$, be Poisson processes of respective rate θ^n_t and μ^n_t at time t, which are responsible for reneging and service, respectively, in the sense made precise by equation (4.2) below. The processes A^n, $L^{n,k}$, and $B^{n,k}$ are considered as random elements of space $\mathbb{D}$ and are assumed to be mutually independent. They are also independent of the initial queue length.

We fix a real-valued sequence b_n such that $b_n \to \infty$ and $b_n/\sqrt{n} \to 0$ as $n \to \infty$. It is assumed that for some $\mathbb{R}_+$-valued locally integrable functions λ^n_t the processes $\big((A^n_t - \int_0^t \lambda^n_s\, ds)/(b_n\sqrt{n}),\, t \in \mathbb{R}_+\big)$ LD converge in distribution at rate b_n^2 as $n \to \infty$ to an idempotent diffusion $\big(\int_0^t \sigma_s \dot{W}_s\, ds,\, t \in \mathbb{R}_+\big)$, where the function σ_t is Lebesgue measurable, $\int_0^t \sigma_s^2\, ds < \infty$, and $W = (W_t,\, t \in \mathbb{R}_+)$ denotes a standard idempotent Wiener process. For example, the A^n can be time-non-homogeneous Poisson processes of rates λ^n_t satisfying equality (4.1) below, alternatively they can be time and space scaled renewal processes or superpositions of renewal processes, see Puhalskii and Whitt [39] for more detail. We further assume that the following expansions hold

$$
\begin{aligned}
\lambda^n_t &= n\lambda_{0,t} + \sqrt{n}b_n\lambda_{1,t} + \sqrt{n}\lambda_{2,t},\\
\mu^n_t &= \mu_{0,t} + \frac{b_n}{\sqrt{n}}\mu_{1,t} + \frac{1}{\sqrt{n}}\,\mu_{2,t},\\
K^n_t &= n\kappa_{0,t} + \sqrt{n}b_n\kappa_{1,t} + \sqrt{n}\kappa_{2,t},\\
\theta^n_t &= \theta_{0,t} + \frac{b_n}{\sqrt{n}}\theta_{1,t} + \frac{1}{\sqrt{n}}\,\theta_{2,t},
\end{aligned}
\tag{4.1}
$$

where the functions $\lambda_{0,t}$, $\mu_{0,t}$ and $\theta_{0,t}$ are $\mathbb{R}_+$-valued and locally integrable, the function $\kappa_{0,t}$ is $\mathbb{R}_+ \cup \{+\infty\}$-valued and Lebesgue measurable, the functions $\lambda_{1,t}$, $\lambda_{2,t}$, $\mu_{1,t}$, $\mu_{2,t}$, $\theta_{1,t}$, and $\theta_{2,t}$ are locally integrable, the functions $\kappa_{1,t}$ and $\kappa_{2,t}$ are Lebesgue measurable and are such that the function $(|\kappa_{1,t}| + |\kappa_{2,t}|)(\mu_{0,t} + \theta_{0,t})$ is locally integrable. We also let x^+ denote the positive part of $x \in \mathbb{R}$.

Denoting by Q^n_t the number of customers in the nth system at time t, we have that distributionally the process $Q^n = (Q^n_t,\, t \in \mathbb{R}_+)$ satisfies the equation

$$
Q^n_t = Q^n_0 + A^n_t - \sum_{k=1}^{\infty} \int_0^t \mathbf{1}(Q^n_{s-} \ge k + K^n_{s-})\, dL^{n,k}_s - \sum_{k=1}^{\infty} \int_0^t \mathbf{1}(Q^n_{s-} \wedge K^n_{s-} \ge k)\, dB^{n,k}_s. \tag{4.2}
$$

Let an absolutely continuous function $(q_t,\, t \in \mathbb{R}_+)$ satisfy the differential equation $\dot{q}_t = \lambda_{0,t} - \theta_{0,t}(q_t - \kappa_{0,t})^+ - \mu_{0,t}(q_t \wedge \kappa_{0,t})$ a.e. and stochastic processes $X^n = (X^n_t,\, t \in \mathbb{R}_+)$ be defined by $X^n_t = \big(\sqrt{n}/b_n\big)\big(Q^n_t/n - q_t\big)$. Let $X = (X_t,\, t \in \mathbb{R}_+)$ be the idempotent diffusion specified by the equation

$$
\begin{aligned}
\dot{X}_t = \lambda_{1,t} &- \theta_{1,t}(q_t - \kappa_{0,t})^+ - \theta_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})(X_t - \kappa_{1,t}) + \mathbf{1}(q_t = \kappa_{0,t})(X_t - \kappa_{1,t})^+\big)\\
&- \mu_{1,t}(q_t \wedge \kappa_{0,t}) - \mu_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})\kappa_{1,t} + \mathbf{1}(q_t = \kappa_{0,t})(X_t \wedge \kappa_{1,t}) + \mathbf{1}(q_t < \kappa_{0,t})X_t\big)\\
&+ \sqrt{\sigma_t^2 + \theta_{0,t}(q_t - \kappa_{0,t})^+ + \mu_{0,t}(q_t \wedge \kappa_{0,t})}\,\dot{W}_t, \quad X_0 = x_0 \in \mathbb{R}.
\end{aligned}
\tag{4.3}
$$

We note that by Corollary 3.2 the latter equation has a unique Luzin strong solution, which is also a unique weak solution and is a unique Luzin weak solution.

Theorem 4.3. *If* $X^n_0 \xrightarrow{\mathbf{P}_n^{1/b_n^2}} x_0$ *as* $n \to \infty$*, then* $X^n \xrightarrow{ld} X$ *at rate* b_n^2 *as* $n \to \infty$*.*

For the proof, we recall a lemma on conditional LD convergence, see Chaganty [4] and Puhalskii [34, Theorem 2.2] (a similar argument appears also in Puhalskii [37, Theorem 1]).

Lemma 4.1. *Let $\mathbf{\Pi}_{\upsilon'}$ be a collection of deviabilities on a metric space Υ indexed by elements υ' of a metric space Υ' and let μ be a deviability on Υ'. Let $\mathbf{P}^n_{\upsilon'}$, $\upsilon' \in \Upsilon'$, be probability transition kernels from Υ' to Υ and m^n be probabilities on Υ'. Let the idempotent probability $\mathbf{\Theta}$ on Υ defined by $\mathbf{\Theta}(\upsilon) = \sup_{\upsilon\in\Upsilon'} \mathbf{\Pi}_{\upsilon'}(\upsilon)\mu(\upsilon')$, $\upsilon \in \Upsilon$, be a deviability. Let probabilities $\mathbf{Q}^n$ on Υ be defined by $\mathbf{Q}^n(\Gamma) = \int_{\Upsilon'} \mathbf{Q}^n_{\upsilon'}(\Gamma)m^n(d\upsilon')$ for Borel subsets Γ of Υ. Let $k_n \to \infty$ as $n \to \infty$. If $m^n \stackrel{ld}{\to} \mu$ at rate k_n as $n \to \infty$, and $\mathbf{Q}^n_{\upsilon'_n} \stackrel{ld}{\to} \mathbf{\Pi}_{\upsilon'}$ at rate k_n for μ-almost all $\upsilon' \in \Upsilon'$ and all sequences $\upsilon'_n \to \upsilon'$, then $\mathbf{Q}^n \stackrel{ld}{\to} \mathbf{\Theta}$ at rate k_n.*

Proof of Theorem 4.3. Let us assume at first that the A^n are deterministic functions $a^n = (a^n_t,\, t \in \mathbb{R}_+)$ such that the functions $\hat{a}^n = (\hat{a}^n_t,\, t \in \mathbb{R}_+)$ defined by $\hat{a}^n_t = (a^n_t - \int_0^t \lambda^n_s\, ds)/(b_n\sqrt{n})$ converge as $n \to \infty$ in $\mathbb{D}$ to an absolutely continuous function $\hat{a} = (\hat{a}_t,\, t \in \mathbb{R}_+)$ with $\hat{a}_0 = 0$. Let σ-algebras $\mathcal{F}^n_t$ be defined as the σ-algebras generated by the random variables Q^n_0, $B^{n,k}_s$, $L^{n,k}_s$, $k = 1, 2, \ldots$, $s \le t$, completed with sets of $\mathbf{P}_n$-measure zero, and let $\mathbf{F}^n = (\mathcal{F}^n_t,\, t \in \mathbb{R}_+)$. By (4.2), the definitions of the X^n and q_t, and the fact that the processes $(B^{n,k}_t - \int_0^t \mu^n_s\, ds,\, t \in \mathbb{R}_+)$ and $(L^{n,k}_t - \int_0^t \theta^n_s\, ds,\, t \in \mathbb{R}_+)$ are locally square integrable martingales relative to $\mathbf{F}^n$ with predictable quadratic variation processes $(\int_0^t \mu^n_s\, ds,\, t \in \mathbb{R}_+)$ and $(\int_0^t \theta^n_s\, ds,\, t \in \mathbb{R}_+)$, respectively, we can write that

$$
\begin{aligned}
X^n_t = X^n_0 + \hat{a}^n_t + \frac{\sqrt{n}}{b_n}\int_0^t \Big(\frac{\lambda^n_s}{n} - \lambda_{0,s}\Big)\, ds - \int_0^t \Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s - \frac{K^n_s}{n}\Big)^+ \frac{\sqrt{n}}{b_n}(\theta^n_s - \theta_{0,s})\, ds \\
- \frac{\sqrt{n}}{b_n}\int_0^t \Big(\Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s - \frac{K^n_s}{n}\Big)^+ - (q_s - \kappa_{0,s})^+\Big)\theta_{0,s}\, ds \\
- \int_0^t \Big(\Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s\Big) \wedge \frac{K^n_s}{n}\Big)\frac{\sqrt{n}}{b_n}(\mu^n_s - \mu_{0,s})\, ds \\
- \frac{\sqrt{n}}{b_n}\int_0^t \Big(\Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s\Big) \wedge \frac{K^n_s}{n} - q_s \wedge \kappa_{0,s}\Big)\mu_{0,s}\, ds + M^n_t, \quad (4.4)
\end{aligned}
$$

where the $M^n = (M^n_t,\, t \in \mathbb{R}_+)$ are locally square integrable martingales relative to the $\mathbf{F}^n$ with predictable quadratic variation processes $(\langle M^n\rangle_t,\, t \in \mathbb{R}_+)$ given by

$$
\langle M^n\rangle_t = \frac{1}{b_n^2}\int_0^t \Big(\Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s - \frac{K^n_s}{n}\Big)^+ \theta^n_s + \Big(\Big(\frac{b_n}{\sqrt{n}}X^n_s + q_s\Big) \wedge \frac{K^n_s}{n}\Big)\mu^n_s\Big)\, ds\,. \quad (4.5)
$$

As in the proof of Theorem 4.2, we now verify convergence of the predictable characteristics, conditions on the jumps, and uniqueness of a solution to the limit maxingale problem. More specifically, since we are within a Markov setting with linearly growing coefficients, we verify the hypotheses of Theorem 5.4.4 in Puhalskii [36]. A similar argument has been used in the proof of Theorem 6.2.3 in Puhalskii [36].

Introducing the processes $Y^n = (Y^n_t,\, t \in \mathbb{R}_+)$ by $Y^n_t = X^n_t - \hat{a}^n_t$, we have that they are special semimartingales, whose first characteristics without truncation $B'^n = (B'^n_t,\, t \in \mathbb{R}_+)$ are given by the equality $B'^n_t = \int_0^t b^n_s(Y^n_s)\,ds$, where

$$b^n_s(u) = \frac{\sqrt{n}}{b_n}\Big(\frac{\lambda^n_s}{n} - \lambda_{0,s}\Big) - \Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s - \frac{K^n_s}{n}\Big)^+ \frac{\sqrt{n}}{b_n}(\theta^n_s - \theta_{0,s})$$
$$- \frac{\sqrt{n}}{b_n}\bigg(\Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s - \frac{K^n_s}{n}\Big)^+ - (q_s - \kappa_{0,s})^+\bigg)\theta_{0,s} - \bigg(\Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s\Big) \wedge \frac{K^n_s}{n}\bigg)\frac{\sqrt{n}}{b_n}(\mu^n_s - \mu_{0,s})$$
$$- \frac{\sqrt{n}}{b_n}\bigg(\Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s\Big) \wedge \frac{K^n_s}{n} - q_s \wedge \kappa_{0,s}\bigg)\mu_{0,s}.$$

The modified second characteristics without truncation $\tilde{C}'^n = (\tilde{C}'^n_t,\, t \in \mathbb{R}_+)$ coincide with the predictable quadratic-variation processes of the M^n and by (4.5) have the form $\tilde{C}'^n_t = b_n^{-2}\int_0^t \tilde{c}'^n_s(Y^n_s)\,ds$, where

$$\tilde{c}'^n_s(u) = \Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s - \frac{K^n_s}{n}\Big)^+\theta^n_s + \Big(\Big(\frac{b_n}{\sqrt{n}}(u - \hat{a}^n_s) + q_s\Big) \wedge \frac{K^n_s}{n}\Big)\mu^n_s.$$

Easy calculations show that for $t \in \mathbb{R}_+$ and $v \in \mathbb{R}_+$

$$\lim_{n\to\infty}\int_0^t \operatorname*{ess\,sup}_{|u|\le v} |b^n_s(u) - b_s(u)|\,ds = 0,$$
$$\lim_{n\to\infty}\int_0^t \operatorname*{ess\,sup}_{|u|\le v} |\tilde{c}'^n_s(u) - c_s|\,ds = 0,$$

where

$$b_s(u) = \lambda_{1,s} - \theta_{1,s}(q_s - \kappa_{0,s})^+ - \theta_{0,s}\big(\mathbf{1}(q_s > \kappa_{0,s})(u - \hat{a}_s - \kappa_{1,s}) + \mathbf{1}(q_s = \kappa_{0,s})(u - \hat{a}_s - \kappa_{1,s})^+\big)$$
$$-\mu_{1,s}(q_s \wedge \kappa_{0,s}) - \mu_{0,s}\big(\mathbf{1}(q_s > \kappa_{0,s})\kappa_{1,s} + \mathbf{1}(q_s = \kappa_{0,s})((u - \hat{a}_s) \wedge \kappa_{1,s}) + \mathbf{1}(q_s < \kappa_{0,s})(u - \hat{a}_s)\big),$$
$$c_s = \theta_{0,s}(q_s - \kappa_{0,s})^+ + \mu_{0,s}(q_s \wedge \kappa_{0,s}).$$

This checks the convergence conditions in the statement of Theorem 5.4.4 in Puhalskii [36]. Since the jumps of the Y^n are not greater than $1/(b_n\sqrt{n})$, the jump condition (SE) of Theorem 5.4.4 also holds. Next, by Corollary 3.2 the equation

$$\dot{Y}_t = \lambda_{1,t} - \theta_{1,t}(q_t - \kappa_{0,t})^+ - \theta_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})(Y_t - \hat{a}_t - \kappa_{1,t}) + \mathbf{1}(q_t = \kappa_{0,t})(Y_t - \hat{a}_t - \kappa_{1,t})^+\big)$$
$$- \mu_{1,t}(q_t \wedge \kappa_{0,t}) - \mu_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})\kappa_{1,t} + \mathbf{1}(q_t = \kappa_{0,t})((Y_t - \hat{a}_t) \wedge \kappa_{1,t}) + \mathbf{1}(q_t < \kappa_{0,t})(Y_t - \hat{a}_t)\big)$$
$$+ \sqrt{\theta_{0,t}(q_t - \kappa_{0,t})^+ + \mu_{0,t}(q_t \wedge \kappa_{0,t})}\,\dot{W}_t, \quad Y_0 = x_0,$$

has a unique Luzin strong solution, so the limit idempotent law is specified uniquely. Thus, by Theorem 5.4.4 in Puhalskii [36] the Y^n LD converge in distribution at rate b_n^2 to Y, so the X^n LD converge in distribution to the Luzin strong solution of the equation

$$\dot{\hat{X}}_t = \dot{\hat{a}}_t + \lambda_{1,t} - \theta_{1,t}(q_t - \kappa_{0,t})^+ - \theta_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})(\hat{X}_t - \kappa_{1,t}) + \mathbf{1}(q_t = \kappa_{0,t})(\hat{X}_t - \kappa_{1,t})^+\big)$$
$$- \mu_{1,t}(q_t \wedge \kappa_{0,t}) - \mu_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})\kappa_{1,t} + \mathbf{1}(q_t = \kappa_{0,t})(\hat{X}_t \wedge \kappa_{1,t}) + \mathbf{1}(q_t < \kappa_{0,t})\hat{X}_t\big)$$
$$+ \sqrt{\theta_{0,t}(q_t - \kappa_{0,t})^+ + \mu_{0,t}(q_t \wedge \kappa_{0,t})}\,\dot{W}_t, \quad \hat{X}_0 = x_0. \quad (4.6)$$

We now consider the general case. Let $\hat{A}^n_t = (A^n_t - \int_0^t \lambda^n_s\,ds)/(b_n\sqrt{n})$, $\hat{A}^n = (\hat{A}^n_t,\, t \in \mathbb{R}_+)$, $\mathbf{Q}^{X^n}$ denote the probability distribution of X^n, $\mathbf{Q}^{X^n}_a$ denote the regular conditional distribution of X^n

given $\hat{A}^n = a \in \mathbb{D}$, $\mathbf{Q}^{\hat{A}^n}$ denote the probability distribution of $\hat{A}^n$, $\mathbf{\Theta}$ denote the Luzin weak solution of (4.3), and let $\mathbf{\Theta}_{\hat{a}}$ denote the Luzin weak solution of (4.6) if $\hat{a}$ is absolutely continuous and an arbitrary deviability on $\mathbb{C}$ otherwise. Since $\hat{A}^n$ is independent of Q_0^n, $L^{n,k}$ and $B^{n,k}$, we have that $\mathbf{Q}_{\hat{a}^n}^{X^n}$ coincides with the law of the solution of (4.4) for $\mathbf{Q}^{\hat{A}^n}$-almost all $\hat{a}^n$. By the properties of the idempotent Wiener process, we can write (4.3) in the following form

$$\begin{aligned}\dot{X}_t = \sigma_t \dot{W}_t^{(1)} + \lambda_{1,t} - \theta_{1,t}(q_t - \kappa_{0,t})^+ - \theta_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})(X_t - \kappa_{1,t}) + \mathbf{1}(q_t = \kappa_{0,t})(X_t - \kappa_{1,t})^+\big)\\ - \mu_{1,t}(q_t \wedge \kappa_{0,t}) - \mu_{0,t}\big(\mathbf{1}(q_t > \kappa_{0,t})\kappa_{1,t} + \mathbf{1}(q_t = \kappa_{0,t})(X_t \wedge \kappa_{1,t}) + \mathbf{1}(q_t < \kappa_{0,t})X_t\big)\\ + \sqrt{\theta_{0,t}(q_t - \kappa_{0,t})^+ + \mu_{0,t}(q_t \wedge \kappa_{0,t})}\,\dot{W}_t, \quad X_0 = x_0,\end{aligned}$$

where $W^{(1)}$ is a standard idempotent Wiener process independent of W. Denoting as $\mathbf{\Theta}^A$ the idempotent distribution of $\big(\int_0^t \sigma_s \dot{W}_s^{(1)}\,ds,\ t \in \mathbb{R}_+\big)$, we have by (4.6) and weak uniqueness for (4.3) that $\mathbf{\Theta}(\Gamma) = \sup_{\hat{a} \in \mathbb{C}} \mathbf{\Theta}_{\hat{a}}(\Gamma)\, \mathbf{\Theta}^A(\hat{a})$ for $\Gamma \subset \mathbb{C}$. Also $\mathbf{Q}^{X^n}(\Gamma) = \int_{\mathbb{D}} \mathbf{Q}_{\hat{a}}^{X^n}(\Gamma)\, \mathbf{Q}^{\hat{A}^n}(d\hat{a})$ for Borel subsets Γ of $\mathbb{D}$. Since, as we have proved, the convergence $\hat{a}^n \to \hat{a}$, where $\hat{a}$ is absolutely continuous with $\hat{a}_0 = 0$, implies that $\mathbf{Q}_{\hat{a}^n}^{X^n} \overset{ld}{\to} \mathbf{\Theta}_{\hat{a}}$ at rate b_n^2, $\mathbf{\Theta}^A(\hat{a}) = 0$ if $\hat{a}$ is not absolutely continuous, and by hypotheses $\mathbf{Q}^{\hat{A}^n} \overset{ld}{\to} \mathbf{\Theta}^A$ at rate b_n^2, the proof is completed by an application of Lemma 4.1. □

Acknowledgement. The author is deeply grateful to the referee for numerous helpful comments and suggestions, which have contributed to improving the quality of presentation.

References

[1] R. Azencott. Grandes deviations et applications. In *Lecture Notes Math.*, volume 774, pages 1–176. Springer, 1980.

[2] P. Baldi and M. Chaleyat-Maurel. An extension of Ventcel-Freidlin estimates. In *Lecture Notes Math.*, volume 1316, pages 305–327. Springer, 1988.

[3] M. Boué, P. Dupuis, and R.S. Ellis. Large deviations for small noise diffusions with discontinuous statistics. *Probab. Theory Related Fields*, 116(1):125–149, 2000.

[4] N.R. Chaganty. Large deviations for joint distributions and statistical applications. *Sankhyā Ser. A*, 59(2):147–166, 1997.

[5] E.A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill, 1955.

[6] N.J. Cutland. An extension of the Ventcel-Freidlin large deviation principle. *Stochastics*, 24:121–149, 1988.

[7] D.A. Dawson and S. Feng. Large deviations for the Fleming-Viot process with neutral mutation and selection. *Stochastic Process. Appl.*, 77(2):207–232, 1998.

[8] D.A. Dawson and S. Feng. Large deviations for the Fleming-Viot process with neutral mutation and selection. II. *Stochastic Process. Appl.*, 92(1):131–162, 2001.

[9] A. de Acosta. A general non-convex large deviation result with applications to stochastic equations. *Probab. Theory Related Fields*, 118(4):483–521, 2000.

[10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, 2nd edition, 1998.

[11] P. Dupuis and R. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, 1997.

[12] S.N. Ethier and T.G. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, 1986.

[13] S. Fang and T. Zhang. A class of stochastic differential equations with non-Lipschitzian coefficients: pathwise uniqueness and no explosion. *C. R. Math. Acad. Sci. Paris*, 337(11):737–740, 2003.

[14] J. Feng. Martingale problems for large deviations of Markov processes. *Stoch. Proc. Appl.*, 81(2):165–216, 1999.

[15] J. Feng and T.G. Kurtz. Large deviations for stochastic processes. (preliminary manuscript), 2004.

[16] S. Feng. The behaviour near the boundary of some degenerate diffusions under random perturbation. In *Stochastic models (Ottawa, ON, 1998)*, volume 26 of *CMS Conf. Proc.*, pages 115–123. Amer. Math. Soc., Providence, RI, 2000.

[17] S. Feng and J. Xiong. Large deviations and quasi-potential of a Fleming-Viot process. *Electron. Comm. Probab.*, 7:13–25 (electronic), 2002.

[18] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems*. Nauka, 1979. In Russian, English translation: Springer, 1984.

[19] A. Friedman. *Stochastic Differential Equations and Applications*, volume 2. Academic Press, 1976.

[20] I.I. Gihman and A.V. Skorohod. *Stokhasticheskie differentsialnye uravneniya i ikh prilozheniya*. Naukova Dumka, 1982. (in Russian).

[21] P. Hartman. *Ordinary Differential Equations*. Wiley, 1964.

[22] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North Holland, 2nd edition, 1989.

[23] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 1987.

[24] A.P. Korostelëv and S.L. Leonov. An action functional for a diffusion process with discontinuous drift. *Teor. Veroyatnost. i Primenen.*, 37(3):570–576, 1992. translation: Theory Probab. Appl. 37 (1992), no. 3, 543–550.

[25] A.P. Korostelëv and S.L. Leonov. Action functional for diffusions in discontinuous media. *Probab. Theory Related Fields*, 94(3):317–333, 1993.

[26] R. Liptser. Large deviations for two scaled diffusions. *Probab. Theory Related Fields*, 106(1):71–104, 1996.

[27] R.Sh. Liptser and A. Pukhalskii. Limit theorems on large deviations for semimartingales. *Stoch. Stoch. Rep.*, 38:201–249, 1992.

[28] R.Sh. Liptser and A.N. Shiryayev. *Theory of Martingales*. Kluwer, 1989.

[29] T. Mikami. Some generalizations of Wentzell's lower estimates on large deviations. *Stochastics*, 24(4):269–284, 1988.

[30] K. Narita. On explosion and growth order of inhomogeneous diffusion processes. *Yokohama Math. J.*, 28(1-2):45–57, 1980.

[31] K. Narita. Large deviation principle for diffusion processes. *Tsukuba J. Math.*, 12(1):211–229, 1988.

[32] A. Puhalskii. Weak convergence theory approach to large deviations. In *Large deviations and applications*, Oberwolfach, 1992.

[33] A. Puhalskii. The method of stochastic exponentials for large deviations. *Stochastic Process. Appl.*, 54(1):45–70, 1994.

[34] A. Puhalskii. Large deviation analysis of the single server queue. *Queueing Systems Theory Appl.*, 21(1-2):5–66, 1995.

[35] A. Puhalskii. Large deviations of semimartingales: a maxingale problem approach. I. Limits as solutions to a maxingale problem. *Stoch. Stoch. Rep.*, 61:141–243, 1997.

[36] A. Puhalskii. *Large Deviations and Idempotent Probability*. Chapman & Hall/CRC, 2001.

[37] A. Puhalskii. On large deviation convergence of invariant measures. *J. Theoret. Probab.*, 16(3):689–724, 2003.

[38] A.A. Puhalskii. Stochastic processes in random graphs. (to appear), 2005.

[39] A.A. Puhalskii and W. Whitt. Functional large deviation principles for first-passage-time processes. *Ann. Appl. Probab.*, 7(2):362–381, 1997.

[40] D.W. Stroock. *An Introduction to the Theory of Large Deviations*. Springer, 1984.

[41] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Springer, 1979.

[42] S.R.S. Varadhan. *Large Deviations and Applications*. SIAM, 1984.

[43] A.D. Wentzell. *Limit Theorems on Large Deviations for Markov Stochastic Processes*. Nauka, 1986. (in Russian, English translation: Kluwer, 1990).