Kiefer-Wolfowitz Type Stochastic Approximation with Semimartingales

Von der Fakultät Mathematik und Physik
der Universität Stuttgart
zur Erlangung der Würde eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

von

Philipp Schnizler

geboren in Schwäbisch Hall

Hauptberichter:	Priv.-Doz. Dr. J. Dippon
Mitberichter:	Prof. Dr. I. Steinwart
Tag der Einreichung:	06. Oktober 2010
Tag der mündlichen Prüfung:	29. Oktober 2010

Institut für Stochastik und Anwendungen

2010

Danksagung

Die vorliegende Arbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Institut für Stochastik und Anwendungen der Universität Stuttgart und wurde im Sommersemester 2010 von der dortigen Fakultät der Mathematik und Physik als Dissertation angenommen.

Herrn PD Dr. Jürgen Dippon danke ich für die wissenschaftliche Betreuung der Arbeit sowie für die Übernahme des Erstgutachtens. Seine wertvollen Anregungen, seine Geduld und sein Vertrauen aber vor allem seine Hilfsbereitschaft bei jeglichen anfallenden Problemen waren mir eine große Hilfe. Danken möchte ich auch Herrn Professor Dr. Ingo Steinwart für die Erstellung des Zweitgutachtens und Herrn Professor Dr. Walk, der immer ein offenes Ohr hatte.

Ein besonderer Dank gebührt meiner Familie und meiner Freundin Noemi, der Konstanten in meinem Leben, für die unermüdliche Unterstützung und den unerschütterlichen Glauben an mich und meine Fähigkeiten.

06. Oktober 2010 Philipp Schnizler

Jeder Mensch kann irren.

Im Irrtum verharren wird jedoch nur der Tor.

Cicero

Deutsche Zusammenfassung

In der vorliegenden Arbeit wird ein verallgemeinerter zeitstetiger Kiefer-Wolfowitz-Prozess vorgeschlagen, der sich als Lösung der folgenden stochastischen Integralgleichung ergibt:

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s} \Big\{ f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) \Big\}_{i \in \{1,\dots,d\}} \mathrm{d}R_s - \int_0^t \frac{a_s}{2c_s} M(\mathrm{d}s, Z_{s-}).$$

Basierend auf den Arbeiten von Melnikov und Valkeila [22], Lazrieva et al. [18] und vielen weiteren Arbeiten, die sich mit dem zeitdiskreten Kiefer-Wolfowitz-Prozess beschäftigen, werden in dieser Arbeit nicht nur die asymptotischen Eigenschaften des Prozesses Z_t, sondern auch die des gemittelten Prozesses $\bar{Z}_t$ betrachtet.

Der Robbins-Monro- und der Kiefer-Wolfowitz-Prozess werden im ersten Kapitel sowohl als zeitdiskrete als auch als zeitstetige stochastische Approximationsverfahren vorgestellt. Diese Darstellung soll für ein besseres Verständnis des Gebiets und der in dieser Arbeit vorgenommenen Verallgemeinerung, bezogen auf die obige stochastische Integralgleichung, sorgen.

Das zweite Kapitel geht auf Fragen nach der Konsistenz und der Konvergenzgeschwindigkeit des Prozesses Z_t ein. Diese Fragen konnten ohne erhebliche Einschränkungen hinsichtlich der Gestalt von R_t, a_t und c_t geklärt werden, da nicht angenommen werden musste, dass diese deterministisch sind. Beim Nachweis beider Resultate spielt ein Lemma eine zentrale Rolle, das Aussagen über Konvergenzmengen von positiven speziellen Semimartingalen macht und im Wesentlichen auf deren multiplikativen Zerlegung beruht. Die Resultate der Konsistenz und der Konvergenzgeschwindigkeit werden im Itô-Fall und im zeitdiskreten Fall diskutiert. Hierbei ergeben sich drei bereits bekannte und ein neues Resultat.

Der nächste Teil der Arbeit widmet sich der asymptotischen Normalität, in dem jedoch nur deterministische Prozesse R_t und Dämpfungsprozesse der Gestalt $a_t := a(1+R_t)^{-1}$ zugelassen werden. Im Fall von zwei bzw. dreimal differenzierbaren Regressionsfunktionen f wird zunächst die Konvergenzgeschwindigkeit des Prozesses Z_t in einem fast L^2-Sinne betrachtet und anschließend die Frage nach dessen asymptotischer Normalität geklärt. Zum Nachweis der asymptotischen Normalität werden eine geeignete Darstellung des Prozesses Z_t, die fast L^2-Konvergenzgeschwindigkeit unter Verwendung der Markov-Ungleichung und ein zentraler Grenzwertsatz verwendet. Diese Resultate werden wiederum im Itô-Fall und im zeitdiskreten Fall diskutiert. Im zeitdiskreten Fall zeigt sich, dass das Resultat über die asymptotische Normalität mit bereits bekannten Resultaten übereinstimmt. Ein entsprechendes Resultat im Itô-Fall war in der Literatur bisher nicht zu finden.

Im vierten Kapitel wird, aufbauend auf den Arbeiten [3], [4] und [5] von Dippon und Renz im zeitdiskreten Fall, ein gemittelter Prozess $\bar{Z}_t$ unter Verwendung von schwächeren Dämpfungsprozessen der Gestalt $a_t := a(1 + R_t)^{-\alpha}$ mit $\frac{5}{6} < \alpha < 1$ präsentiert und sowohl dessen Konsistenz als auch dessen asymptotische Normalität diskutiert. Wie im vorangegangenen Kapitel wird auch hier davon ausgegangen, dass der Prozess R_t deterministisch ist. Anschließend werden wieder der Itô-Fall und der zeitdiskrete Fall diskutiert. Dabei ergibt sich im zeitdiskreten Fall ein bereits bekanntes und im Itô-Fall ein neues Resultat.

Abschließend beschäftigt sich die Arbeit mit der Frage, wie sich der Kiefer-Wolfowitz-Prozess bzw. der gemittelte Prozess asymptotisch verhält, wenn der Dämpfungsprozess zu einer Konstanten ausgeartet ist. Diese Frage war bei zeitstetigen Varianten des Robbins-Monro-Prozesses bisher noch nicht geklärt. Das Augenmerk richtet sich auf lineare Regressionsfunktionen; die Einschränkung, dass der Prozess R_t deterministisch ist, wird aufgehoben. Unter Verwendung eines einfachen Beispieles wird gezeigt, dass der Robbins-Monro-Prozess in diesem Fall nicht konsistent ist. Allerdings kann vom gemittelten Prozess gezeigt werden, dass dieser asymptotisch normal ist. Auf diesen Resultaten aufbauend wird die eigentliche Fragestellung nach dem asymptotischen Verhalten des Kiefer-Wolfowitz-Prozesses für den Fall einer quadratischen Regressionsfunktion geklärt. Anschließend werden der Itô-Fall, der zeitdiskrete Fall und der Fall $R_t := t + \lfloor t \rfloor$ diskutiert. Für den zeitdiskreten Fall ergibt sich erneut ein bekanntes und für den Itô-Fall ein neues Resultat. Von besonderem Interesse erweist sich die Betrachtung von $R_t := t + \lfloor t \rfloor$, da hier, im Gegensatz zum zeitdiskreten und zum Itô-Fall, der konstante Dämpfungsparameter a in die Kovarianzmatrix der asymptotischen Verteilung eingeht.

Contents

1 Introduction

From Recursions to Semimartingale Approaches in Stochastic Approximation

The pioneering work [35] by Robbins and Monro in 1951 can be seen as the origin of the field of stochastic approximation. Motivated by the problem of finding roots of a function $f : \mathbb{R} \to \mathbb{R}$, where the precise form of f is not known and the experimenter is only able to take "noisy" measurements, they suggested the recursion

$$Z_{n+1} - Z_n = -a_n \Big(f(Z_n) + M_n \Big) . \tag{1}$$

According to the recursion, the experimenter obtains the next point Z_{n+1}, if he takes the observation at the point Z_n that contains noise and is represented by $f(Z_n) + M_n$, damps it with the factor a_n and subtracts the result from the point Z_n. A typical weight sequence is $a_n := a\, n^{-1}$, and its damping influence is obvious. In the same work they proved that the process generated by the recursion converges in probability to the root of the function.

Inspired by this work, Kiefer and Wolfowitz [13] suggested in 1952 the recursion

$$Z_{n+1} - Z_n = -\frac{a_n}{2c_n} \Big(f(Z_n + c_n) - f(Z_n - c_n) + M_n \Big)$$

to find the stationary point instead of the root of the function f. The underlying conception of the recursion is the use of the Robbins-Monro method to find the root of the function $\nabla f(x)$. Since neither $\nabla f(x)$ nor noisy observations of it are available, the practical idea is to approximate $\nabla f(x)$ by $\frac{f(x+c_n)-f(x-c_n)}{2c_n}$. Here the weight sequence a_n takes the damping part as in the Robbins-Monro method and the weight sequence c_n is needed in the approximation of the gradient. Typical choices of the sequences a_n and c_n are $a_n := a\, n^{-1}$ and $c_n := c\, n^{-\gamma}$ with $\gamma \in (0, \frac{1}{2})$. As in the Robbins-Monro recursion, they showed that the process generated by the recursion converges in probability to the stationary point of the function f.

In 1954 Blum showed not only another concept of convergence [1], but he also suggested a generalization for both recursions. He considered the Robbins-Monro recursion

$$Z_{n+1} - Z_n = -a_n \Big(f(Z_n) + M_n \Big)$$

in the case of a function $f : \mathbb{R}^d \to \mathbb{R}^d$ and the Kiefer-Wolfowitz recursion

$$Z_{n+1} - Z_n = -\frac{a_n}{c_n} \Big\{ f(Z_n + c_n e_i) - f(Z_n - c_n e_i) + M_n^i \Big\}_{i \in \{1,\dots,d\}}$$

in the case of a function $f : \mathbb{R}^d \to \mathbb{R}$. In 1957 Sacks proved asymptotic normality for both processes in [36].

In 1967 Fabian found a weakness in the multidimensional version of the Kiefer-Wolfowitz recursion suggested by Blum. He suggested a recursion with a modified design in [9] and showed that the speed of convergence reaches the speed of the Robbins-Monro process, which is $\sqrt{n}$, if the considered function f is in any order differentiable. Furthermore the algorithm avoids the appearance of a bias, which is clearly visible in [3]. Since then, many different designs have been suggested and the resulting processes were discussed.

A truly amazing new idea was introduced by Polyak concerning the Robbins-Monro process in [31]. He used slowly decaying weights $a_n := an^{-\alpha}$ with $0 < \alpha < 1$ in the Robbins-Monro recursion and considered the averaged Robbins-Monro process instead of the Robbins-Monro process itself. Surprisingly the averaged process offers some key benefits—in terms of the asymptotic behavior and the stability—over the Robbins-Monro process itself. Later on, Dippon and Renz suggested to consider a weighted averaged process of the Kiefer-Wolfowitz process and showed that it has some advantages over the traditional Kiefer-Wolfowitz process in [3], [4], and [5]. We refer the reader to Chapter 4 or to the references above for details on the benefits.

Although Itô, with his groundbreaking work in the forties, opened the possibility to consider stochastic differential or integral equations, his theory was not in the focus of applied mathematics until the seventies. Perhaps the most famous application of the theory is the work of Black, Scholes and Merton 1973 in the context of finance. In the course of this development, in 1973, Nevel'son and Has'minskiĭ did not only study stochastic approximation processes generated by recursions [30], but also stochastic approximation processes that were generated by stochastic integral equations of Itô type. More precisely, they considered the solution of the d-dimensional stochastic integral equation

$$Z_t = Z_0 - \int_0^t a_s f(Z_s)\,\mathrm{d}s - \int_0^t a_s \sigma_s(Z_s)\,\mathrm{d}W_s \qquad (2)$$

as a continuous-time version of the Robbins-Monro process and the one-dimensional stochastic integral equation

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s}\Big(f(Z_s + c_s) - f(Z_s - c_s)\Big)\,\mathrm{d}s - \int_0^t \frac{a_s}{2c_s}\sigma_s(Z_s)\,\mathrm{d}W_s$$

as a continuous-time version of the Kiefer-Wolfowitz process. Here W_t represents a Brownian motion. In the same monograph they investigated consistency for both processes and, in the case of the Robbins-Monro process, speed of convergence and asymptotic normality. Here, in the proof of the asymptotic normality, they used among others the well-known and helpful fact that an integral with the Brownian motion as integrator is normally distributed. The issues of whether the Kiefer-Wolfowitz process is asymptotic normal or how fast it converges, were not considered.

Until about 1990, the continuous-time processes generated by Itô-type integral equations and the discrete-time processes generated by recursions were considered separately. Initially, further development of stochastic analysis and a generalization of stochastic integral equations of Itô type towards broader semimartingale integral equations made a unification of both concepts possible. Modern stochastic integration theory leads to a self-contained integration theory. The concept of semimartingales is essential in this theory. Roughly speaking, a semimartingale can be represented by a sum of a process of finite variation and a local martingale. This approach makes the embedding of both discrete- and continuous-time stochastic approximation procedures in one framework possible, since the class of semimartingales contains jump processes. This new perspective offers, on the one hand, the chance to consider both process types using one general theory, and the chance to gain and consider further processes on the other hand. In the eighties, Melnikov dealt in detail with strong solutions of stochastic integral equations with respect to semimartingales and recognized the possibility of the generalization. In [25] and [26], he suggested the solution of the stochastic integral equation

$$Z_t = Z_0 - \int_0^t a_s f(Z_{s-})\,\mathrm{d}R_s - \int_0^t a_s\,\mathrm{d}M_s \tag{3}$$

as a general continuous-time version of the Robbins-Monro process. Here R_t is a predictable, increasing càdlàg process and a_t is a weight process, both of which continuous in time. Typical choices are $R_t := t$, $R_t := \lfloor t \rfloor$ and $a_t := a(1+R_t)^{-\alpha}$ with $0 < \alpha \leq 1$. The choices $R_t := t$ and $M_t := W_t$ with a Brownian motion W_t lead to an Itô setting. Together with Rodkina [24], Melnikov also suggested the solution of the stochastic integral equation

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s}\big\{f(Z_s + c_s e_i) - f(Z_s - c_s e_i)\big\}_{i\in\{1,\dots,d\}}\,\mathrm{d}R_s - \int_0^t \frac{a_s}{2c_s}\sigma_s(Z_s)\,\mathrm{d}M_s$$

as a general continuous-time version of the Kiefer-Wolfowitz process. However, they considered only continuous processes R_t, M_t and consequently Z_t. This strong continuity restriction does not allow the embedding of discrete-time stochastic approximation procedures. Under the assumption of the existence of a strong solution on $[0,\infty)$ the consistency of the Kiefer-Wolfowitz process is treated in [24] and the consistency and the asymptotic normality of the Robbins-Monro process as well as of the averaged process is treated in [22]. The conditions used are very technical, strong and hard to verify.

Lazrieva, Sharia and Toronjadze suggested—in [16], [17], and summary [18]—the solution of the stochastic integral equation

$$Z_t = Z_0 - \int_0^t H_s(Z_{s-})\,\mathrm{d}R_s - \int_0^t M(\mathrm{d}s, Z_{s-}) \tag{4}$$

as a general continuous-time version of the Robbins-Monro process. The choices $H_s(Z_{s-}) := a_s f(Z_{s-})$ and $M(\mathrm{d}s, Z_{s-}) := a_s \,\mathrm{d}M_s$ show clearly the embedding of (3) in (4). They proved consistency of the process using a multiplicative decomposition theorem, hence eliminating the technical, strong, and unnatural conditions. Of course, they also assumed the existence of a strong solution of the equation. The interested reader can find some notes on the existence and uniqueness of such stochastic integral equations for instance in [6], [7], [10], [21], [27], [28], [29], [33], and [34]. Furthermore, they proved asymptotic normality of the Robbins-Monro process and its associated averaged process. In both models (3) and (4) asymptotic normality is considered only in the case of a deterministic process R_t.

Why Bother with Semimartingale Models

Stochastic approximation arises from the problem of finding roots and locating stationary points of a function. Quite frequently, either the function has a very complicated form, or its exact expression is not known explicitly. Hence we have to rely on measurements only, which are typically corrupted by noise. As we assume noise and not inherent errors, it seems quite natural to assume that in each observation the expectation of the noise is zero, and the "best prediction" of the noise in the next observation—knowing all observations and errors that appeared before—is also zero. Let us consider the rewritten recursion (1)

$$Z_n = Z_1 - \sum_{j=1}^{n-1} a_j f(Z_j) - \sum_{j=1}^{n-1} a_j M_j.$$

The assumptions concerning the expectation and the "best prediction" are equivalent to the fact that the sum $\sum_{j=1}^{n-1} a_j M_j$ is a martingale. A martingale can be interpreted as fair game, which means here that the error does not deliberately lead us on the wrong track. If we want to consider a continuous-time version of the recursion above, it should be clear that the sums turn into integrals. In a first step, we will get the stochastic integral equation (2) and in further steps, we will obtain (3) and (4). We have pointed out that the cumulated error should be a martingale or a local martingale, respectively; of course this feature should also hold in a continuous-time version. Given the fact that the semimartingale theory is the broadest closure integration theory that handles such integrals with the desired martingale feature, it seems natural to prefer the general theory and consider the stochastic approximation process Z_n as a semimartingale. From a mathematical point of view, a generalization is worthwhile, especially since the usage of a semimartingale model allows the embedding of the discrete-time setting and the Itô setting, which were treated separately before. Furthermore, it offers the opportunity to discuss more general settings, for example situations in which the experimenter is only able to take observations at random times.

In many applications, a continuous-time model (2) or, more generally (3) and (4) seems natural. In the context of finance it is common to model the price processes of the traded assets using positive semimartingales. Here it is plausible to use semimartingale approximation processes instead of stochastic approximation recursions. Further example where continuous-time models occur are signal models.

Considering continuous-time stochastic approximation processes can be justified from another practical point of view. These processes can be used to approximate discrete-time settings if the observations are taken rather frequently. Furthermore, a discussion of continuous-time models raises the understanding of problems with high-frequency sampling.

Summary of this Thesis

In this thesis we suggest a general continuous-time Kiefer-Wolfowitz process. The Kiefer-Wolfowitz process is a solution of the stochastic integral equation

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s} \Big\{ f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) \Big\}_{i \in \{1,\dots,d\}} \mathrm{d}R_s - \int_0^t \frac{a_s}{2c_s} M(\mathrm{d}s, Z_{s-}).$$

Discussion of the asymptotic behavior of the process Z_t and the averaged process $\bar{Z}_t$ is based on the work done by Lazrieva et al. in [18], Melnikov and Valkeila in [22] and many papers dealing with the Kiefer-Wolfowitz process discrete in time.

This thesis consists of four further chapters, each with three sections. The theorems, required lemmata and their proofs are presented in the first two sections of each chapter; in the last one, we discuss the theorems using special settings. The discussion of special settings illustrates the presented theorems and allows a comparison with known results.

In the second chapter, we dwell on the question of consistency and speed of convergence of the process Z_t. This question is answered without major restrictions on the form of R_t, a_t, and c_t. In particular, the processes may not even be deterministic. In the proof of both theorems a lemma about convergence sets of positive special semimartingales is crucial while on the other hand the lemma is based on a multiplicative decomposition of semimartingales. It follows a discussion of the theorems in the Itô setting and in the discrete-time setting, yielding three known and one new result.

The third chapter is devoted to asymptotic normality. Here we assume that the process R_t is deterministic and the weight process a_t is equal to $a(1+R_t)^{-1}$. First we establish an almost L^2-convergence rate and afterwards asymptotic normality in the case of a two or three times differentiable regression function f. To manage the proof of asymptotic normality, we present a handy representation of the process, which enables us to use the almost L^2-convergence rate applying Markov's inequality and a central limit theorem. Afterwards we discuss asymptotic normality in the Itô setting and in the discrete-time setting, reaching a well-known and a new result, respectively.

In the fourth chapter, we present an averaged process using slowly decaying weights $a_t := a(1 + R_t)^{-\alpha}$ with $\frac{5}{6} < \alpha < 1$, which is based on the work done in [3], [4], and [5] by Dippon and Renz. In the first section, we discuss consistency as well as asymptotic normality. Here, as in the section before, we assume that the process R_t is deterministic. In the last section, we consider the discrete-time and the Itô setting, obtaining a known and a new result, respectively.

The last chapter investigates how the Kiefer-Wolfowitz process and the averaged process behave asymptotically, if the weight process a_t is degenerated to a constant a. Since this question is not answered in the case of the continuous-time Robbins-Monro process, this is our starting point. Here the restriction that the process R_t is deterministic is abrogated, but the discussion contains only linear regression functions. The contemplation of a simple example shows that the Robbins-Monro process is not consistent in general. However we prove asymptotic normality of the averaged process. Then we answer the main question of the asymptotic behavior of the averaged Kiefer-Wolfowitz process with constant weights in the case of a quadratic regression function. In the last section, we consider the results in the Itô setting, in the discrete-time setting, and in the case of a process $R_t := t + \lfloor t \rfloor$. As in the chapters before, we obtain in the discrete-time setting a known result and in the Itô setting a new one. It turns out that the process $R_t := t + \lfloor t \rfloor$ is of particular interest since here—in contrast to the discrete-time and the Itô setting—the constant a enters into the covariance matrix of the asymptotic distribution.

2 Asymptotic Properties of the Process Using General Weights

We suggest a solution of a stochastic integral equation as a general Kiefer-Wolfowitz process continuous in time and discuss its asymptotic behavior. To consider a stochastic integral equation, it is essential to talk about the stochastic basis and the objects or processes and how they are connected. Let us fix a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ satisfying the usual conditions, that is, $\mathcal{F}_0$ contains all $\mathbf{P}$-null sets of $\mathcal{F}$ and the filtration $\mathbb{F}$ is right-continuous. On this basis the objects $(R_t)_{t\geq 0}$, $(a_t)_{t\geq 0}$, $(c_t)_{t\geq 0}$, Z_0, and $(M(t,u))_{t\geq 0}$ are given. Here Z_0 is a $\mathcal{F}_0$-measurable random variable. The process $(R_t)_{t\geq 0}$ is increasing, càdlàg, adapted, predictable with respect to the filtration $\mathbb{F}$, and further $R_0 = 0$ and $\Delta R_0 = 0$ hold. The processes $(a_t)_{t\geq 0}$ and $(c_t)_{t\geq 0}$ are adapted and predictable with respect to the filtration $\mathbb{F}$ and furthermore the process $(\frac{a_t}{c_t})_{t\geq 0}$ is locally bounded. The random field $M(t,u)$ is $\mathbb{F}$-adapted, $(M(t,u))_{t\geq 0} \in \mathcal{M}^2_{\text{loc}}(\mathbf{P})$ holds for every $u \in \mathbb{R}^d$ and furthermore $(\int_0^t \frac{a_s}{c_s} M(\mathrm{d}s, Z_{s-}))_{t\geq 0} \in \mathcal{M}^2_{\text{loc}}(\mathbf{P})$. Here $\mathcal{M}^2_{\text{loc}}(\mathbf{P})$ denotes the class of $\mathbb{F}$-adapted, càdlàg processes for which a localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ exists such that the stopped process $(X_t^{\tau_n})_{t\geq 0}$ is, for every $n \in \mathbb{N}$, a square-integrable martingale with respect to the probability measure $\mathbf{P}$ and the filtration $\mathbb{F}$.

As the setting is given, we are ready to set up the stochastic integral equation

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s} \left\{ f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) \right\}_{i\in\{1,\dots,d\}} \mathrm{d}R_s - \int_0^t \frac{a_s}{2c_s} M(\mathrm{d}s, Z_{s-}). \tag{5}$$

In this thesis, we focus on the asymptotic behavior of the process Z_t. To this end we assume the existence of a unique strong solution on $[0,\infty)$. The existence and uniqueness of such stochastic integral equations is well-investigated for instance in [6], [7], [10], [21], [27], [28], [29], [33], and [34]. We propose this unique strong solution as general Kiefer-Wolfowitz process continuous in time.

Considering the stochastic integral equation, a further question arises. What is the meaning of each process which appears? Z_0 represents the starting point and is a random variable or a fixed point. Often the statistician who searches for the stationary point chooses Z_0. As already mentioned in the introduction, the idea of the algorithm as well as of the integral equation is the simultaneous approximation of $x^\star$, where $\nabla f(x^\star) = 0$, using a process of Robbins-Monro type and of $\nabla f(x)$ by $\frac{1}{2c_s}\{f(x + c_s e_i) - f(x - c_s e_i)\}_{i\in\{1,\dots,d\}}$. The process a_t is a weight process that has a damping effect and is needed in the Robbins-Monro part of the approximation. The process c_t is needed in the approximation of $\nabla f(x)$ by $\frac{1}{2c_s}\{f(x + c_s e_i) - f(x - c_s e_i)\}_{i\in\{1,\dots,d\}}$. Considering the meaning of the processes a_s and c_s, we will see later on that the statistician has to choose them among others positive and decreasing to find the stationary point. Chapter 5 gives an exception of the choice of decreasing weights. Typical choices for a_t and c_t

are $a_t := a\,(1+R_t)^{-1}$ and $c_t := c\,(1+R_t)^{-\gamma}$. Furthermore the process $(\frac{a_t}{c_t})_{t\geq 0}$ is assumed to be predictable and locally bounded to ensure that the integral $\int_0^t \frac{a_s}{c_s} M(\mathrm{d}s, Z_{s-})$ exists and is well defined in the context of semimartingales. The process R_t itself describes the observation rate. Commonly the process $R_t := \lfloor t \rfloor$ or $R_t := t$ is chosen. Thus in the case $R_t := \lfloor t \rfloor$ new observations arrive only at times $t \in \mathbb{N}$ and thus only at these times the data are updated and influence the process Z_t. In the case $R_t := t$ the observations are continuous in time and hence the updating is continuous in time. But there are much more interesting cases, for example stochastic processes R_t. The integral $\int_0^t \frac{a_s}{2c_s}\, M(\mathrm{d}s, Z_{s-})$ represents the noise in the data in which the damping factor is already included. If we assume that the observation noise is white noise, we have the integral $\int_0^t \frac{a_s}{2c_s}\sigma_s(Z_{s-})\,\mathrm{d}W_s$ with a matrix-valued process $\sigma_s(\cdot)$ and the d-dimensional Brownian motion W_s. This is one example of the quite general integral $\int_0^t \frac{a_s}{2c_s}\, M(\mathrm{d}s, Z_{s-})$.

In the following, we use $\mathcal{C}$ as a constant that may vary from inequality to inequality. If $\mathcal{C}$ depends on ω, we denote it by $\mathcal{C}_\omega$. Furthermore we use the well-known Landau symbols o, $\mathcal{O}$ to describe the asymptotic behavior of processes and the symbols o_b, $\mathcal{O}_\mathrm{b}$ to note that they are bounded in addition.

Given that the sense in which most statements should be interpreted is obvious, we accentuate it only in the statements of results or if it is not easy to distinguish. Most of the statements involving random variables or stochastic processes should be interpreted as almost sure if there is nothing mentioned.

2.1 Consistency

Here we address the issue of consistency of the process Z_t or more precisely strong consistency. The question of whether an estimator is consistent is quite natural in the background of estimation theory and it is usually one of the first one should ask. A discussion of consistency is a good starting point because further research is unnecessary if the estimator is not consistent and, if the estimator is consistent, such a result is often needed or helpful in doing further research, for example, when investigating speed of convergence. We consider the following set of conditions:

(A) $f : \mathbb{R}^d \to \mathbb{R}$ has a Lipschitz-continuous gradient.

(B) There exists an $x^\star$ with $\nabla f(x^\star) = 0$.

(C) The gradient satisfies

$$\underset{\epsilon>0}{\forall} \quad \underset{C(\epsilon)>0}{\exists} \quad \underset{\{x\in\mathbb{R}^d \mid \epsilon\leq\|x-x^\star\|\leq 1/\epsilon\}}{\forall} \qquad \langle \nabla f(x), x - x^\star \rangle \geq C(\epsilon).$$

(D) The processes a_s and c_s, which the statistician has to choose, are left-continuous and satisfy

$$a_s, c_s > 0 \qquad\qquad a_s, c_s \downarrow 0$$
$$\int_0^\infty a_s \,\mathrm{d}R_s = \infty \qquad\qquad \int_0^\infty a_s c_s \,\mathrm{d}R_s < \infty.$$

(E) For every $i \in \{1, .., d\}$ and $x \in \mathbb{R}^d$, we have

$$\int_0^\infty \frac{{a_s}^2}{{c_s}^2} \frac{h_s^{ii}(Z_{s-})}{1 + \|Z_{s-}\|^2} \,\mathrm{d}R_s < \infty, \quad \text{where} \quad h_s^{ii}(x) := \frac{\mathrm{d}\left\lceil \int_0^\cdot M_i(\mathrm{d}t, x)\right\rceil_s}{\mathrm{d}R_s}.$$

(F) If the process R_s is not continuous, then the following condition should also hold:

$$\int_0^\infty {a_s}^2 \Delta R_s \,\mathrm{d}R_s^d < \infty.$$

Now we explain the intuitive meaning of the conditions. To get a first impression, it is helpful to reduce to the 1-dimensional case. Considering the conditions (A) and (B), there is nothing to mention. Let $x^\star = 0$, it follows $f'(x)x > 0$ from condition (C), assuring that the gradient shows us the right direction to go. To recognize the need for the first part of condition (D) we assume $\int_0^\infty a_s \,\mathrm{d}R_s \leq I < \infty$. Now we regard the integral equation in the case of a function f with bounded first derivative (e.g. $f(x) = 1 - \exp(-0.5x^2)$) and the absence of "noise". We use the mean-value theorem and get

$$\begin{aligned} |Z_t - Z_0| &= |\int_0^t \frac{a_s}{2c_s} \left(f(Z_{s-} + c_s) - f(Z_{s-} - c_s)\right) \mathrm{d}R_s| \\ &= \int_0^t a_s |f'(\xi_s)| \,\mathrm{d}R_s \leq \mathcal{C} \int_0^t a_s \,\mathrm{d}R_s \leq \mathcal{C}. \end{aligned}$$

If the statistician chooses a bad starting point Z_0 that is further away from $x^\star$ than $\mathcal{C}$, the process Z_t cannot reach $x^\star$. Since often the statistician has no or only vague information about $x^\star$, the choice of Z_0 should not be crucial. The second condition in (D), namely $\int_0^\infty a_s c_s \,\mathrm{d}R_s < \infty$, and the condition (E), which is often implied by $\int_0^\infty \frac{{a_s}^2}{{c_s}^2} \,\mathrm{d}R_s < \infty$, affects the speed of the processes a_t, c_t and their interplay. If we interpret the Kiefer-Wolfowitz method as a simultaneous approximation of the gradient and of the stationary point using a Robbins-Monro-type method, these conditions ensure a balance. Furthermore, the condition (E) connects R_t and $M_i(\mathrm{d}t, x)$. This is quite natural since the "noise" is connected with observations, and the increasing rate of the process R_t specifies the observation rate. Considering a time interval on which R_t is constant, this connection is clearly apparent. Since the connection is given by the predictable quadratic variation, it is important to note here that it is almost surely unique. Another important feature of the condition (E) is the fact that it asymptotically damps the effect of the "noise". The condition (F) ensures that the damped jumps $a_s \Delta R_s$ converge to zero.

In the following, we assume the existence of a strong solution of the stochastic integral equation (5) on $\mathbb{R}_+ \times \Omega$.

Theorem 2.1. *Assume that the conditions (A)–(F) are fulfilled. Then the solution Z_t of the stochastic integral equation* (5) *converges almost surely towards the stationary point of the function f.*

Remark 2.1. Of course, convergence of the solution of (5), as claimed in Theorem 2.1, implies consistency of the solution.

Remark 2.2. Assumption (E) holds, if, e.g., $h_s^{ii}(x) \leq K_s^i(1 + \|x\|^2)$ with $\int_0^\infty \frac{a_s^2}{c_s^2} K_s^i \, \mathrm{d}R_s < \infty$ or especially $h_s^{ii}(x) \leq \mathcal{C}$ and $\int_0^\infty \frac{a_s^2}{c_s^2} \, \mathrm{d}R_s < \infty$ are satisfied.

The following lemma, from Lazrieva et al. (see [16], Corollary 2.3), plays an important role in the proof of Theorem 2.1. We recall this lemma and give a detailed proof.

Lemma 2.1. *For an arbitrary non-negative special semimartingale $X_t = X_0 + A_t + M_t$ satisfying $A_t \in \mathcal{V} \cap \mathcal{P}$, $M_t \in \mathcal{M}_{\mathrm{loc}}$,*

$$A_t \leq A_t^1 - A_t^2, \qquad A_t^1, A_t^2 \in \mathcal{V}^+ \cap \mathcal{P}, \qquad \textit{and} \qquad A_t^1 - A_t \in \mathcal{V}^+,$$

we have

$$\left\{ \int_0^\infty \frac{1}{1 + X_{s-}} \, \mathrm{d}A_s^1 < \infty \right\} \subseteq \{X_s \to\} \cap \left\{ A_\infty^2 < \infty \right\}.$$

Proof of Lemma 2.1

To begin with, the semimartingale X_t can be written in the form

$$X_t = X_0 + A_t + M_t = X_0 + A_t^1 - (A_t^1 - A_t) + M_t = X_0 + A_t^1 - \tilde{A}_t^2 + M_t.$$

In view of the assumptions, it is easy to discover that A_t^1 and $\tilde{A}_t^2 := A_t^1 - A_t$ are elements of $\mathcal{V}^+$ and that the following implications

$$A_t \leq A_t^1 - A_t^2 \qquad \Rightarrow \qquad A_t - A_t^1 \leq -A_t^2 \qquad \Rightarrow \qquad A_t^2 \leq A_t^1 - A_t$$

hold. Hence $\tilde{A}_t^2$ is non-negative and further the inequality

$$A_\infty^2 \leq A_\infty^1 - A_\infty = \tilde{A}_\infty^2 \tag{6}$$

is evident.

The rest of the proof is organized as follows: First we use the fact that a semimartingale Z_t with $Z_t \geq 0$ and $Z_t = Z_0 + B_t^1 - B_t^2 + \tilde{M}_t$, $B_t^1, B_t^2 \in \mathcal{V}^+$, $\tilde{M}_t \in \mathcal{M}_{\text{loc}}$, satisfies

$$\left\{\int_0^\infty \frac{1}{1 + Z_{s-} + B_{s-}^2}\,\mathrm{d}B_s^1 < \infty\right\} \subseteq \{Z_s \rightarrow\} \cap \left\{B_\infty^2 < \infty\right\} \tag{7}$$

to prove Lemma 2.1. Afterwards we will verify statement (7).

Applying (7) to X_s and using (6), A_t^1, and $\tilde{A}_t^2$, we obtain the statement of Lemma 2.1:

$$\begin{aligned}\left\{\int_0^\infty \frac{1}{1 + X_{s-}}\,\mathrm{d}A_s^1 < \infty\right\} &\subseteq \left\{\int_0^\infty \frac{1}{1 + X_{s-} + \tilde{A}_{s-}^2}\,\mathrm{d}A_s^1 < \infty\right\} \\ &\overset{(7)}{\subseteq} \{X_s \rightarrow\} \cap \{\tilde{A}_\infty^2 < \infty\} \overset{(6)}{\subseteq} \{X_s \rightarrow\} \cap \{A_\infty^2 < \infty\}.\end{aligned}$$

To complete the proof of the lemma, it remains to show (7). To this end we require the following theorem on a multiplicative decomposition of positive semimartingales (see [19], §5, Theorem 1, pp. 127):

If L_s is a special semimartingale with an additive decomposition

$$L_t = L_0 + C_t + N_t$$

where $C_t \in \mathcal{V} \cap \mathcal{P}$ and $N_t \in \mathcal{M}_{\text{loc}}$, and

$$\inf_{s \leq t} L_s > 0 \qquad \text{and} \qquad \inf_{0 < s \leq t} \frac{1}{L_{s-}} \Delta C_s > -1,$$

then L_s can be represented in a multiplicative way as

$$L_s = L_0 \mathcal{E}_s(\hat{C}) \mathcal{E}_s(\hat{N}), \tag{8}$$

where the processes $\hat{C}_s$ and $\hat{N}_s$ are defined as

$$\hat{C}_t := \int_0^t \frac{1}{L_{s-}}\,\mathrm{d}C_s \qquad \text{and} \qquad \hat{N}_t := \int_0^t \frac{1}{L_{s-} + \Delta C_s}\,\mathrm{d}N_s$$

and $\mathcal{E}_s(\hat{B})$, $\mathcal{E}_s(\hat{N})$ are the stochastic exponentials thereof.

To use this result effectively, we look at the process Y_t, given by

$$Y_t := Y_0 + B_t^1 + \tilde{M}_t \qquad \text{or} \qquad Y_t := 1 + Z_t + B_t^2$$

with $Y_0 := 1 + Z_0$. To apply the above theorem to Y_t, we verify its assumptions. The cumulative decomposition follows directly from the definition of Y_t. Thus, it suffices to check the two inequalities.

Owing to

$$Y_t = 1 + Z_t + B_t^2 \geq 1,$$

the first inequality

$$\inf_{s \leq t} Y_s > 0$$

is evident.

Now we prove the second inequality. Since B_t^1 is increasing, the jump process ΔB_t^1 naturally fulfills $\Delta B_t^1 \geq 0$. As mentioned above, Y_t is strongly positive, so

$$\frac{1}{Y_{s-}} \Delta B_s^1 \geq 0 \qquad \Rightarrow \qquad \inf_{0<s\leq t} \left(\frac{1}{Y_{s-}} \Delta B_s^1 \right) > -1$$

holds. But that is the proof of the second inequality, so with (8) we get

$$Y_t = Y_0 \mathcal{E}_t(\hat{A}) \mathcal{E}_t(\hat{M}) \tag{9}$$

where

$$\hat{A}_t := \int_0^t \frac{1}{Y_{s-}} \mathrm{d}B_s^1 \qquad \text{and} \qquad \hat{M}_t := \int_0^t \frac{1}{Y_{s-} + \Delta B_s^1} \mathrm{d}\tilde{M}_s.$$

Now we show that $\mathcal{E}_t(\hat{A})$ and $\mathcal{E}_t(\hat{M})$ almost surely converge. Then with these statements and (9), we get the almost surely convergence of Y_t.
To prove almost sure convergence of $\mathcal{E}_t(\hat{M})$, we apply a well-known convergence theorem on positive supermartingales. First we verify non-negativity of $\mathcal{E}_t(\hat{M})$. According to a formula concerning stochastic exponentials, $\mathcal{E}_t(\hat{M})$ can be written as

$$\mathcal{E}_t(\hat{M}) = \exp\left(\hat{M}_t - \frac{1}{2}[\hat{M}, \hat{M}]_t\right) \prod_{0<s\leq t} \left(1 + \Delta\hat{M}_s\right) \exp\left(-\Delta\hat{M}_s + \frac{1}{2}(\Delta\hat{M}_s)^2\right).$$

To show positivity of the right side of the above equation, it suffices to show $\Delta\hat{M}_s > -1$, which follows from

$$\begin{aligned} \Delta\hat{M}_s &= \Delta\left(\int_0^s \frac{1}{Y_{r-} + \Delta B_r^1} \mathrm{d}\tilde{M}_r\right) = \frac{1}{Y_{s-} + \Delta B_s^1}(\Delta\tilde{M}_s) \\ &= \frac{Y_s}{Y_{s-} + \Delta B_s^1} - \frac{Y_{s-} + \Delta B_s^1}{Y_{s-} + \Delta B_s^1} = \frac{Y_s}{Y_{s-} + \Delta B_s^1} - 1 > -1 \end{aligned}$$

using $\Delta\tilde{M}_s = \Delta(Y_s - B_s^1) = \Delta Y_s - \Delta B_s^1 = Y_s - (Y_{s-} + \Delta B_s^1)$. Thus $\mathcal{E}_t(\hat{M}) > 0$ holds.

Recall that

$$M_t \text{ is a non-negative local martingale} \quad \Rightarrow \quad M_t \text{ is a supermartingale} \tag{10}$$

Since $\mathcal{E}_t(\hat{M})$ is given as the solution of a stochastic integral equation driven by $\tilde{M}_t$, it is a local martingale, and, as shown above, it is positive. Furthermore, by (10), $\mathcal{E}_t(\hat{M})$ is a non-negative supermartingale. This proves almost sure convergence of the stochastic exponential of $\hat{M}_t$ towards a real-valued random variable due to the supermartingale convergence theorem, in short

$$\left\{\omega \in \Omega \mid \mathcal{E}_t(\hat{M})(\omega) \to\right\} = \Omega.$$

Coming back to the process $\mathcal{E}_t(\hat{A})$, we know that $\hat{A}_t$ is monotonically increasing. According to a formula concerning stochastic exponentials, $\mathcal{E}_t(\hat{A})$ can be written as

$$\mathcal{E}_t(\hat{A}) = \exp\left(\hat{A}_t\right) \prod_{0<s\leq t} \left(1 + \Delta\hat{A}_s\right) \exp\left(-\Delta\hat{A}_s\right).$$

Since $\hat{A}_t$ is monotonously increasing, the inequalities

$$0 \leq \left(1 + \Delta\hat{A}_s\right) \exp\left(-\Delta\hat{A}_s\right) \leq 1$$

hold and we get

$$\left\{\hat{A}_\infty < \infty\right\} = \left\{\hat{A}_t \to\right\} = \left\{\mathcal{E}_t(\hat{A}) \to\right\}.$$

Combining the above results, we obtain

$$\begin{aligned}\left\{\hat{A}_\infty < \infty\right\} &= \left\{\hat{A}_\infty < \infty\right\} \cap \left\{\mathcal{E}_t(\hat{M}) \to\right\} = \left\{\hat{A}_t \to\right\} \cap \left\{\mathcal{E}_t(\hat{M}) \to\right\} \\ &= \left\{\mathcal{E}_t(\hat{A}) \to\right\} \cap \left\{\mathcal{E}_t(\hat{M}) \to\right\} \subseteq \left\{Y_t \to\right\}.\end{aligned}$$

We know that B_t^2 is increasing and that Z_t and B_t^2 fulfill $Z_t, B_t^2 \geq 0$. Hence, with $Y_t = 1 + Z_t + B_t^2$, we get

$$\{Y_t \to\} \subseteq \{Z_t \to\} \cap \left\{B_\infty^2 < \infty\right\}.$$

This shows (7) and completes the proof of Lemma 2.1. □

Proof of Theorem 2.1

(a) Proof of almost sure convergence of Z_t. We assume without loss of generality that the stationary point of the function f is $x^\star = 0$. Using Itô's formula we get

$$\mathrm{d}\left\langle Z_s, Z_s\right\rangle = \sum_{i=1}^{d} \mathrm{d}\left(Z_s^i\right)^2 = \sum_{i=1}^{d} \left(2Z_{s-}^i \,\mathrm{d}Z_s^i + \mathrm{d}[Z^i]_s\right). \tag{11}$$

Now we are studying the terms appearing in (11). Using the stochastic integral equation (5) to represent $\mathrm{d}Z_s^i$, we get

$$Z_{s-}^i \,\mathrm{d}Z_s^i = -a_s Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i \,\mathrm{d}R_s - \frac{a_s}{2c_s} Z_{s-}^i \, M_i(\mathrm{d}s, Z_{s-}) \tag{12}$$

for the first summand and

$$\mathrm{d}[Z^i]_s = \frac{{a_s}^2}{4{c_s}^2}\Big(\{\cdot\}_i^2 \Delta R_s \,\mathrm{d}R_s^d + 2\{\cdot\}_i \Delta R_s \, M_i^d(\mathrm{d}s, Z_{s-}) + \,\mathrm{d}[\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})]_s\Big) \tag{13}$$

for the second one. Hence, by (12) and (13) the differential in (11) can be written as

$$\begin{aligned}\mathrm{d}\,\langle Z_s, Z_s\rangle = {} & -2\sum_{i=1}^d a_s Z_{s-}^i \frac{\{\cdot\}_i}{2c_s}\,\mathrm{d}R_s - 2\sum_{i=1}^d \frac{a_s}{2c_s} Z_{s-}^i \, M_i(\mathrm{d}s, Z_{s-}) + \sum_{i=1}^d \frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i^2 \Delta R_s \,\mathrm{d}R_s^d \\ & + 2\sum_{i=1}^d \frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i \Delta R_s \, M_i^d(\mathrm{d}s, Z_{s-}) + \sum_{i=1}^d \frac{{a_s}^2}{4{c_s}^2}\,\mathrm{d}[\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})]_s.\end{aligned}$$

With the two processes A_t and $\tilde{M}_t$, given by

$$\begin{aligned}A_t := {} & -2\sum_{i=1}^d \int_0^t a_s Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i \,\mathrm{d}R_s + \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i^2 \Delta R_s \,\mathrm{d}R_s^d \\ & + \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2}\,\mathrm{d}\lceil\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})\rceil_s \\ \tilde{M}_t := {} & -2\sum_{i=1}^d \int_0^t \frac{a_s}{2c_s} Z_{s-}^i \, M_i(\mathrm{d}s, Z_{s-}) + 2\sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i \Delta R_s \, M_i^d(\mathrm{d}s, Z_{s-}) \\ & + \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2}\,\mathrm{d}\left([\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})]_s - \lceil\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})\rceil_s\right)\end{aligned}$$

we obtain

$$\langle Z_t, Z_t\rangle = \langle Z_0, Z_0\rangle + A_t + \tilde{M}_t.$$

To finish the proof of Theorem 2.1, we apply Lemma 2.1 to the process $\langle Z_t, Z_t\rangle$. First we verify that $A_t \in \mathcal{V} \cap \mathcal{P}$ and $\tilde{M}_t \in \mathcal{M}_{\mathrm{loc}}$. Considering the integrals that appear in the definition of A_t, we see that all integrators are increasing, hence almost all paths of A_t are of finite variation on each compact interval of $\mathbb{R}_+$. In the definition of $\tilde{M}_t$ there are only local martingales as integrators, thus $\tilde{M}_t$ itself must be a local martingale.

Now, we seek two processes $A_t^1, A_t^2 \in \mathcal{V}^+ \cap \mathcal{P}$ that satisfy $A_t \leq A_t^1 - A_t^2$ and $A_t^1 - A_t \in \mathcal{V}^+$. These two processes will be defined later on. Beforehand we give some inequalities required to understand the definitions of the processes. The first inequality is

$$\begin{aligned}-2a_s \sum_{i=1}^d \Big(Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i\Big) &= -2a_s \Big\langle\, Z_{s-}, \frac{1}{2c_s}\{\cdot\}_\cdot \,\Big\rangle \\ &= -2a_s \Big\langle\, Z_{s-}, \frac{1}{2c_s}\{\cdot\}_\cdot - \nabla f(Z_{s-}) + \nabla f(Z_{s-}) \,\Big\rangle\end{aligned}$$

$$
\begin{aligned}
&\leq -2a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle + 2a_s |\langle Z_{s-}, \frac{1}{2c_s}\{\cdot\}. - \nabla f(Z_{s-})\rangle| \\
&\overset{(\alpha)}{\leq} -2a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle + 2a_s \|Z_{s-}\| \, \|\frac{1}{2c_s}\{\cdot\}. - \nabla f(Z_{s-})\| \\
&\overset{(\beta)}{\leq} -2a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle + \sqrt{d} L a_s c_s \|Z_{s-}\|, \qquad (14)
\end{aligned}
$$

where (α) follows by the Cauchy-Schwarz inequality and (β) follows by

$$
\begin{aligned}
\left\|\frac{1}{2c_s}\{\cdot\}. - \nabla f(Z_{s-})\right\| &= \frac{1}{2c_s} \|\{\cdot\}. - 2c_s \nabla f(Z_{s-})\| \\
&= \frac{1}{2c_s} \|\{f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) - \langle 2c_s e_i, \nabla f(Z_{s-})\rangle\}.\| \\
&= \frac{1}{2c_s} \sqrt{\sum_{i=1}^{d} |f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) - \langle 2c_s e_i, \nabla f(Z_{s-})\rangle|^2} \\
&\leq \frac{1}{2c_s} \sqrt{dL^2 c_s^4} = \frac{1}{2}\sqrt{d} L c_s \qquad (15)
\end{aligned}
$$

with a Taylor expansion (see (17)).

The other inequality strived for is

$$
\begin{aligned}
\sum_{i=1}^{d} \frac{{a_s}^2}{4{c_s}^2} \{\cdot\}_i^2 \Delta R_s &= \sum_{i=1}^{d} a_s^2 \left|\frac{1}{2c_s}\{\cdot\}_i\right|^2 \Delta R_s \overset{(\gamma)}{\leq} \sum_{i=1}^{d} a_s^2 \left(|\nabla_i f(Z_{s-})| + \frac{1}{2} L c_s\right)^2 \Delta R_s \\
&= a_s^2 \|\nabla f(Z_{s-})\|^2 \Delta R_s + L a_s^2 c_s \sum_{i=1}^{d} |\nabla_i f(Z_{s-})| \Delta R_s + \frac{1}{4} d L^2 a_s^2 c_s^2 \Delta R_s \\
&\overset{(\delta)}{\leq} a_s^2 \|\nabla f(Z_{s-})\|^2 \Delta R_s + L d a_s^2 c_s \|\nabla f(Z_{s-})\| \Delta R_s + \frac{1}{4} d L^2 a_s^2 c_s^2 \Delta R_s, \qquad (16)
\end{aligned}
$$

where we get (γ) by

$$
\begin{aligned}
\left|\frac{1}{2c_s}\{\cdot\}_i\right| &= \left|\frac{1}{2c_s}\Big(f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) - \langle 2c_s e_i, \nabla f(Z_{s-})\rangle\Big) + \langle e_i, \nabla f(Z_{s-})\rangle\right| \\
&\leq \frac{1}{2c_s} |f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i) - \langle 2c_s e_i, \nabla f(Z_{s-})\rangle| + |\langle e_i, \nabla f(Z_{s-})\rangle| \\
&= \frac{1}{2c_s} \left|\int_{-1}^{1} \langle c_s e_i, \nabla f(Z_{s-} + t c_s e_i) - \nabla f(Z_{s-})\rangle \, \mathrm{d}t\right| + |\nabla_i f(Z_{s-})| \\
&\leq \frac{1}{2c_s} \|c_s e_i\| \int_{-1}^{1} L|t| \|c_s e_i\| \mathrm{d}t + |\nabla_i f(Z_{s-})| \\
&= \frac{1}{2} L c_s + |\nabla_i f(Z_{s-})| \qquad (17)
\end{aligned}
$$

once again with a Taylor expansion (see (15)). The inequality (δ) follows from Chebyshev's inequality.

We define two increasing processes by

$$\begin{aligned} A_t^1 &:= \sqrt{d}L\int_0^t a_s c_s \left\|Z_{s-}\right\| \,\mathrm{d}R_s + \int_0^t a_s^2 \|\nabla f(Z_{s-})\|^2 \Delta R_s \,\mathrm{d}R_s^d \\ &\quad + Ld\int_0^t a_s^2 c_s \|\nabla f(Z_{s-})\| \Delta R_s \,\mathrm{d}R_s^d + \frac{1}{4}dL^2 \int_0^t a_s^2 c_s^2 \Delta R_s \,\mathrm{d}R_s^d \\ &\quad + \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2} \,\mathrm{d}\lceil \int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-}) \rceil_s \\ A_t^2 &:= 2\int_0^t a_s \left\langle Z_{s-}, \nabla f(Z_{s-})\right\rangle \,\mathrm{d}R_s. \end{aligned}$$

Now, according to inequalities (14) and (16), we can ensure $A_t \leq A_t^1 - A_t^2$. To apply Lemma 2.1, we only have to verify $A_t^1 - A_t \in \mathcal{V}^+$. Hence, we focus our attention on

$$\begin{aligned} A_t^1 - A_t &= \sqrt{d}L\int_0^t a_s c_s \left\|Z_{s-}\right\| \,\mathrm{d}R_s + \int_0^t a_s^2 \|\nabla f(Z_{s-})\|^2 \Delta R_s \,\mathrm{d}R_s^d \\ &\quad + Ld\int_0^t a_s^2 c_s \|\nabla f(Z_{s-})\| \Delta R_s \,\mathrm{d}R_s^d + \frac{1}{4}dL^2 \int_0^t a_s^2 c_s^2 \Delta R_s \,\mathrm{d}R_s^d \\ &\quad + \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2} \,\mathrm{d}\lceil \int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-}) \rceil_s + 2\sum_{i=1}^d \int_0^t a_s Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i \,\mathrm{d}R_s \\ &\quad - \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2} \{\cdot\}_i^2 \Delta R_s \,\mathrm{d}R_s^d - \sum_{i=1}^d \int_0^t \frac{{a_s}^2}{4{c_s}^2} \,\mathrm{d}\lceil \int_0^{\cdot} M_i(Z_{t-}, \mathrm{d}t) \rceil_s \\ &= \int_0^t a_s \left(\sqrt{d}Lc_s \left\|Z_{s-}\right\| + 2\sum_{i=1}^d Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i \right) \,\mathrm{d}R_s \\ &\quad + \int_0^t a_s^2 \left(\|\nabla f(Z_{s-})\|^2 + Ldc_s \|\nabla f(Z_{s-})\| + \frac{dL^2}{4} c_s^2 - \sum_{i=1}^d \frac{\{\cdot\}_i^2}{4{c_s}^2} \right) \Delta R_s \,\mathrm{d}R_s. \end{aligned} \tag{18}$$

To prove that $A_t^1 - A_t$ is increasing, it will suffice to show that the expressions in both brackets in (18) are non-negative, because the integrators R_t and R_t^d are increasing. Concerning the expression in the first bracket, the inequality we have is

$$\sqrt{d}Lc_s \left\|Z_{s-}\right\| + 2\sum_{i=1}^d Z_{s-}^i \frac{1}{2c_s}\{\cdot\}_i \overset{(14)}{\geq} 2\left\langle Z_{s-}, \nabla f(Z_{s-})\right\rangle \geq 0.$$

The second bracket satisfies

$$\|\nabla f(Z_{s-})\|^2 + Ldc_s \|\nabla f(Z_{s-})\| + \frac{1}{4}dL^2 c_s^2 - \sum_{i=1}^d \frac{1}{4{c_s}^2}\{\cdot\}_i^2$$

$$= \|\nabla f(Z_{s-})\|^2 + Ldc_s\|\nabla f(Z_s)\| + \frac{1}{4}dL^2c_s^2 - \sum_{i=1}^{d}\left(\frac{1}{2c_s}\{\cdot\}_i\right)^2$$
$$\overset{(16)}{\geq} 0.$$

Hence Lemma 2.1 is applicable to $\|Z_s\|^2$, and we obtain

$$\left\{\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}\,\mathrm{d}A_s^1 < \infty\right\} \subseteq \left\{\|Z_s\|^2 \to\right\} \cap \left\{A_\infty^2 < \infty\right\}. \tag{19}$$

To complete the proof of convergence of the process Z_s, it is sufficient to show

$$\left\{\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}\,\mathrm{d}A_s^1 < \infty\right\} = \Omega. \tag{20}$$

Since

$$\begin{aligned}
&\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}\,\mathrm{d}A_s^1 \\
&\quad= \sqrt{d}L\int_0^\infty \frac{\|Z_{s-}\|}{1+\|Z_{s-}\|^2}a_sc_s\,\mathrm{d}R_s + \int_0^\infty \frac{\|\nabla f(Z_{s-})\|^2}{1+\|Z_{s-}\|^2}a_s^2\Delta R_s\,\mathrm{d}R_s^d \\
&\quad+ Ld\int_0^\infty \frac{\|\nabla f(Z_{s-})\|}{1+\|Z_{s-}\|^2}a_s^2c_s\Delta R_s\,\mathrm{d}R_s^d + \frac{1}{4}dL^2\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}a_s^2c_s^2\Delta R_s\,\mathrm{d}R_s^d \\
&\quad+ \sum_{i=1}^{d}\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}\frac{{a_s}^2}{4{c_s}^2}\,\mathrm{d}\left[\int_0^\cdot M_i(\mathrm{d}t, Z_{t-})\right]_s,
\end{aligned}$$

it suffices to show that all integrals appearing on the right-hand side are almost surely finite:

$$\begin{aligned}
\int_0^\infty \frac{\|Z_{s-}\|}{1+\|Z_{s-}\|^2}a_sc_s\,\mathrm{d}R_s &\leq \int_0^\infty a_sc_s\,\mathrm{d}R_s < \infty \\
\int_0^\infty \frac{\|\nabla f(Z_{s-})\|^2}{1+\|Z_{s-}\|^2}a_s^2\Delta R_s\,\mathrm{d}R_s^d &\leq \mathcal{C}\int_0^\infty \frac{\|Z_{s-}\|^2}{1+\|Z_{s-}\|^2}a_s^2\Delta R_s\,\mathrm{d}R_s^d \\
&\leq \mathcal{C}\int_0^\infty a_s^2\Delta R_s\,\mathrm{d}R_s^d < \infty \\
\int_0^\infty \frac{\|\nabla f(Z_{s-})\|}{1+\|Z_{s-}\|^2}a_s^2c_s\Delta R_s\,\mathrm{d}R_s^d &\leq \mathcal{C}\int_0^\infty \frac{\|Z_{s-}\|}{1+\|Z_{s-}\|^2}a_s^2c_s\Delta R_s\,\mathrm{d}R_s^d \\
&\leq \mathcal{C}\int_0^\infty a_s^2\Delta R_s\,\mathrm{d}R_s^d < \infty \\
\int_0^\infty \frac{1}{1+\|Z_{s-}\|^2}a_s^2c_s^2\Delta R_s\,\mathrm{d}R_s^d &\leq \int_0^\infty a_s^2c_s^2\Delta R_s\,\mathrm{d}R_s^d \leq \mathcal{C}\int_0^\infty a_s^2\Delta R_s\,\mathrm{d}R_s^d < \infty
\end{aligned}$$

and finally

$$\int_0^\infty \frac{a_s{}^2}{4c_s{}^2} \frac{1}{1+\|Z_{s-}\|^2} \,\mathrm{d}\lceil \int_0^\cdot M_i(\mathrm{d}t, Z_{t-})\rceil_s = \int_0^\infty \frac{a_s{}^2}{4c_s{}^2} \frac{h_s^{ii}(Z_{s-})}{1+\|Z_{s-}\|^2} \,\mathrm{d}R_s < \infty.$$

This proves almost sure convergence of $\|Z_t\|^2$ and consequently of Z_t.

(b) Proof of almost sure convergence of Z_t to the stationary point of f. We have to show that Z_s converges almost surely to 0. From (19) and (20) in the proof of part (a), we know that

$$\Omega = \{\|Z_s\|^2 \to\} \cap \{A_\infty^2 < \infty\}$$

and

$$\Omega = \{A_\infty^2 < \infty\},$$

hold. The convergence of Z_s to 0 will be proved by contradiction. Therefore we assume that there exists a set N of non-zero probability on which the solution of the stochastic integral equation does not converge to 0. Now we will deduce a contradiction to

$$\Omega = \{A_\infty^2 < \infty\} \tag{21}$$

that holds according to (19) and (20) in the proof of part (a). Note that

$$A_\infty^2 = \int_0^\infty \mathrm{d}A_s^2 + A_0^2 = 2\int_0^\infty a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle \,\mathrm{d}R_s + A_0^2.$$

As proved in part (a), Z_s converges for almost all $\omega \in \Omega$, but for all $\omega \in N$ the process does not converge to 0. Thus it follows that for almost all $\omega \in N$

$$\underset{\epsilon^\star>0}{\exists} \quad \underset{s_0}{\exists} \quad \underset{s\geq s_0}{\forall} \qquad \epsilon^\star \leq \|Z_s\| \leq 1/\epsilon^\star. \tag{22}$$

Using (22) we find that for almost all $\omega \in N$

$$\begin{aligned} A_\infty^2 &= 2\int_0^\infty a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle \,\mathrm{d}R_s + A_0^2 \\ &= 2\int_0^{s_0} a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle \,\mathrm{d}R_s + 2\int_{s_0+}^\infty a_s \langle Z_{s-}, \nabla f(Z_{s-})\rangle \,\mathrm{d}R_s + A_0^2 \\ &\geq \mathcal{C}_\omega + 2C(\epsilon^\star)\int_{s_0+}^\infty a_s \,\mathrm{d}R_s = \infty \end{aligned}$$

which contradicts (21). Hence such a set N cannot exist. This completes the proof. □

2.2 Almost Sure Convergence Rate

Since we have verified consistency, it raises the question how fast Z_t converges. We seek a result of the form

$$\gamma_t(\delta)\|Z_t - x^\star\| \to 0 \quad \text{a.s.} \qquad \text{where} \quad \gamma_t(\delta) := \mathcal{E}_t(\delta \int_0^\cdot a_s \,\mathrm{d}R_s)$$

and $\mathcal{E}_t(\cdot)$ is the stochastic exponential. Naturally stronger assumptions are needed to get convergence rates than to get the consistency. Therefore we consider the following additional set of conditions here:

(D^2)

$$\int_0^\infty \gamma_{s-}(\delta) a_s c_s \,\mathrm{d}R_s < \infty$$

(D^3)

$$\int_0^\infty \gamma_{s-}(\delta) a_s c_s^2 \,\mathrm{d}R_s < \infty$$

$(\tilde{E})$ For every $i \in \{1, .., d\}$ and $x \in \mathbb{R}^d$, we have

$$\int_0^\infty \frac{a_s^2}{c_s^2} \frac{\gamma_{s-}^2(\delta) h_s^{ii}(Z_{s-})}{1 + \gamma_{s-}^2(\delta)\|Z_{s-}\|^2} \,\mathrm{d}R_s < \infty, \quad \text{where} \quad h_s^{ii}(x) := \frac{\mathrm{d}\left[\int_0^\cdot M_i(\mathrm{d}t, x)\right]_s}{\mathrm{d}R_s}.$$

Except for the fact that the process $\gamma_{s-}(\delta)$ describes the speed, the conditions above are quite similar to the conditions presented in the section before. Therefore we do not give further explanations here.

Remark 2.3. Considering processes $\gamma_{s-}(\delta)$ obeying condition (D^2) or (D^3), it turns out that only weight processes a_t that are mainly of the form $a_t := a(1 + R_t)^{-1}$ come into consideration. But here we do not assume that the process R_t and thus the weight processes a_t and c_t are deterministic. This is the first reason to present Theorem 2.2 in this chapter. The other reason is the fact that the theorem, and especially the proof, can be used as a starting point for considerations on weight processes of different design. To make an entry into such a consideration easier we use general representations in the proof of the theorem.

Theorem 2.2. *Let the conditions (A)–(F), which ensure the consistency, be fulfilled. We assume that f is two or three times differentiable at $x^\star$, in both cases with a continuous Hessian H around $x^\star$. Furthermore, let conditions $(\tilde{E})$ and (D^2) or (D^3) hold, if f is two or three times differentiable at $x^\star$, respectively. Then we get for all* $0 \le \delta < \lambda_{\min}$

$$\gamma_t(\delta)\,\|Z_t - x^\star\| \overset{t\to\infty}{\longrightarrow} 0 \quad \textit{a.s.}.$$

Proof of Theorem 2.2

In both parts we follow the path suggested in Theorem 2.1. Without loss of generality we assume $x^\star = 0$. Then we examine the positive process $\gamma_t^2(\delta)\|Z_t\|^2$ instead of $\|Z_t\|^2$. By

$$\begin{aligned}\gamma_t^2(\delta) &= \gamma_t(\delta)\gamma_t(\delta) = \mathcal{E}_t(2\delta\int_0^{\cdot} a_s\,\mathrm{d}R_s + \delta^2\int_0^{\cdot} a_s^2\,\mathrm{d}[R,R]_s)\\ &= \mathcal{E}_t(2\delta\int_0^{\cdot} a_s\,\mathrm{d}R_s + \delta^2\int_0^{\cdot} a_s^2\Delta R_s\,\mathrm{d}R_s)\end{aligned}$$

we find

$$\begin{aligned}&\gamma_t^2(\delta)\langle Z_t, Z_t\rangle - \gamma_0^2(\delta)\langle Z_0, Z_0\rangle\\ &\quad= \sum_{i=1}^d\Big(\int_0^t \gamma_{s-}^2(\delta)\,\mathrm{d}Z_s^{i^2} + \int_0^t Z_{s-}^{i^2}\,\mathrm{d}\gamma_s^2(\delta) + \int_0^t \mathrm{d}[\gamma^2(\delta), Z^{i^2}]_s\Big)\\ &\quad= \sum_{i=1}^d\Big(\int_0^t \gamma_{s-}^2(\delta)\,\mathrm{d}Z_s^{i^2} + \int_0^t Z_{s-}^{i^2}\,\mathrm{d}\gamma_s^2(\delta) + \int_0^t \Delta\gamma_s^2(\delta)\,\mathrm{d}Z_s^{i^2}\Big)\\ &\quad= \sum_{i=1}^d\Big(\int_0^t \gamma_{s-}^2(\delta)\,\mathrm{d}Z_s^{i^2} + \int_0^t Z_{s-}^{i^2}\,\mathrm{d}\gamma_s^2(\delta) + \int_0^t (\gamma_s^2(\delta) - \gamma_{s-}^2(\delta))\,\mathrm{d}Z_s^{i^2}\Big)\\ &\quad= -2\int_0^t \gamma_s^2(\delta)a_s\sum_{i=1}^d Z_{s-}^i\frac{1}{2c_s}\{\cdot\}_i\,\mathrm{d}R_s + \int_0^t \gamma_s^2(\delta)a_s^2\sum_{i=1}^d\frac{1}{4{c_s}^2}\{\cdot\}_i^2\Delta R_s\,\mathrm{d}R_s^d\\ &\qquad+\sum_{i=1}^d\int_0^t \gamma_s^2(\delta)\frac{{a_s}^2}{4{c_s}^2}\,\mathrm{d}\big[\int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-})\big]_s + 2\delta\int_0^t a_s\gamma_{s-}^2(\delta)\|Z_{s-}\|^2\,\mathrm{d}R_s\\ &\qquad+\delta^2\int_0^t \gamma_{s-}^2(\delta)\|Z_{s-}\|^2a_s^2\Delta R_s\,\mathrm{d}R_s + \int_0^t \mathrm{d}\tilde{M}_s\end{aligned}\tag{23}$$

and seek, as in the proof of Theorem 2.1, two suitable processes, A_t^1 and A_t^2, to apply Lemma 2.1. Before we define these processes, we require some inequalities. Since f is two or three times differentiable at 0, the Lipschitz-continuous gradient ∇f and the Hessian H_0 exist. We obtain

$$\begin{aligned}\frac{1}{2c_s}\{\cdot\}_i &= \frac{1}{2}\int_{-1}^1 \nabla_i f(Z_{s-} + tc_se_i)\,\mathrm{d}t\\ &= \frac{1}{2}\int_{-1}^1 \nabla_i f(Z_{s-})\,\mathrm{d}t + \overbrace{\frac{1}{2}\int_{-1}^1 (\nabla_i f(Z_{s-} + tc_se_i) - \nabla_i f(Z_{s-}))\,\mathrm{d}t}^{=:\,C_s^i}\\ &= \sum_{j=1}^d H_0^{ij}Z_{s-}^j + \underbrace{\nabla_i f(Z_{s-}) - \sum_{j=1}^d H_0^{ij}Z_{s-}^j}_{=:\,B_s^i} + C_s^i\\ &= \sum_{j=1}^d H_0^{ij}Z_{s-}^j + B_s^i + C_s^i.\end{aligned}\tag{24}$$

Therefore we get

$$\begin{aligned}\frac{1}{4c_s^2}\sum_{i=1}^{d}\{\cdot\}_i^2 &= \|\frac{1}{2c_s}\{\cdot\}_.\|^2 = \|H_0 Z_{s-} + B_s + C_s\|^2 \le 3\|H_0 Z_{s-}\|^2 + 3\|B_s\|^2 + 3\|C_s\|^2 \\ &\le 3\lambda_{\max}^2\|Z_{s-}\|^2 + 3\|B_s\|^2 + 3\|C_s\|^2 \end{aligned} \tag{25}$$

and

$$\begin{aligned}-\frac{1}{2c_s}\sum_{i=1}^{d}\{\cdot\}_i Z_{s-}^i &= -\sum_{i=1}^{d}\left(\sum_{j=1}^{d} H_0^{ij} Z_{s-}^j + B_s^i + C_s^i\right) Z_{s-}^i \\ &\le \underbrace{-Z_{s-}^T H_0 Z_{s-}}_{} + |\langle B_s, Z_{s-}\rangle| + |\langle C_s, Z_{s-}\rangle| \\ &\qquad = -\langle T^T D T Z_{s-}, Z_{s-}\rangle \le -\lambda_{\min}\|T Z_{s-}\|^2 = -\lambda_{\min}\|Z_{s-}\|^2 \\ &\le -\lambda_{\min}\|Z_{s-}\|^2 + (\|B_s\| + \|C_s\|)\|Z_{s-}\|. \end{aligned} \tag{26}$$

Later on it will be advantageous to have, on the one hand, decreasing rates and on the other hand, bounds for the processes $\|B_s\|$ and $\|C_s\|$. Since we assume that Z_s is a strong solution of the stochastic integral equation on $[0,\infty)$, we know that there are no explosion times. Furthermore, we know from Theorem 2.1 that Z_s converges to zero. If we combine these two statements, we get $\sup_s \|Z_s\| \le C(\omega) < \infty$, which we need to prove bounds. Since f has a Lipschitz-continuous gradient, we obtain the bounds

$$\begin{aligned}\|B_s\| &= \|\nabla f(Z_{s-}) - H_0 Z_{s-}\| \le \|\nabla f(Z_{s-})\| + \|H_0\|\,\|Z_{s-}\| \\ &\le (L + \lambda_{\max})\,\|Z_{s-}\| \le \mathcal{C}\,\|Z_{s-}\| \le \mathcal{C}_\omega < \infty\end{aligned}$$

and

$$\begin{aligned}\|C_s\| &= \left(\sum_{i=1}^{d}|C_s^i|^2\right)^{0.5} \le \frac{1}{2}\left(\sum_{i=1}^{d}\left(\int_{-1}^{1}|\nabla_i f(Z_{s-} + t c_s e_i) - \nabla_i f(Z_{s-})|\,\mathrm{d}t\right)^2\right)^{0.5} \\ &\le \frac{1}{2}\left(\sum_{i=1}^{d}\left(L c_s \int_{-1}^{1}|t|\,\mathrm{d}t\right)^2\right)^{0.5} \le \frac{1}{2}\sqrt{dL^2 c_s^2} \le L\sqrt{d}\,c_s \le L\sqrt{d}\,c_0.\end{aligned}$$

If f is two times differentiable at 0, we use the statement $\|C_s\| \le L\sqrt{d}\,c_s$ above to obtain the asymptotic behavior of $\|C_s\|$. Now we consider the asymptotic behavior of $\|C_s\|$ in the case where f is three times differentiable at 0. Since $\|Z_s\|$ and $\|c_s\|$ converge to zero, they are for sufficiently large time in a small neighborhood of 0. In addition, we know from the assumptions that the Hessian is continuous in a small neighborhood of 0 and therefore we may use a Taylor expansion at 0. Furthermore, we use the fact that the second partial derivative is differentiable at 0. We obtain

$$C_s^i = \frac{1}{2c_s}\{\cdot\}_i - \nabla_i f(Z_{s-})$$

$$
\begin{aligned}
&= \frac{1}{2c_s}\sum_{l,k}\left(\frac{\partial^2 f}{\partial x_l \partial x_k}(Z_{s-}+\theta_s^1 c_s e_i) - \frac{\partial^2 f}{\partial x_l \partial x_k}(Z_{s-}-\theta_s^2 c_s e_i)\right)(c_s e_i)_l (c_s e_i)_k \\
&= \frac{c_s}{4}\left(\frac{\partial^2 f}{(\partial x_i)^2}(Z_{s-}+\theta_s^1 c_s e_i) - \frac{\partial^2 f}{(\partial x_i)^2}(Z_{s-}-\theta_s^2 c_s e_i)\right) \\
&= \frac{\theta_s^1+\theta_s^2}{4}\frac{\partial^3 f}{(\partial x_i)^3}(0)c_s^2 + \mathrm{o}(c_s\|Z_{s-}\| + c_s^2) = \mathcal{O}(c_s^2) + \mathrm{o}(c_s\|Z_{s-}\|) + \mathrm{o}(c_s^2) \\
&= \mathrm{o}(c_s\|Z_{s-}\|) + \mathcal{O}(c_s^2).
\end{aligned}
$$

and $\|C_s\| = \mathrm{o}(c_s\|Z_{s-}\|) + \mathcal{O}(c_s^2)$.

The asymptotic behavior of $\|B_s\|$ is given by

$$
\|B_s\| = \|\nabla f(Z_{s-}) - H_0 Z_{s-}\| = \|\nabla f(0) + H_0 Z_{s-} + \mathrm{o}(\|Z_{s-}\|) - H_0 Z_{s-}\| = \mathrm{o}(\|Z_{s-}\|)
$$

and holds in the cases of a two and three times differentiable f. We notice that the faster convergence rate in case of a three times differentiable f follows from the difference in the asymptotic behavior of $\|B_s\| + \|C_s\|$, that is

$$
\|B_s\| + \|C_s\| = \begin{cases} \mathrm{o}(\|Z_{s-}\|) + L\sqrt{d}c_s & = \mathrm{o}(\|Z_{s-}\|) + \mathcal{O}(c_s), \text{ if } f \text{ is } 2\times \text{ diff.} \\ \mathrm{o}(\|Z_{s-}\|) + \mathrm{o}(c_s\|Z_{s-}\|) + \mathcal{O}(c_s^2) & = \mathrm{o}(\|Z_{s-}\|) + \mathcal{O}(c_s^2), \text{ if } f \text{ is } 3\times \text{ diff.} \end{cases}
$$

Coming back to (23) we seek two processes A_t^1 and A_t^2 satisfying the assumptions of Lemma 2.1 and

$$
\gamma_t(\delta)\langle Z_t, Z_t\rangle - \gamma_0(\delta)\langle Z_0, Z_0\rangle \quad \leq A_t^1 - A_t^2 + \tilde{M}_t.
$$

These can be find using the above inequalities:

$$
\begin{aligned}
A_t^2 &:= 2\int_0^t \gamma_s^2(\delta) a_s \left((\delta - \lambda_{\min})\|Z_{s-}\|^2 + \|B_s\|\|Z_{s-}\|\right)^- \,\mathrm{d}R_s \\
A_t^1 &:= 2\int_0^t \gamma_s^2(\delta) a_s \left((\delta - \lambda_{\min})\|Z_{s-}\|^2 + \|B_s\|\|Z_{s-}\|\right)^+ \,\mathrm{d}R_s \\
&\quad + 3\int_0^t \gamma_s^2(\delta) a_s^2 \|B_s\|^2 \Delta R_s \,\mathrm{d}R_s^d + (3\lambda_{\max}^2 + \delta^2)\int_0^t \gamma_s^2(\delta) a_s^2 \|Z_{s-}\|^2 \Delta R_s \,\mathrm{d}R_s^d \\
&\quad + 2\int_0^t \gamma_s^2(\delta) a_s \|C_s\|\|Z_{s-}\| \,\mathrm{d}R_s + 3\int_0^t \gamma_s^2(\delta) a_s^2 \|C_s\|^2 \Delta R_s \,\mathrm{d}R_s^d \\
&\quad + \sum_{i=1}^d \int_0^t \gamma_s^2(\delta) \frac{{a_s}^2}{4{c_s}^2} \,\mathrm{d}\lceil \int_0^\cdot M_i(\mathrm{d}t, Z_{t-}) \rceil_s.
\end{aligned}
$$

Now it is sufficient to prove

$$
\int_0^\infty \frac{1}{1 + \gamma_{s-}^2(\delta)\langle Z_{s-}, Z_{s-}\rangle} \,\mathrm{d}A_s^1 < \infty. \tag{27}
$$

The following representation helps us to use $1/(1+\gamma_{s-}^2(\delta)\langle Z_{s-}, Z_{s-}\rangle)$ effectively.

$$\begin{aligned}
\gamma_t(\delta) &= \mathcal{E}_t(\delta \int_0^{\cdot} a_s \,\mathrm{d}R_s) = \exp(\delta \int_0^t a_s \,\mathrm{d}R_s) \prod_{0<s\leq t} (1+\delta a_s \Delta R_s) \exp(-\delta a_s \Delta R_s) \\
&= \exp(\delta a_t \Delta R_t) \Big(\exp(\delta \int_0^{t-} a_s \,\mathrm{d}R_s) \prod_{0<s<t} (1+\delta a_s \Delta R_s) \exp(-\delta a_s \Delta R_s) \Big) \\
&\quad \cdot (1+\delta a_t \Delta R_t) \exp(-\delta a_t \Delta R_t) \\
&= \gamma_{t-}(\delta)(1+\delta a_t \Delta R_t)
\end{aligned} \tag{28}$$

From the assumptions $\int_0^\infty a_s^2 \Delta R_s \,\mathrm{d}R_s^d < \infty$, $\int_0^\infty a_s \,\mathrm{d}R_s = \infty$, and (28) it follows

$$\frac{\gamma_t(\delta)}{\gamma_{t-}(\delta)} = (1+\delta a_t \Delta R_t) = (1+\mathrm{o_b}(1)) = \mathcal{C}_\omega.$$

In view of the first summand in the expanded form of (27), it will suffice to show in both cases that

$$\underset{s_0(\omega)}{\exists} \quad \underset{s\geq s_0(\omega)}{\forall} \qquad \Big((\delta - \lambda_{\min}) \|Z_{s-}\|^2 + \|B_s\| \|Z_{s-}\| \Big)^+ = 0.$$

Using $\lambda_{\min} > \delta$ and the fact that $\|B_s\| = \mathrm{o}(\|Z_{s-}\|)$ for increasing s, the term in the bracket is negative for sufficiently large s_0. Furthermore, we have

$$\int_0^{s_0(\omega)} \Big((\delta - \lambda_{\min}) \|Z_{s-}\|^2 + \|B_s\| \|Z_{s-}\| \Big)^+ \gamma_s^2(\delta) a_s \,\mathrm{d}R_s \leq \mathcal{C}_\omega \int_0^{s_0(\omega)} \gamma_s^2(\delta) a_s \,\mathrm{d}R_s < \infty.$$

In both cases, the third summand can be estimated by

$$\int_0^\infty \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)} \right)^2 \frac{\gamma_{s-}^2(\delta) \|Z_{s-}\|^2}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s^2 \Delta R_s \,\mathrm{d}R_s^d \leq \mathcal{C}_\omega \int_0^\infty a_s^2 \Delta R_s \,\mathrm{d}R_s^d < \infty.$$

In both cases, the last summand can be estimated by

$$\begin{aligned}
\int_0^\infty \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)} \right)^2 &\frac{\gamma_{s-}^2(\delta)}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} \frac{{a_s}^2}{{c_s}^2} \,\mathrm{d}\lceil \int_0^{\cdot} M_i(\mathrm{d}t, Z_{t-}) \rceil_s \\
&\leq \mathcal{C}_\omega \int_0^\infty \frac{\gamma_{s-}^2(\delta) h_s^{ii}(Z_{s-})}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} \frac{{a_s}^2}{{c_s}^2} \,\mathrm{d}R_s < \infty.
\end{aligned}$$

For the remaining summands we use the fact that the asymptotic behavior of $\|B_s\|$ and $\|C_s\|$ is known and that both are bounded. Since $\|Z_s\|$ and c_s converge to zero there exists a time $\tau(\omega) < \infty$ such that the estimates $\mathcal{C}c_t$ and $\mathcal{C}\|Z_t\|$ follows from $\mathrm{o}(c_t)$ and $\mathcal{O}(\|Z_t\|)$, respectively, for all $t \geq \tau(\omega)$ and for almost all $\omega \in \Omega$. For the second summand we get

$$\int_0^\infty \frac{\gamma_s^2(\delta)\|B_s\|^2}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s^2 \Delta R_s \,\mathrm{d}R_s^d$$

$$\leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_{\tau(\omega)}^{\infty} \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)}\right)^2 \frac{\gamma_{s-}^2(\delta)\|Z_{s-}\|^2}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s^2 \Delta R_s \,\mathrm{d}R_s^d$$
$$\leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_0^{\infty} a_s^2 \Delta R_s \,\mathrm{d}R_s^d < \infty.$$

Consider the fourth summand. In the case where f is two times differentiable at 0, we have

$$\int_0^{\infty} \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)}\right)^2 \frac{\gamma_{s-}(\delta)\|Z_{s-}\|}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} \|C_s\| \gamma_{s-}(\delta) a_s \,\mathrm{d}R_s \leq \mathcal{C}_\omega \int_0^{\infty} \gamma_{s-}(\delta) a_s c_s \,\mathrm{d}R_s < \infty$$

and in the case where f is three times differentiable at 0,

$$\int_0^{\infty} \frac{\gamma_s^2(\delta)\|C_s\|\|Z_{s-}\|}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s \,\mathrm{d}R_s$$
$$\leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_{\tau(\omega)}^{\infty} \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)}\right)^2 \frac{(c_s\|Z_{s-}\| + c_s^2)\gamma_{s-}^2(\delta)\|Z_{s-}\|}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s \,\mathrm{d}R_s$$
$$\leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_0^{\infty} a_s c_s \,\mathrm{d}R_s + \mathcal{C}_\omega \int_0^{\infty} \gamma_{s-}(\delta) a_s c_s^2 \,\mathrm{d}R_s < \infty.$$

In the case where f is two times differentiable at 0, we get for the fifth summand

$$\int_0^{\infty} \gamma_s^2(\delta)\|C_s\|^2 a_s^2 \Delta R_s \,\mathrm{d}R_s \leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_{\tau(\omega)}^{\infty} (\gamma_{s-}(\delta)c_s)^2 a_s^2 \Delta R_s \,\mathrm{d}R_s^d < \infty$$

using $\int_0^\infty \gamma_{s-}(\delta) a_s c_s \,\mathrm{d}R_s < \infty$, $\int_0^\infty a_s \,\mathrm{d}R_s = \infty$ and hence $\gamma_{s-}(\delta)c_s \to 0$. In the case where f is three times differentiable at 0, we get

$$\int_0^{\infty} \left(\frac{\gamma_s(\delta)}{\gamma_{s-}(\delta)}\right)^2 \frac{\gamma_{s-}^2(\delta)\|C_s\|^2}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2} a_s^2 \Delta R_s \,\mathrm{d}R_s$$
$$\leq \mathcal{C}_\omega + \mathcal{C}_\omega \int_{\tau(\omega)}^{\infty} \left(c_s^2 + \gamma_{s-}^2(\delta)\, c_s^4\right) a_s^2 \Delta R_s \,\mathrm{d}R_s < \infty$$

as $\|C_s\|^2 \leq \mathcal{C}_\omega(c_s^2\|Z_{s-}\|^2 + c_s^4)$ holds for s larger than $\tau(\omega)$. Now, the convergence of $\gamma_t^2(\delta)\|Z_t\|^2$ follows from Lemma 2.1.

To show almost sure convergence of $\gamma_t^2(\delta)\|Z_t\|^2$ to 0 we use the whole conclusion of Lemma 2.1

$$\Omega = \{\gamma_s^2(\delta)\|Z_s\|^2 \to\} \cap \{A_\infty^2 < \infty\}.$$

Therefore we know that $\Omega = \{A_\infty^2 < \infty\}$ holds. Convergence of $\gamma_t^2(\delta)\|Z_t\|^2$ to 0 will be proved by contradiction. To this end, we assume that there exists a set N of non-zero probability on which $\gamma_t^2(\delta)\|Z_t\|^2$ does not converge to 0. Now we will deduce a contradiction to $\Omega = \{A_\infty^2 < \infty\}$. Note

$$A_\infty^2 = \int_0^{\infty} \mathrm{d}A_s^2 + A_0^2 = 2\int_0^{\infty} \gamma_s^2(\delta) a_s \left(\left(\delta - \lambda_{\min}\right)\|Z_{s-}\|^2 + \|B_s\|\|Z_{s-}\|\right)^- \mathrm{d}R_s + A_0^2$$

and further, since $\|B_s\| = \mathrm{o}(\|Z_{s-}\|)$ holds, the existence of $s_0^1(\omega) < \infty$ such that for all $s \geq s_0^1(\omega)$ the relation $\|B_s\| \leq \frac{1}{2}(\lambda_{\min} - \delta)\|Z_{s-}\|$ holds. We assess the term in brackets for $s \geq s_0^1(\omega)$ and observe

$$\begin{aligned}\Big((\delta - \lambda_{\min})\,\|Z_{s-}\|^2 + \|B_s\|\|Z_{s-}\|\Big)^- &\geq \Big((\delta - \lambda_{\min})\,\|Z_{s-}\|^2 + \frac{1}{2}(\lambda_{\min} - \delta)\|Z_{s-}\|^2\Big)^- \\ &= \frac{1}{2}(\lambda_{\min} - \delta)\|Z_{s-}\|^2 \quad \text{ with } \lambda_{\min} > \delta.\end{aligned}$$

As proved before, $\gamma_t^2(\delta)\|Z_t\|^2$ converges for almost all $\omega \in \Omega$, but for all $\omega \in N$ the expression does not converge to 0. Thus it follows that, for almost all $\omega \in N$,

$$\underset{\epsilon^\star > 0}{\exists} \quad \underset{s_0^2}{\exists} \quad \underset{s \geq s_0^2}{\forall} \qquad \epsilon^\star \leq \gamma_{t-}^2(\delta)\|Z_{t-}\|^2.$$

Using the equation above with $s_0 := \max\{s_0^1, s_0^2\}$, we have

$$\begin{aligned}A_\infty^2 &= 2\int_0^\infty \Big((\delta - \lambda_{\min})\,\|Z_{s-}\|^2 + \|B_s\|\|Z_{s-}\|\Big)^- a_s\gamma_s^2(\delta)\,\mathrm{d}R_s + A_0^2 \\ &\geq (\lambda_{\min} - \delta)\int_{s_0}^\infty a_s\gamma_{s-}^2(\delta)\|Z_{s-}\|^2\,\mathrm{d}R_s \geq (\lambda_{\min} - \delta)\epsilon^\star\int_{s_0}^\infty a_s\,\mathrm{d}R_s \\ &\geq \underbrace{(\lambda_{\min} - \delta)\epsilon^\star}_{>0}\Bigg(\underbrace{\int_0^\infty a_s\,\mathrm{d}R_s}_{=\infty} - \underbrace{\int_0^{s_0} a_s\,\mathrm{d}R_s}_{\leq (R_{s_0} - R_0) < \infty}\Bigg) = \infty\end{aligned}$$

for almost all $\omega \in N$, which contradicts $\Omega = \{A_\infty^2 < \infty\}$. Hence such a set N does not exist. This completes the proof. □

2.3 Discussion of Special Settings

Here we illustrate the theorems presented in this chapter in the case of two common settings. One treats the discrete-time setting, the other treats the Itô setting. We obtain three known and one new result.

The following assumptions are needed in the special situation of a stochastic integral equation of Itô type, which will be considered in Corollary 2.1.

(D') The processes a_s and c_s, which the statistician has to choose, are left-continuous and satisfy $a_s, c_s > 0$, $a_s, c_s \downarrow 0$ and

$$\int_0^\infty a_s\,\mathrm{d}s = \infty \qquad \int_0^\infty a_s c_s\,\mathrm{d}s < \infty.$$

(E') For every $i, j \in \{1, .., d\}$ we have

$$\int_0^\infty \frac{\sigma_s^{ij}(Z_s)^2}{1 + \|Z_s\|^2} \frac{{a_s}^2}{{c_s}^2} \,\mathrm{d}s < \infty.$$

Corollary 2.1. *Consider the Itô-type stochastic integral equation*

$$Z_t = Z_0 - \int_0^t \frac{a_s}{2c_s} \left\{ f(Z_s + c_s e_i) - f(Z_s - c_s e_i) \,\mathrm{d}s + \sum_{j=1}^d \sigma_s^{ij}(Z_s) \,\mathrm{d}W_s^j \right\}_{i \in \{1,\ldots,d\}} \tag{29}$$

with $\sigma^{ij} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ *and independent standard Brownian motions* W_t^j *defined on our standard stochastic basis. If assumptions (A), (B), (C), (D′), and (E′) hold, then the strong solution* Z_t *of* (29) *converges almost surely to the stationary point of the function* f*.*

Remark 2.4. If in Corollary 2.1 we assume the stronger restrictions $\sigma_t^{ij}(x) \le C(1+\|x\|)$ and $\int_0^\infty {a_s}^2 {c_s}^{-2} \,\mathrm{d}s < \infty$ instead of (E'), Corollary 2.1 is well-known (see e.g. [30]).

Figure 1 shows a simulated path of the process Z_t given in (29). We choose $f(x) = 0.5x^T x$, $x \in \mathbb{R}^2$, $Z_0 = (5, 6)^T$, $a_s = (1+s)^{-3/4}$, $c_s = (1+s)^{-1/4}$, $\sigma_s(\cdot) = I$ and use the Milstein scheme (see, e.g., [14]) to simulate the path. Furthermore, Figure 2 illustrates how the process $f(Z_t)$ behaves on the surface of the considered function f by taking the same path of Z_t as in Figure 1.

The next assumptions are needed in the special situation of a recursive scheme that will be considered in Corollary 2.2.

(D'') The sequences a_n and c_n, which the statistician has to choose, satisfy $a_n, c_n > 0$, $a_n, c_n \downarrow 0$ and

$$\sum_{n=1}^\infty a_n = \infty \qquad \sum_{n=1}^\infty a_n c_n < \infty \qquad \sum_{n=1}^\infty \frac{a_n^2}{c_n^2} < \infty.$$

(E'') We have

$$\sup_{n \in \mathbb{N}} \mathbf{E}\|V_n\|^2 < \infty \qquad \text{and} \qquad \mathbf{E}\left(V_n | \mathcal{F}_{n-1}\right) = 0$$

where $\mathcal{F}_n := \mathcal{F}(Z_1, V_1, \ldots, Z_n, V_n)$.

Corollary 2.2. *If the assumptions (A), (B), (C), (D″), and (E″) hold, then the sequence* Z_n *generated by the recursive scheme*

$$Z_n - Z_{n-1} = -\frac{a_n}{2c_n} \left\{ f(Z_{n-1} + c_n e_i) - f(Z_{n-1} - c_n e_i) + V_n^i \right\}_{i \in \{1,\ldots,d\}} \tag{30}$$

converges almost surely to the stationary point of the function f*.*

This Corollary 2.2 is well-known (see e.g. [20]).

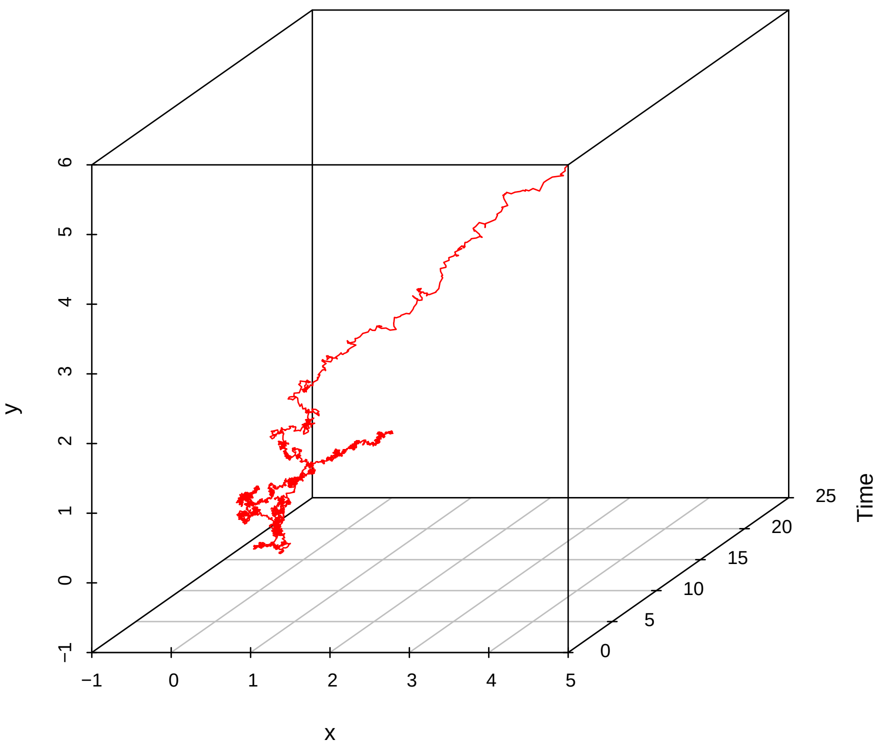

Figure 1: Simulation of a path of the process Z_t of (29).

Corollary 2.3. *Consider the stochastic integral equation* (29) *of Itô type as given in Corollary 2.1 with the restriction* $\sum_{j=1}^{d} \sigma_s^{ij}(x) \le C(1 + \|x\|)$ *for all* $i \in \{1, \dots, d\}$ *and assume* (A), (B), (C), *and* (E').

(a) If f *is twice differentiable at* $x^\star$, *Hessian* H *is continuous around* $x^\star$, *and* $\lambda_{\min} > \frac{1}{4a}$ *holds, we get*

$$(1+t)^\delta \|Z_t - x^\star\| \to 0 \quad a.s.$$

for all $\delta \in (0, \frac{1}{4})$, *where* $c_s := c(1+s)^{-\frac{1}{4}}$ *and* $a_s := a(1+s)^{-1}$.

(b) If f *is three times differentiable at* $x^\star$, *Hessian* H *is continuous around* $x^\star$, *and* $\lambda_{\min} > \frac{1}{3a}$ *holds, we get*

$$(1+t)^\delta \|Z_t - x^\star\| \to 0 \quad a.s.$$

for all $\delta \in (0, \frac{1}{3})$, *where* $c_s := c(1+s)^{-\frac{1}{6}}$ *and* $a_s := a(1+s)^{-1}$.

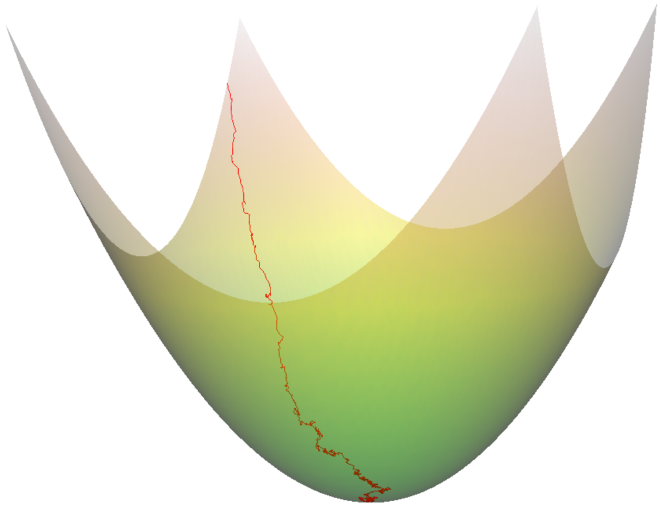

Figure 2: Simulation of a path of the process $f(Z_t)$, with the same path of Z_t as depicted in Figure 1.

Corollary 2.4. *Consider recursion* (30) *of Corollary 2.2 and assume* (A), (B), (C), *and* (E'').

(a) *If f is twice differentiable at $x^\star$, Hessian H is continuous around $x^\star$, and $\lambda_{\min} > \frac{1}{4a}$ holds, we get*

$$n^\delta \|Z_n - x^\star\| \to 0 \quad a.s.$$

for all $\delta \in (0, \frac{1}{4})$, where $c_n := cn^{-\frac{1}{4}}$ and $a_n := an^{-1}$.

(b) *If f is three times differentiable at $x^\star$, Hessian H is continuous around $x^\star$, and $\lambda_{\min} > \frac{1}{3a}$ holds, we get*

$$n^\delta \|Z_n - x^\star\| \to 0 \quad a.s.$$

for all $\delta \in (0, \frac{1}{3})$, where $c_n := cn^{-\frac{1}{6}}$ and $a_n := an^{-1}$.

Proof of Corollary 2.1

If we use the continuous processes $R_s := s$ and $M_i(\mathrm{d}s, x) := \sum_{j=1}^{d} \sigma_s^{ij}(x)\,\mathrm{d}W_s^j$ in (5), the solution of (5) is also continuous and we get (29). Hence it is sufficient to verify the assumptions needed in Theorem 2.1 with respect to these special processes R_s and $M_i(\mathrm{d}s, x)$. Assumption (F) directly follows from the continuity of R_s. By

$$\lceil \int_0^\cdot M_i(\mathrm{d}s, J_s) \rceil_t = \lceil \int_0^\cdot \sum_{j=1}^{d} \sigma_s^{ij}(J_s)\,\mathrm{d}W_s^j \rceil_t = \sum_{j,l=1}^{d} \lceil \int_0^\cdot \sigma_s^{ij}(J_s)\,\mathrm{d}W_s^j, \int_0^\cdot \sigma_s^{il}(J_s)\,\mathrm{d}W_s^l \rceil_t$$

$$
\begin{aligned}
&= \int_0^t \sum_{j,l=1}^{d} \sigma_s^{ij}(J_s)\sigma_s^{il}(J_s)\,\mathrm{d}\lceil W^j, W^l\rceil_s = \int_0^t \sum_{j,l=1}^{d} \sigma_s^{ij}(J_s)\sigma_s^{il}(J_s)\mathbb{1}_{\{j=l\}}\,\mathrm{d}s \\
&= \int_0^t \sum_{j=1}^{d} \sigma_s^{ij}(J_s)^2\,\mathrm{d}s
\end{aligned}
$$

we obtain $h_s^{ii}(\cdot) = \sum_{j=1}^d \sigma_s^{ij}(\cdot)^2$, and therefore validity of (E) follows from (E'). □

Proof of Corollary 2.2

We extend the sequence V_n to the process $\tilde{V}_t$ continuous in time by

$$
\tilde{V}_t := \begin{cases} V_1 & \text{for} \quad t = 0 \\ V_n & \text{for} \quad n-1 < t \le n \qquad \text{and} \qquad n \in \mathbb{N} \end{cases}
$$

and introduce the following processes:

$$
R_s := \max_{n\in\mathbb{N}_0, n\le s}(n) = \lfloor s \rfloor \quad \text{for} \quad s \ge 0 \qquad\qquad \text{and} \qquad\qquad M(\mathrm{d}s, x) := \tilde{V}_s\,\mathrm{d}R_s.
$$

With these definitions we have

$$
\int_0^t M_i(\mathrm{d}s, x) = \int_0^t \tilde{V}_s^i\,\mathrm{d}R_s = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \tilde{V}_n^i(\Delta R_n) = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \tilde{V}_n^i = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} V_n^i =: H_t^i.
$$

First we show that $\int_0^t M(\mathrm{d}s, x)$ is a martingale with respect to $\tilde{\mathcal{F}}_t := \mathcal{F}_{R_t}$, $t \ge 0$.

$$
\begin{aligned}
\mathbf{E}\left(H_t \,\middle|\, \tilde{\mathcal{F}}_s\right) &= \mathbf{E}\left(H_t \,\middle|\, \mathcal{F}_{\lfloor s\rfloor}\right) = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n \,\middle|\, \mathcal{F}_{\lfloor s\rfloor}\right) \\
&= \sum_{\substack{n\le \lfloor s\rfloor\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n \,\middle|\, \mathcal{F}_{\lfloor s\rfloor}\right) + \sum_{\substack{\lfloor s\rfloor < n\le t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n \,\middle|\, \mathcal{F}_{\lfloor s\rfloor}\right) = \sum_{\substack{n\le \lfloor s\rfloor\\ n\in\mathbb{N}}} V_n + 0 = \sum_{\substack{n\le s\\ n\in\mathbb{N}}} V_n \\
&= H_s
\end{aligned}
$$

With (5) and the processes above defined we get for $n \in \mathbb{N}$

$$
\begin{aligned}
Z_n^i - Z_0^i &= -\int_0^n \frac{a_s}{2c_s}\left(f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i)\right)\mathrm{d}R_s - \int_0^n \frac{a_s}{2c_s} M_i(\mathrm{d}s, Z_{s-}) \\
&= -\sum_{j=1}^{n} \frac{a_j}{2c_j}\left(f(Z_{j-1} + c_j e_i) - f(Z_{j-1} - c_j e_i)\right)(\Delta R_j) - \sum_{j=1}^{n} \frac{a_j}{2c_j} V_j^i(\Delta R_j) \\
&= -\sum_{j=1}^{n} \frac{a_j}{2c_j}\left(f(Z_{j-1} + c_j e_i) - f(Z_{j-1} - c_j e_i)\right) - \sum_{j=1}^{n} \frac{a_j}{2c_j} V_j^i.
\end{aligned}
$$

The recursive scheme (30) can be written in the form

$$Z_n^i - Z_0^i = \sum_{j=1}^n (Z_j^i - Z_{j-1}^i) = -\sum_{j=1}^n \frac{a_j}{2c_j}\left(f(Z_{j-1}+c_je_i) - f(Z_{j-1}-c_je_i)\right) - \sum_{j=1}^n \frac{a_j}{2c_j} V_j^i.$$

Therefore it suffices to verify the assumptions needed in Theorem 2.1. The assumptions (D) and (F) can be verified in a similar way to

$$\int_0^\infty a_s \,\mathrm{d}R_s = \sum_{j=1}^\infty a_j(\Delta R_j) = \sum_{j=1}^\infty a_j < \infty.$$

Using

$$\lceil \int_0^\cdot M_i(\mathrm{d}s, x)\rceil_t = \lceil \int_0^\cdot \tilde{V}_s^i \,\mathrm{d}R_s \rceil_t = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2}(\Delta R_n)^2 \,|\mathcal{F}_{n-1}\right) = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2}\,|\mathcal{F}_{n-1}\right),$$

it remains to verify

$$\int_0^\infty \frac{{a_s}^2}{{c_s}^2}\frac{h_s^{ii}(Z_{s-})}{1+\|Z_{s-}\|^2}\,\mathrm{d}R_s \le \sum_{n\in\mathbb{N}} \frac{{a_n}^2}{{c_n}^2}\mathbf{E}\left(V_n^{i^2}|\mathcal{F}_{n-1}\right) < \infty.$$

Since all summands and thus the sum itself are positive, it is sufficient to show that the expectation

$$\mathbf{E}\sum_{n\in\mathbb{N}} \frac{{a_n}^2}{{c_n}^2}\mathbf{E}\left(V_n^{i^2}|\mathcal{F}_{n-1}\right) \overset{(\star)}{=} \sum_{n\in\mathbb{N}} \frac{{a_n}^2}{{c_n}^2}\mathbf{E}|V_n^i|^2 \le \left(\sup_{n\in\mathbb{N}} \mathbf{E}\|V_n\|^2\right) \sum_{n\in\mathbb{N}} \frac{{a_n}^2}{{c_n}^2} < \infty$$

is finite, where $(\star)$ follows from the monotone convergence theorem. □

Proof of Corollary 2.3

We verify only part (a), since part (b) follows analogously. First we define some processes needed later to apply Theorem 2.2. The continuous processes $R_s := s$ and $M_i(\mathrm{d}s, x) := \sum_{j=1}^d \sigma_s^{ij}(x)\,\mathrm{d}W^j(s)$ are defined as in the proof of Corollary 2.1. It is obvious that (D') holds. Since R_s is continuous, $\gamma_t(\delta)$ is also continuous. We show that $\gamma_t(\delta)$, which determines the speed of convergence, is given by

$$\gamma_t(\delta) = \mathcal{E}_t(\delta\int_0^\cdot a_s\,\mathrm{d}R_s) = \exp\left(\delta a\int_0^t \frac{1}{1+s}\,\mathrm{d}s\right) = \exp\left(\delta a \ln(1+t)\right) = (1+t)^{a\delta}.$$

To obtain the desired result, we have to verify that for all $\delta \in (0, \frac{1}{4a})$ assumptions $(\tilde{E})$ and (D^2) hold. Due to the assumption on $\lambda_{\min}$ we have $0 \le \delta < \lambda_{\min}$. Assumption (D^2) follows from

$$\int_0^\infty \gamma_{s-}(\delta)a_sc_s\,\mathrm{d}R_s = ac\int_0^\infty (1+s)^{a\delta-\frac{5}{4}}\,\mathrm{d}s \le ac\int_0^\infty (1+s)^{\frac{1}{4}-\epsilon-\frac{5}{4}}\,\mathrm{d}s < \infty.$$

Finally, assumption $(\tilde{E})$ follows from

$$\int_0^\infty \frac{\gamma_{s-}^2(\delta)h_s^{ii}(Z_{s-})}{1+\gamma_{s-}^2(\delta)\|Z_{s-}\|^2}\frac{a_s^2}{c_s^2}\,\mathrm{d}R_s = \sum_{j=1}^d \int_0^\infty \frac{\gamma_s^2(\delta)\sigma_s^{ij}(Z_s)^2}{1+\gamma_s^2(\delta)\|Z_s\|^2}\frac{a_s^2}{c_s^2}\,\mathrm{d}s$$
$$\leq \mathcal{C}\int_0^\infty \frac{\gamma_s^2(\delta)(1+\|Z_s\|)^2}{1+\gamma_s^2(\delta)\|Z_s\|^2}\frac{a_s^2}{c_s^2}\,\mathrm{d}s \leq \mathcal{C}\int_0^\infty \gamma_s^2(\delta)\frac{a_s^2}{c_s^2}\,\mathrm{d}s$$
$$\leq \mathcal{C}\int_0^\infty (1+s)^{\frac{1}{2}-\epsilon-\frac{3}{2}}\,\mathrm{d}s \leq \mathcal{C}\int_0^\infty (1+s)^{-1-\epsilon}\,\mathrm{d}s < \infty.$$

□

Proof of Corollary 2.4
As in the proof of Corollary 2.3, we verify only part (a), since part (b) follows analogously. We define the processes $\tilde{V}_t$, R_s, and $M(\mathrm{d}s,x)$ as in the proof of Corollary 2.2. Now we apply Theorem 2.2. From the choices $R_s := \lfloor s \rfloor$ and $a_s := as^{-1}$ we obtain by Taylor expansion

$$\gamma_t(\delta) = \mathcal{E}_t(a\delta\int_0^\cdot \frac{1}{s}\,\mathrm{d}R_s) = \prod_{i=1}^{\lfloor t \rfloor}(1+\frac{a\delta}{i}) = \exp\,(\sum_{i=1}^{\lfloor t \rfloor}\ln\,(1+\frac{a\delta}{i})) = C_{\lfloor t \rfloor}\exp\,(a\delta\sum_{i=1}^{\lfloor t \rfloor}\frac{1}{i})$$

where $C_{\lfloor t \rfloor} \to C$. Furthermore, we get the following two inequalities

$$\exp\,(a\delta\sum_{i=1}^{\lfloor t \rfloor}\frac{1}{i}) \geq \exp\,(a\delta\int_1^{\lfloor t \rfloor}\frac{1}{x}\,\mathrm{d}x) = \lfloor t \rfloor^{a\delta}$$

and

$$\exp\,(a\delta\sum_{i=1}^{\lfloor t \rfloor}\frac{1}{i}) = \exp\,(a\delta + a\delta\sum_{i=2}^{\lfloor t \rfloor}\frac{1}{i}) \leq \exp\,(a\delta + a\delta\int_1^{\lfloor t \rfloor}\frac{1}{x}\,\mathrm{d}x) = \exp\,(a\delta)\lfloor t \rfloor^{a\delta}.$$

Hence, it is sufficient to verify the assumptions of Theorem 2.2 for $\lfloor t \rfloor^{a\delta}$ instead of $\gamma_{t-}(\delta)$. To obtain the desired result, we have to check that, for all $\delta \in (0, \frac{1}{4a})$, the assumptions $(\tilde{E})$ and (D^2) are satisfied. Due to the assumption on $\lambda_{\min}$ we have $0 \leq \delta < \lambda_{\min}$. The assumption (D^2) follows from

$$\int_0^\infty \lfloor s \rfloor^{a\delta}a_s c_s\,\mathrm{d}R_s \leq ac\sum_{n\in\mathbb{N}} n^{a\delta}n^{-\frac{5}{4}} = \mathcal{C}\sum_{n\in\mathbb{N}} n^{\frac{1}{4}-\epsilon-\frac{5}{4}} < \infty.$$

Finally the validation of assumption $(\tilde{E})$ is missing. The processes $\lfloor s \rfloor^{2a\delta}\frac{a_s^2}{c_s^2}h_s^{ii}(Z_{s-})$ and ΔR_s itself, which appear in the assumption, are always positive. Hence, the assumption $(\tilde{E})$ follows from $\mathbf{E}\int_0^\infty h_s^{ii}(Z_{s-})\lfloor s \rfloor^{2a\delta}\frac{a_s^2}{c_s^2}\,\mathrm{d}R_s < \infty$. Therefore we show

$$\mathbf{E}\int_0^\infty h_s^{ii}(Z_{s-})\lfloor s \rfloor^{2a\delta}\frac{a_s^2}{c_s^2}\,\mathrm{d}R_s \leq \mathcal{C}\,\mathbf{E}\sum_{n\in\mathbb{N}}\mathbf{E}\left(V_n^{i2}|\mathcal{F}_{n-1}\right)n^{2a\delta}n^{-\frac{3}{2}}$$

$$\overset{(\alpha)}{\leq} \mathcal{C} \left(\sup_{n \in \mathbb{N}} \mathbf{E} \|V_n\|^2 \right) \sum_{n \in \mathbb{N}} n^{\frac{1}{2} - \epsilon - \frac{3}{2}} \leq \mathcal{C} \sum_{n \in \mathbb{N}} n^{-1-\epsilon} < \infty$$

where we get $h_s^{ii}(\cdot)$ as in the proof of Corollary 2.2 and (α) using the monotone convergence theorem. □

3 Asymptotic Properties of the Process Using Special Weights

This chapter is devoted to asymptotic normality which is of major interest since we get the speed of convergence in a special sense. Furthermore, the knowledge of the limit distribution is very useful. Indeed, if the limit distribution is known, it can be used, for example, to identify optimal design parameters. Here, the design parameters are, for example, a and c. All theorems presented in this chapter assume that the process R_t is deterministic but there are also some lemmata which do not need this assumption. The theorem dealing with the asymptotic normality is presented in Section 3.2.

3.1 Almost L^2-Convergence Rate

In this section, we establish a result on an almost L^2-convergence rate. To this end we say that a process Z_t converges in the almost L^2-convergence sense, if, for any $\epsilon > 0$, $Z_t \mathbb{1}_{A_\epsilon}$ converges in L^2 for a suitable set A_ϵ of probability $\geq 1 - \epsilon$.

In the next section we will see that a theorem on an almost L^2-convergence rate is a useful tool to prove asymptotic normality. Here we make the restrictions that the process R_t is deterministic and a_s, c_s are chosen as

$$a_s := \frac{a}{(1 + R_{s-})^\alpha} \qquad \text{and} \qquad c_s := \frac{c}{(1 + R_{s-})^\gamma}$$

for some fixed $a, c > 0$ and $0 < \alpha, \gamma \leq 1$. In this setting, which is a special case of (5), the following theorem presents an almost L^2-convergence rate.

Theorem 3.1. *We assume the existence of a positive, deterministic, monotonously increasing process $(S_s)_{s\geq 0}$ with $S_s \uparrow \infty$ and $S_s \|Z_s\| \to 0$. Let conditions (A), (B), and (D)–(F) be valid. Assume that f is two or three times differentiable at $x^\star$ with a continuous Hessian around $x^\star$ and*

$$(E^\star) \qquad \begin{cases} \underset{i,j\in\{1,\dots,d\}}{\forall} \ \underset{0<S<\infty}{\forall} \ \underset{0<K<\infty}{\exists} & \|x\| \leq S \Rightarrow |h_t^{ij}(x)| \leq K \\ & \int_0^\infty a_s^2\, c_s^{-2}\, \mathrm{d}R_s < \infty. \end{cases}$$

In the case $\alpha < 1$, we assume that the Hessian is positive definite at $x^\star$, and in the special, yet important, case $\alpha = 1$ we further stipulate $\lambda_{\min} > \frac{1-2\gamma}{2a}$. Then, for all $\epsilon > 0$, there exists a process Y_t such that

$$\mathbf{P}\left[\underset{t\geq 0}{\forall}\ Y_t = Z_t\right] \geq 1 - \epsilon$$

and

$$\mathbf{E}\|Y_t - x^\star\|^2 = \mathcal{O}((1+R_t)^\beta)$$

with

$$\beta := \begin{cases} \max\{1-\alpha-2\gamma, 1-2\alpha+2\gamma\}, & \textit{if } f \textit{ is } 2\times \textit{ differentiable at } x^\star \\ \max\{1-\alpha-4\gamma, 1-2\alpha+2\gamma\}, & \textit{if } f \textit{ is } 3\times \textit{ differentiable at } x^\star. \end{cases}$$

Remark 3.1. The assumption $(E^\star)$ is no major restriction. The second part holds, for example, if $h_t(x)$ is bounded in both arguments or if $h_t^{ij}(x) \leq C(1+\|x\|)$ holds.

Remark 3.2. Considering the assumptions in Theorem 3.1, we see that the process with standard weights, that is $\alpha = 1$, leads to a restriction on the smallest eigenvalue. This restriction on the smallest eigenvalue is $\lambda_{\min} > \frac{1-2\gamma}{2a}$ and appears also when we treat asymptotic normality using standard weights. Observing the proof, we easily see why the assumption disappears in the case $\alpha < 1$. In this case one process appearing in the insight of the bracket term of (37) is $(1+R_t)^{\alpha-1}$, converges to zero and therefore the restriction on the eigenvalue turns out to be superfluous.

Remark 3.3. The theorem above is of particular interest when we consider convergence in probability. Indeed, in this case the processes Z_t and Y_t are interchangeable.

Lemma 3.1. *If there is a strictly positive, monotone increasing process $(S_t)_{t\geq 0}$ that satisfies $S_t \uparrow \infty$ and $S_t\|Z_t\| \to 0$, then, for all $\epsilon, \delta > 0$, a deterministic time $T(\epsilon,\delta)$ exists with*

$$\mathbf{P}\left[\sup_{t\geq T(\epsilon,\delta)} \|Z_t\| > \delta\right] < \epsilon.$$

Proof of Lemma 3.1

We write $T := T(\epsilon,\delta)$ to shorten notations. Since S_t is monotonously increasing and strictly positive, we obtain

$$\sup_{t\geq T}\|Z_t\| = \sup_{t\geq T}\left(S_t^{-1}S_t\|Z_t\|\right) \leq S_T^{-1}\underbrace{\sup_t\left(S_t\|Z_t\|\right)}_{\overset{(\star)}{\leq} C(\omega) < \infty}.$$

Because we assume that Z_s is a strong solution of the stochastic integral equation on $[0,\infty)$, we know that there are no explosion times. We further know from the assumptions that $S_t\|Z_t\|$ converges to 0. If we combine these two statements we get $(\star)$. Therefore, we have

$$\mathbf{P}\left[\sup_{t\geq T}\|Z_t\| > \delta\right] \leq \mathbf{P}\left[C(\omega) > \delta \underbrace{S_T}_{\uparrow\infty}\right]$$

and thus it suffices to verify

$$\underset{T\in\mathbb{R}_+}{\exists} \qquad \mathbf{P}\left[C(\omega) > \delta S_T\right] < \epsilon \qquad\qquad \text{(T independent of } \omega)$$

to prove the lemma. This will be proved by contradiction. We assume

$$\underset{T\in\mathbb{R}_+}{\forall} \qquad \mathbf{P}\left[C(\omega) > \delta S_T\right] \geq \epsilon.$$

Since the above holds for all $T \in \mathbb{R}_+$, it holds in particular for all $T \in \mathbb{N}$. Moreover, $\mathbf{P}$ as a probability measure, is continuous, and it follows

$$\mathbf{P}\left(\bigcap_{n=1}^{\infty}\left[C(\omega) > \delta S_n\right]\right) = \lim_{N\to\infty}\mathbf{P}\left(\bigcap_{n=1}^{N}\left[C(\omega) > \delta S_n\right]\right) = \lim_{N\to\infty}\underbrace{\mathbf{P}\left[C(\omega) > \delta S_N\right]}_{\geq\,\epsilon\quad\underset{N\in\mathbb{N}}{\forall}} \geq \epsilon.$$

Thus we have

$$\bigcap_{n=1}^{\infty}\left[C(\omega) > \delta S_n\right] = \left[C(\omega) = \infty\right]$$

and in particular

$$\mathbf{P}\left[C(\omega) = \infty\right] \geq \epsilon,$$

which contradicts the fact that C is almost surely bounded (i.e. $C(\omega) < \infty$ almost surely). □

Proof of Theorem 3.1
Without loss of generality we assume $x^\star = 0$. We will prove the assertion of the theorem by constructing the process Y_t and verifying its L^2-convergence rate. However, before we construct the process, we need some approximations and notations. We know that f is differentiable on $\mathbb{R}^d$ and at least two times differentiable at $x^\star = 0$. This ensures the existence of $\nabla f(x)$ and H_0. Furthermore we get for $i \in \{1,\dots,d\}$

$$\frac{1}{2c}\{f(x+ce_i) - f(x-ce_i)\} = (H_0x)_i + \left(\underbrace{\frac{1}{2}\int_{-1}^{1}\nabla_i f(x+tce_i)\,\mathrm{d}t - (H_0x)_i}_{=:\,V^i(x,c)}\right).$$

Note that for an f two or three times differentiable at $x^\star = 0$, the gradient is one or two times differentiable at $x^\star = 0$. For a fixed $c > 0$ we obtain on the one hand,

$$V^i(x,c) = \frac{1}{2}\int_{-1}^{1} \underbrace{\nabla_i f(x + tce_i)}\,\mathrm{d}t - (H_0 x)_i = \mathrm{o}(\|x\| + c) = \mathrm{o}(\|x\|) + \mathrm{o}(c) \tag{31}$$
$$= \nabla_i f(0) + \langle H_0^{\cdot i}, x + tce_i\rangle + \mathrm{o}(\|x + tce_i\|)$$

and on the other hand

$$V^i(x,c) = \frac{1}{2}\int_{-1}^{1} \underbrace{\nabla_i f(x + tce_i)}\,\mathrm{d}t - (H_0 x)_i = \mathcal{O}(\|x\|^2) + \mathcal{O}(c^2) + \mathrm{o}(\|x + c\|^2)$$
$$= \nabla_i f(0) + \langle H_0^{\cdot i}, x + tce_i\rangle + \mathrm{o}(\|x + tce_i\|^2)$$
$$+ \frac{1}{2}\sum_{l,k} \frac{\partial^3 f}{\partial x_i \partial x_l \partial x_k}(0)(x + tce_i)_l (x + tce_i)_k$$
$$= \mathcal{O}(\|x\|^2) + \mathcal{O}(c^2) + \mathrm{o}(\|x\|^2) + \mathrm{o}(c^2) = \mathcal{O}(\|x\|^2) + \mathcal{O}(c^2). \tag{32}$$

By combining (31) and (32), we have

$$\frac{1}{2c}\{\cdot\}. = \begin{cases} H_0 x + V(x,c) & \text{with } V(x,c) = \mathrm{o}(\|x\|) + \mathrm{o}(c) \quad \text{, if } f \text{ is } 2\times \text{ diff. at } x^\star = 0 \\ H_0 x + V(x,c) & \text{with } V(x,c) = \mathcal{O}(\|x\|^2) + \mathcal{O}(c^2)\text{, if } f \text{ is } 3\times \text{ diff. at } x^\star = 0. \end{cases}$$

This describes the behavior of $V(x,c)$ for small x and c. In the following we consider a process very similar to the process $(1 + R_t)^{p^\star}\mathbf{E}\|Z_t\|^2$. To this end we need some further definitions. We distinguish between the cases $\alpha = 1$ and $\alpha < 1$. In the case $\alpha = 1$ we have $2a\lambda_{\min} > 1 - 2\gamma$, thus there exists a strictly positive $p^\star < 1$ with $2a\lambda_{\min} > p^\star > 1 - 2\gamma$. In the case $\alpha < 1$ there exists a strictly positive $p^\star < 1$ with $p^\star > 1 - 2\gamma$.

(a) Case: f is two times differentiable at $x^\star = 0$. We choose

$$\kappa := \begin{cases} \frac{2a\lambda_{\min} - p^\star}{6a} & \text{if } \alpha = 1 \\ \frac{2}{9}\lambda_{\min} & \text{if } \alpha < 1 \end{cases}, \qquad \text{thus we have } \kappa > 0.$$

Using $V(x,c) = \mathrm{o}(\|x\|) + \mathrm{o}(c)$, we get

$$\underset{\rho>0}{\forall}\ \underset{\delta_1>0}{\exists}\ \underset{\delta_2>0}{\exists}\ \underset{\|x\|<\delta_1}{\forall}\ \underset{c<\delta_2}{\forall} \quad \|V(x,c)\| \le \rho\|x\| + \rho c.$$

Now we choose $\rho = \kappa$. This is possible because $\kappa > 0$ holds. Since $\epsilon > 0$ is fixed and we choose $\delta := \min\{\delta_1, 1\} > 0$, the statement of Lemma 3.1 guarantees the existence of a deterministic (!) time $T(\epsilon,\delta) < \infty$ with

$$\mathbf{P}\left[\sup_{t \ge T(\epsilon,\delta)} \|Z_t\| \le \delta\right] \ge 1 - \epsilon.$$

Set $a_s := a(1 + R_{s-})^{-\alpha}$. Since $(R_s)_{s\geq 0}$ is a deterministic process and

$$\int_0^\infty a_s \, \mathrm{d}R_s = \infty \quad \wedge \quad \int_0^\infty a_s^2 \Delta R_s \, \mathrm{d}R_s < \infty \qquad \Longrightarrow \qquad \frac{\Delta R_s}{(1 + R_{s-})^\alpha} \to 0,$$

we get for $\alpha = 1$

$$\underset{s_0}{\exists} \ \underset{s\geq s_0}{\forall} \quad \left(a_s \Delta R_s \leq \frac{2a\lambda_{\min} - p^\star}{4a(\lambda_{\max} + \kappa)^2} \quad \wedge \quad c_s \leq \delta_2 \right)$$

and for $\alpha < 1$

$$\underset{s_1}{\exists} \ \underset{s\geq s_1}{\forall} \quad \left(a_s \Delta R_s \leq \frac{\lambda_{\min}}{3(\lambda_{\max} + \kappa)^2} \quad \wedge \quad (1 + R_{s-})^{\alpha-1} \leq \frac{2a\lambda_{\min}}{3p^\star} \quad \wedge \quad c_s \leq \delta_2 \right).$$

Recall that the times $T(\epsilon, \delta)$, s_0, and s_1 are all deterministic. Consequently

$$T := \begin{cases} \max\{T(\epsilon, \delta), s_0\} & \text{if } \alpha = 1 \\ \max\{T(\epsilon, \delta), s_1\} & \text{if } \alpha < 1 \end{cases}$$

is also deterministic and hence a stopping time. Then

$$D := \inf\{t > T \ : \ \|Z_t\| > \delta\}$$

is also a stopping time. Now we define the process $(Y_t)_{t\geq 0}$ by

$$\begin{aligned} Y_t : &= Z_t \mathbb{1}_{[0,T)}(t) + Z_T^D \mathbb{1}_{[T,\infty)}(t) \mathbb{1}_{[T\neq D]} + \int_{T+}^{t\wedge D} \mathrm{d}Z_s \\ &= Z_t \mathbb{1}_{[0,T)}(t) + Z_T \mathbb{1}_{[T,\infty)}(t) \mathbb{1}_{[T\neq D]} - \int_{T+}^{t} \frac{a_s}{2c_s} \{\cdot\}.\mathbb{1}_{(T,D]}(s) \, \mathrm{d}R_s \\ &\quad - \int_{T+}^{t} \frac{a_s}{2c_s} \mathbb{1}_{(T,D]}(s) \, M(\mathrm{d}s, Z_{s-}). \end{aligned}$$

If we consider the process Y_t on the set $[D = \infty]$

$$Z_t \mathbb{1}_{[0,T)}(t) + \mathbb{1}_{[T,\infty)}(t) \left(Z_T + \int_{T+}^{t} \mathrm{d}Z_s \right) = Z_t \mathbb{1}_{[0,T)}(t) + \mathbb{1}_{[T,\infty)}(t) Z_t = Z_t$$

we see that the process Y_t may differ from the process Z_t only on the set $[D < \infty]$. But the measurement of such a set is at most ϵ, because we have

$$\mathbf{P}\left[Z_t = Y_t \quad \underset{t\geq 0}{\forall} \right] \geq \mathbf{P}[D = \infty] = \mathbf{P}\left[\sup_{t\geq T} \|Z_t\| \leq \delta \right] \geq 1 - \epsilon.$$

In the following, we assume that $t > T$ holds. Using (12) and (13) we obtain

$$\|Y_t\|^2 = \|Y_T\|^2 + \int_{T+}^{t} \mathrm{d}\|Y_s\|^2$$

$$= \left\| Z_T^D 1_{[T\neq D]} \right\|^2 + M_t^\star + \sum_{i=1}^d \Bigg(\int_{T+}^t \frac{{a_s}^2}{4{c_s}^2} h_s^{ii}(Z_{s-}) \mathbb{1}_{(T,D]}(s)\, \mathrm{d}R_s$$

$$-2 \int_{T+}^t a_s \frac{\{\cdot\}_i}{2c_s} Z_{s-}^i \mathbb{1}_{(T,D]}(s)\, \mathrm{d}R_s + \int_{T+}^t a_s^2 \frac{\{\cdot\}_i^2}{4{c_s}^2} \mathbb{1}_{(T,D]}(s) \Delta R_s\, \mathrm{d}R_s^d \Bigg)$$

with the local martingale

$$M_t^\star = \sum_{i=1}^d \Bigg(\int_{T+}^t \frac{{a_s}^2}{2{c_s}^2} \{\cdot\}_i \Delta R_s \mathbb{1}_{(T,D]}(s)\, M_i^d(\mathrm{d}s, Z_{s-}) - \int_{T+}^t \frac{a_s}{c_s} Z_{s-}^i \mathbb{1}_{(T,D]}(s)\, M_i(\mathrm{d}s, Z_{s-})$$

$$+ \int_{T+}^t \frac{{a_s}^2}{4{c_s}^2} \mathbb{1}_{(T,D]}(s) \ \Big(\ [M_i(\mathrm{d}r, Z_{r-})]_s - \lceil M_i(\mathrm{d}r, Z_{r-}) \rceil_s \ \Big) \Bigg).$$

Furthermore, we have

$$\begin{aligned}
\mathbb{1}_{(T,D]}(t) Y_t &= Z_T^D \mathbb{1}_{(T,D]}(t) \mathbb{1}_{[T\neq D]} + \mathbb{1}_{(T,D]}(t) \int_{T+}^{t\wedge D} \mathrm{d}Z_s \\
&= Z_T^D \mathbb{1}_{(T,D]}(t) + \mathbb{1}_{(T,D]}(t) \int_{T+}^{t\wedge D} \mathrm{d}Z_s = \mathbb{1}_{(T,D]}(t) \left(Z_T^D + \int_{T+}^t \mathrm{d}Z_s^D \right) \\
&= \mathbb{1}_{(T,D]}(t) Z_t^D = \mathbb{1}_{(T,D]}(t) Z_t
\end{aligned}$$

and

$$\mathbb{1}_{(T,D]}(t) \|Y_{t-}\| = \mathbb{1}_{(T,D]}(t) \|Z_{t-}\| \leq \delta \mathbb{1}_{(T,D]}(t).$$

Now we need some inequalities to handle the terms $\frac{1}{2c_s} \sum_{i=1}^d Z_{s-}^i \{\cdot\}_i$ and $\frac{1}{4{c_s}^2} \sum_{i=1}^d \{\cdot\}_i^2$. To obtain these inequalities we use the asymptotic results that we derived at the beginning of this proof. Let s be a time such that $s > T$. Then we observe

$$\frac{1}{2c_s} \{\cdot\}_i = \sum_{j=1}^d H_0^{ij} Z_{s-}^j + V^i(Z_{s-}, c_s) \tag{33}$$

and obtain both

$$\begin{aligned}
-\frac{1}{2c_s} \sum_{i=1}^d Z_{s-}^i \{\cdot\}_i &= -\sum_{i=1}^d \left(\sum_{j=1}^d H_0^{ij} Z_{s-}^j + V^i(Z_{s-}, c_s) \right) Z_{s-}^i \\
&\leq -Z_{s-}^T H_0 Z_{s-} + |\langle V(Z_{s-}, c_s), Z_{s-} \rangle| \\
&\leq -\lambda_{\min} \|Z_{s-}\|^2 + \|V(Z_{s-}, c_s)\| \|Z_{s-}\| \\
&\leq -\lambda_{\min} \|Z_{s-}\|^2 + \frac{3}{2}\kappa \|Z_{s-}\|^2 + \frac{1}{2}\kappa c_s^2
\end{aligned} \tag{34}$$

and

$$\frac{1}{4{c_s}^2} \sum_{i=1}^d \{\cdot\}_i^2 = \langle \frac{1}{2c_s}\{\cdot\}_\cdot, \frac{1}{2c_s}\{\cdot\}_\cdot \rangle = \left\| \ \frac{1}{2c_s}\{\cdot\}_\cdot \ \right\|^2 \leq \Big(\|H_0 Z_{s-}\| + \|V(Z_{s-}, c_s)\| \Big)^2$$

$$\leq 2\left(\lambda_{\max}+\kappa\right)^2\|Z_{s-}\|^2+2\kappa^2c_s^2. \tag{35}$$

Now we show that the expectation of the local martingale $(M_t^\star)_{t\geq T}$ is zero. To this end we can use the fact that the expectation of a martingale starting at zero is zero (which, of course, is not generally the case for a local martingale). To show that a local martingale M_t is even a martingale, it suffices to verify $\mathbf{E}\lceil M\rceil_t<\infty$ for all t. The second local martingale in $M_t^\star$ gives

$$\begin{aligned}
&\sup_{t>T}\mathbf{E}\left\lceil\sum_{i=1}^d\int_{T+}^{\cdot}\frac{a_s}{2c_s}Z^i_{s-}\mathbb{1}_{(T,D]}(s)\,M_i(\mathrm{d}s,Z_{s-})\right\rceil_t\\
&\quad=\sup_{t>T}\mathbf{E}\sum_{i,j=1}^d\int_{T+}^t\frac{a_s^2}{4c_s^2}Z^i_{s-}Z^j_{s-}\mathbb{1}_{(T,D]}(s)\;\lceil M_i(\mathrm{d}r,Z_{r-}),M_j(\mathrm{d}r,Z_{r-})\rceil_s\\
&\quad=\sum_{i,j=1}^d\sup_{t>T}\mathbf{E}\int_{T+}^t\frac{a_s^2}{4c_s^2}Z^i_{s-}Z^j_{s-}\mathbb{1}_{(T,D]}(s)h_s^{ij}(Z_{s-})\,\mathrm{d}R_s\leq\mathcal{C}\int_0^\infty\frac{a_s^2}{c_s^2}\,\mathrm{d}R_s<\infty
\end{aligned}$$

hence it is a martingale. In view of the first local martingale we may write $M_i(\mathrm{d}s,Z_{s-})$ instead of $M_i^d(\mathrm{d}s,Z_{s-})$, because the process ΔR_s appears as an integrand.

$$\begin{aligned}
\frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i\Delta R_s\,M_i(\mathrm{d}s,Z_{s-})&=\frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i\Delta R_s\,M_i^c(\mathrm{d}s,Z_{s-})+\frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i\Delta R_s\,M_i^d(\mathrm{d}s,Z_{s-})\\
&=\frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i\Delta R_s\,M_i^d(\mathrm{d}s,Z_{s-})
\end{aligned}$$

The first local martingale in $M_t^\star$ gives

$$\begin{aligned}
&\sup_{t>T}\mathbf{E}\left\lceil\sum_{i=1}^d\int_{T+}^{\cdot}\frac{{a_s}^2}{4{c_s}^2}\{\cdot\}_i\mathbb{1}_{(T,D]}(s)\Delta R_s\,M_i(\mathrm{d}s,Z_{s-})\right\rceil_t\\
&\quad=\sup_{t>T}\mathbf{E}\sum_{i,j=1}^d\int_{T+}^t\frac{a_s^4}{16c_s^4}\{\cdot\}_i\{\cdot\}_j(\Delta R_s)^2h_s^{ij}(Z_{s-})\mathbb{1}_{(T,D]}(s)\,\mathrm{d}R_s\\
&\quad=\mathcal{C}\sup_{t>T}\mathbf{E}\int_{T+}^t\frac{a_s^4}{c_s^2}\left(\sum_{l=1}^d\frac{1}{2c_s}\{\cdot\}_l\right)^2\mathbb{1}_{(T,D]}(s)(\Delta R_s)^2\,\mathrm{d}R_s\\
&\quad\overset{(35)}{\leq}\mathcal{C}\int_0^\infty\frac{a_s^4}{c_s^2}(\Delta R_s)^2\,\mathrm{d}R_s\leq\mathcal{C}\int_0^\infty\frac{{a_s}^2}{{c_s}^2}(a_s\Delta R_s)^2\,\mathrm{d}R_s\\
&\quad=\mathcal{C}\int_0^\infty a_s^2c_s^{-2}\mathrm{o_b}(1)\,\mathrm{d}R_s<\mathcal{C}\int_0^\infty a_s^2c_s^{-2}\,\mathrm{d}R_s<\infty
\end{aligned}$$

and it follows that the local martingale is even a martingale. In order to show that the expectation of the last local martingale in $M_t^\star$ is zero, we use the following criterion: For a locally square integrable martingale M_t starting at zero we have $\mathbf{E}[M]_t=\mathbf{E}\lceil M\rceil_t$ for all t. For the last local martingale in $M_t^\star$ we have

$$\mathbf{E}\int_{T+}^t\frac{{a_r}^2}{4{c_r}^2}\mathbb{1}_{(T,D]}(r)\;\Big(\;[M_i(\mathrm{d}l,Z_{l-})]_r-\lceil M_i(\mathrm{d}l,Z_{l-})\rceil_r\;\Big)$$

$$= \mathbf{E}\Bigg[\underbrace{\int_{T+}^{\cdot} \frac{a_r}{2c_r} \mathbb{1}_{(T,D]}(r)\, M_i(\mathrm{d}r, Z_{r-})}_{\in \mathcal{M}^2_{\mathrm{loc}}}\Bigg]_t - \mathbf{E}\Bigg[\underbrace{\int_{T+}^{\cdot} \frac{a_r}{2c_r} \mathbb{1}_{(T,D]}(r)\, M_i(\mathrm{d}r, Z_{r-})}_{\in \mathcal{M}^2_{\mathrm{loc}}}\Bigg]_t$$
$$= 0.$$

Let us now turn to the examination of

$$\mathrm{d}\left((1+R_t)^{p^\star}\mathbf{E}\|Y_t\|^2\right) = (1+R_{t-})^{p^\star}\,\mathrm{d}\mathbf{E}\|Y_t\|^2 + \mathbf{E}\|Y_{t-}\|^2\,\mathrm{d}(1+R_t)^{p^\star} + \mathrm{d}[(1+R)^{p^\star}, \mathbf{E}\|Y\|^2]_t. \tag{36}$$

We seek an estimate of the first summand:

$$\begin{aligned}\mathbf{E}\,\|Y_t\|^2 = \mathbf{E}\left\|Z_T^D \mathbb{1}_{[T\neq D]}\right\|^2 &+ \int_{T+}^t a_s^2 \mathbf{E}\left(\sum_{i=1}^d \frac{\{\cdot\}_i^2}{4{c_s}^2}\mathbb{1}_{(T,D]}(s)\right)\Delta R_s\,\mathrm{d}R_s^d \\ &+ \int_{T+}^t \frac{{a_s}^2}{4{c_s}^2}\mathbf{E}\left(\sum_{i=1}^d h_s^{ii}(Z_{s-})\mathbb{1}_{(T,D]}(s)\right)\mathrm{d}R_s \\ &- 2\int_{T+}^t a_s \mathbf{E}\left(\sum_{i=1}^d \frac{\{\cdot\}_i}{2c_s} Z_{s-}^i \mathbb{1}_{(T,D]}(s)\right)\mathrm{d}R_s.\end{aligned}$$

Using (34) and (35) we get for the first summand in (36),

$$\begin{aligned}&\int_{T+}^t (1+R_{s-})^{p^\star}\mathrm{d}\mathbf{E}\|Y_s\|^2 \\ &\leq 2a\int_{T+}^t \left(-\lambda_{\min} + (\lambda_{\max}+\kappa)^2\frac{a\Delta R_s}{(1+R_{s-})^\alpha} + \frac{3\kappa}{2}\right)(1+R_{s-})^{p^\star-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s \\ &\quad + \mathcal{C}\int_{T+}^t (1+R_{s-})^{p^\star-\alpha-2\gamma}\,\mathrm{d}R_s + \mathcal{C}\int_{T+}^t (1+R_{s-})^{p^\star-2\alpha+2\gamma}\,\mathrm{d}R_s \\ &\leq 2a\int_{T+}^t \left(-\lambda_{\min} + v_s + \frac{3\kappa}{2}\right)(1+R_{s-})^{p^\star-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s \\ &\quad + \mathcal{C}\int_{T+}^t (1+R_{s-})^{p^\star+\beta-1}\mathrm{d}R_s,\end{aligned}$$

where $v_s := (\lambda_{\max}+\kappa)^2\frac{a\Delta R_s}{(1+R_{s-})^\alpha}$ and $\beta := \max\{1-\alpha-2\gamma, 1-2\alpha+2\gamma\}$. Using the standard Itô formula and a Taylor expansion around R_{s-}, we obtain

$$\begin{aligned}(1+R_t)^{p^\star} - 1 &= \int_{0+}^t p^\star(1+R_{s-})^{p^\star-1}\,\mathrm{d}R_s \\ &\quad + \sum_{0<s\leq t}\{(1+R_s)^{p^\star} - (1+R_{s-})^{p^\star} - p^\star(1+R_{s-})^{p^\star-1}\Delta R_s\} \\ &= \int_{0+}^t p^\star(1+R_{s-})^{p^\star-1}\,\mathrm{d}R_s\end{aligned}$$

$$+\sum_{0<s\le t}\Big\{\underbrace{\frac{1}{2}p^\star}_{\ge 0}\underbrace{(p^\star-1)}_{<0}\underbrace{(1+R_{s-}+\vartheta_s\Delta R_s)^{p^\star-2}(\Delta R_s)^2}_{\ge 0}\Big\}$$

$$\le \int_{0+}^t p^\star(1+R_{s-})^{p^\star-1}\,\mathrm{d}R_s$$

and for the second summand in (36)

$$\int_{T+}^t \mathbf{E}\|Y_{s-}\|^2\mathrm{d}(1+R_s)^{p^\star} \le p^\star\int_{T+}^t(1+R_{s-})^{p^\star-1}\mathbf{E}\|Y_{s-}\|^2\mathrm{d}R_s.$$

Since

$$\begin{aligned}0\le \Delta(1+R_t)^{p^\star} &= (1+R_t)^{p^\star}-(1+R_{t-})^{p^\star} = p^\star(1+R_{t-}+\vartheta_t\Delta R_t)^{p^\star-1}\Delta R_t\\ &\le p^\star(1+R_{t-})^{p^\star-1}\Delta R_t\end{aligned}$$

holds, we get for the last summand in (36)

$$\begin{aligned}\int_{T+}^t \mathrm{d}\left[(1+R)^{p^\star},\mathbf{E}\|Y\|^2\right]_s &= \int_{T+}^t \Delta(1+R_s)^{p^\star}\,\mathrm{d}\mathbf{E}\|Y_s\|^2\\ &\le 2a\int_{T+}^t\left(-\lambda_{\min}+v_s+\frac{3\kappa}{2}\right)(1+R_{s-})^{-\alpha}\mathbf{E}\|Y_{s-}\|^2\Delta(1+R_s)^{p^\star}\,\mathrm{d}R_s\\ &\quad+\mathcal{C}\int_{T+}^t(1+R_{s-})^{\beta-1}\Delta(1+R_s)^{p^\star}\,\mathrm{d}R_s\\ &\overset{(\star)}{\le}\mathcal{C}\int_{T+}^t(1+R_{s-})^{p^\star-1+\beta-1}\Delta R_s\,\mathrm{d}R_s\end{aligned}$$

where $(\star)$ is discussed below. Note that all terms appearing above, in particular v_s, are purely deterministic! Combining all estimates we have

$$\begin{aligned}&\int_{T+}^t \mathrm{d}\left((1+R_s)^{p^\star}\mathbf{E}\|Y_s\|^2\right)\\ &\qquad\le\mathcal{C}\int_{T+}^t\left(-\lambda_{\min}+v_s+\frac{3\kappa}{2}+\frac{p^\star}{2a}(1+R_{s-})^{\alpha-1}\right)(1+R_{s-})^{p^\star-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s\\ &\qquad\quad+\mathcal{C}\int_{T+}^t(1+R_{s-})^{p^\star+\beta-1}\mathrm{d}R_s \qquad (37)\\ &\qquad\overset{(\star)}{\le}\mathcal{C}\int_{T+}^t(1+R_{s-})^{p^\star+\beta-1}\mathrm{d}R_s\end{aligned}$$

where we used $\frac{\Delta R_s}{(1+R_s)^\alpha}=\mathrm{o_b}(1)$ and $(\star)$.

We get $(\star)$ in the case $\alpha=1$ by

$$-\lambda_{\min}+(\lambda_{\max}+\kappa)^2\,\frac{a\Delta R_s}{(1+R_{s-})}+\frac{3}{2}\kappa+\frac{p^\star}{2a}$$

$$\leq -\lambda_{\min} + (\lambda_{\max}+\kappa)^2 \frac{2a\lambda_{\min} - p^\star}{4a(\lambda_{\max}+\kappa)^2} + \frac{3}{2}\frac{2a\lambda_{\min} - p^\star}{6a} + \frac{p^\star}{2a} = 0$$

and in the case $\alpha < 1$ by

$$\begin{aligned} &-\lambda_{\min} + (\lambda_{\max}+\kappa)^2 \frac{a\Delta R_s}{(1+R_{s-})^\alpha} + \frac{3\kappa}{2} + \frac{p^\star}{2a}(1+R_{s-})^{\alpha-1} \\ &\qquad \leq -\lambda_{\min} + (\lambda_{\max}+\kappa)^2 \frac{\lambda_{\min}}{3(\lambda_{\max}+\kappa)^2} + \frac{3}{2}\frac{2\lambda_{\min}}{9} + \frac{p^\star}{2a}\frac{2a\lambda_{\min}}{3p^\star} = 0. \end{aligned}$$

Using the statement above and Itô's formula, we find

$$\begin{aligned} \int_{T+}^t \mathrm{d}\left((1+R_s)^{p^\star}\mathbf{E}\|Y_s\|^2\right) &\leq \mathcal{C}\int_{T+}^t (1+R_{s-})^{p^\star+\beta-1}\mathrm{d}R_s \\ &= \mathcal{C}\int_{T+}^t \mathrm{d}(1+R_s)^{p^\star+\beta} + \mathcal{C}\int_{T+}^t (1+R_{s-})^{p^\star+\beta-2}\Delta R_s\,\mathrm{d}R_s. \end{aligned}$$

Since $\int_0^\infty (1+R_{s-})^{-2\alpha}\Delta R_s\,\mathrm{d}R_s < \infty$ holds, we obtain by

$$\int_0^\infty \frac{\Delta R_s}{(1+R_{s-})^{2\alpha}}\,\mathrm{d}R_s < \infty \quad \Longrightarrow \quad \frac{\int_{T+}^t (1+R_{s-})^{p^\star+\beta-2}\Delta R_s\,\mathrm{d}R_s}{(1+R_t)^{p^\star+\beta+2\alpha-2}} \to 0$$

a convergence rate of the second summand. Hence we have

$$(1+R_t)^{p^\star}\mathbf{E}\|Y_t\|^2 \leq 1(1+R_T)^{p^\star} + \mathcal{C}(1+R_t)^{p^\star+\beta} + \mathrm{o}(1)(1+R_t)^{p^\star+\beta+2\alpha-2}$$

and

$$\mathbf{E}\|Y_t\|^2 \leq \mathcal{C}(1+R_t)^{-p^\star} + \mathcal{C}(1+R_t)^\beta + \mathrm{o}(1)(1+R_t)^{\beta+2\alpha-2}.$$

The assumptions and the choice of $p^\star$ ensure $\alpha \leq 1$ and $-p^\star < 2\gamma - 1 \leq 1-2\alpha+2\gamma \leq \beta$. Thus it follows that

$$\mathbf{E}\|Y_t\|^2 \leq \mathcal{O}\left((1+R_t)^\beta\right). \tag{38}$$

(b) Case: f is three times differentiable at $x^\star$. By $V(x,c) = \mathcal{O}(\|x\|^2) + \mathcal{O}(c^2)$ we have

$$\underset{\rho>0}{\exists}\ \underset{\delta_1>0}{\exists}\ \underset{\delta_2>0}{\exists}\ \underset{\|x\|<\delta_1}{\forall}\ \underset{c<\delta_2}{\forall} \qquad \|V(x,c)\| \leq \rho\|x\|^2 + \rho c^2,$$

where ρ is finite. However, since the bound includes $\|x\|^2$ instead of $\|x\|$, we can handle it. We choose κ as in the case of a twice differentiable f and define $\delta := \min\{\delta_1, 1, \frac{\kappa}{\rho}\} > 0$. Therefore we have, for s larger than a sufficiently large T,

$$\|V(Z_{s-}, c_s)\| \leq \rho\|Z_{s-}\|^2 + \rho c_s^2 \leq \rho\frac{\kappa}{\rho}\|Z_{s-}\| + \rho c_s^2 = \kappa\|Z_{s-}\| + \rho c_s^2$$

on a set of $\mathbf{P}$-measure greater than $1-\epsilon$. Proceeding further, as in the case of an f that is twice differentiable at $x^\star$, we get the bound (38) but with a modified $\beta := \max\{1-\alpha-4\gamma, 1-2\alpha+2\gamma\}$. The modification of the first term, in the argument list of the maximum, is due to the term $\mathcal{O}(c^2)$ instead of $\mathrm{o}(c)$. The second argument does not change, as it is related to the process $\int \frac{a_s}{4c_s}M(Z_{s-}, \mathrm{d}s)$. □

3.2 Asymptotic Normality

Before we start to deal with asymptotic normality, we give some notations, assumptions, and lemmata.

3.2.1 General Settings and Required Tools

Now we present some notations, assumptions, and lemmata that we will use in the proof of asymptotic normality of the standard as well as of the averaged Kiefer-Wolfowitz process. The first lemma explains how to deal with the inverse stochastic exponential and the second presents a useful representation of the Kiefer-Wolfowitz process.

In the entire section we assume that the Hessian H of the function f exists at least at $x^\star$ and consequently the representation

$$\frac{1}{2c_s}\Big\{f(Z_{s-}+c_s e_i)-f(Z_{s-}-c_s e_i)\Big\}_{i\in\{1,\dots,d\}} = H_{x^\star}Z_{s-}+V(Z_{s-},c_s)$$

exists (compare Section 3.1 for a further discussion of $V(\cdot,\cdot)$). Furthermore, we assume that the Hessian H is positive definite at $x^\star$ and continuous around $x^\star$, therefore it can be diagonalized $T^T H_{x^\star} T = D$. The diagonal elements λ_i of the diagonal matrix D are the eigenvalues of the matrix $H_{x^\star}$. Of course, the matrix T holds $TT^T = I$. In the following, we need the matrix-valued process ϕ_t whose diagonal elements are defined as

$$\phi_t^{ii} := \mathcal{E}_t\Big(-\int_0^\cdot \overline{a\lambda_i}^s (1+R_{s-})^{-\alpha}\,\mathrm{d}R_s\Big) \quad \text{where} \quad \overline{a\lambda}_i^s := a\lambda_i \mathbb{1}_{\{a\lambda_i\Delta R_s\neq(1+R_{s-})^\alpha\}}$$

and the non-diagonal elements are defined as zero. Furthermore we often need the inverse of the process ϕ_t, whose non-diagonal elements are also zero. The process $\overline{a\lambda}_i^s$ is used to ensure the invertibility of ϕ_t (compare Lemma 3.2). The common stochastic exponential gives

$$\mathcal{E}_t(R) = 1 + \int_0^t \mathcal{E}_{s-}(R)\,\mathrm{d}R_s. \tag{39}$$

We note that in the integral representation of $\mathcal{E}_t^{-1}(R)$ the process $\mathcal{E}_s^{-1}(R)$ appears as opposed to the left-continuous modification $\mathcal{E}_{s-}(R)$ in the representation of $\mathcal{E}_t(R)$.

Lemma 3.2. *Let $(R_t)_{t\in\mathbb{R}_+}$ be a monotonously increasing stochastic process with $R_0 = 0$. Then the inverse of the stochastic exponential $\mathcal{E}_t^{-1}(R)$ gives*

$$\mathcal{E}_t^{-1}(R) = 1 - \int_0^t \mathcal{E}_s^{-1}(R)\,\mathrm{d}R_s \qquad \textit{and} \qquad \mathcal{E}_t^{-1}(R) = \mathcal{E}_t\Big(-\int_0^\cdot \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\Big).$$

Furthermore, if the process $(R_t)_{t\in\mathbb{R}_+}$ *satisfies* $-\Delta R_s \neq 1$ *for all* $s \in \mathbb{R}_+$*, we have*

$$\mathcal{E}_t^{-1}(-R) = 1 + \int_0^t \mathcal{E}_s^{-1}(-R)\,\mathrm{d}R_s \qquad \text{and} \qquad \mathcal{E}_t^{-1}(-R) = \mathcal{E}_t\Big(\int_0^{\cdot} \frac{1}{1-\Delta R_s}\,\mathrm{d}R_s\Big).$$

Now we give another representation of the Kiefer-Wolfowitz process. We consider only processes $a_s := a(1+R_{s-})^{-\alpha}$ and $c_s := c(1+R_{s-})^{-\gamma}$ for fixed $0 \le \alpha, \gamma \le 1$, to get such a representation. This representation is helpful in the proof of the asymptotic normality and gives a better understanding of the insights of the stochastic integral equation.

Lemma 3.3. *The Kiefer-Wolfowitz stochastic integral equation* (5) *is solved by*

$$Z_t = T\phi_t\Big(T^T Z_0 - \frac{a}{2c}\int_0^t (1+R_{s-})^{\gamma-\alpha}\phi_s^{-1}T^T\,M(\mathrm{d}s, x^\star) - \int_0^t \phi_s^{-1}T^T\,\mathrm{d}\tilde{R}_s\Big) \tag{40}$$

where

$$\tilde{R}_t^i := a\int_0^t V^i(Z_{s-}, c_s)(1+R_{s-})^{-\alpha}\,\mathrm{d}R_s + \sum_{j,k} T_{ij}T_{jk}^T \sum_{s\le t} Z_{s-}^k \mathbb{1}_{\{a\lambda_j\Delta R_s = (1+R_{s-})^\alpha\}}$$
$$+ \frac{a}{2c}\int_0^t (1+R_{s-})^{\gamma-\alpha}\Big(M_i(\mathrm{d}s, Z_{s-}) - M_i(\mathrm{d}s, x^\star)\Big).$$

Proof of Lemma 3.2

First we show

$$\mathcal{E}_t^{-1}(R) = \mathcal{E}_t\Big(-\int_0^{\cdot} \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\Big).$$

Since R_t is monotonously increasing, we get

$$[-\int_0^{\cdot} \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s]_t = \sum_{s\le t}\Big(\frac{\Delta R_s}{1+\Delta R_s}\Big)^2 \quad \text{and} \quad \Delta\int_0^t \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s = \frac{\Delta R_t}{1+\Delta R_t}.$$

Hence using the formula for the stochastic exponential we conclude

$$\begin{aligned}
\mathcal{E}_t\Big(-\int_0^{\cdot} \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\Big) &= \exp\Big(-\int_0^t \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s - \frac{1}{2}\sum_{s\le t}\Big(\frac{\Delta R_s}{1+\Delta R_s}\Big)^2\Big) \\
&\quad \cdot \prod_{s\le t}\Big(1 - \frac{\Delta R_s}{1+\Delta R_s}\Big)\exp\Big(\frac{\Delta R_s}{1+\Delta R_s} + \frac{1}{2}\Big(\frac{\Delta R_s}{1+\Delta R_s}\Big)^2\Big) \\
&= \exp\Big(-\int_0^t \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\Big)\prod_{s\le t}\Big(1 - \frac{\Delta R_s}{1+\Delta R_s}\Big)\exp\Big(\frac{\Delta R_s}{1+\Delta R_s}\Big) \\
&= \exp\Big(-\int_0^t \frac{1+\Delta R_s - \Delta R_s}{1+\Delta R_s}\,\mathrm{d}R_s\Big)\prod_{s\le t}\Big(1 - \frac{\Delta R_s}{1+\Delta R_s}\Big)\exp\Big(\frac{\Delta R_s}{1+\Delta R_s}\Big)
\end{aligned}$$

$$\begin{aligned}
&= \exp\big(-\int_0^t \mathrm{d}R_s\big) \prod_{s\le t}\big(\frac{1}{1+\Delta R_s}\big) \exp\big(\frac{\Delta R_s}{1+\Delta R_s} + \frac{(\Delta R_s)^2}{1+\Delta R_s}\big) \\
&= \exp\big(-R_t\big) \prod_{s\le t}\Big(\frac{1}{1+\Delta R_s}\Big) \exp\big(\Delta R_s\big) \qquad (41)\\
&= \Big(\exp\big(R_t\big) \prod_{s\le t}(1+\Delta R_s) \exp\big(-\Delta R_s\big)\Big)^{-1} \\
&= \mathcal{E}_t^{-1}(R). \qquad (42)
\end{aligned}$$

Now we verify

$$\mathcal{E}_t^{-1}(R) = 1 - \int_0^t \mathcal{E}_s^{-1}(R)\,\mathrm{d}R_s.$$

Using (42) and (39), we have

$$\begin{aligned}
\mathcal{E}_t^{-1}(R) &= \mathcal{E}_t\big(-\int_0^\cdot \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\big) = 1 - \int_0^t \mathcal{E}_{s-}\big(-\int_0^\cdot \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\big)\frac{1}{1+\Delta R_s}\,\mathrm{d}R_s \\
&\overset{(\star)}{=} 1 - \int_0^t \mathcal{E}_s^{-1}(R)\,\mathrm{d}R_s,
\end{aligned}$$

where we get $(\star)$ using (41)

$$\begin{aligned}
\frac{1}{1+\Delta R_t}\,\mathcal{E}_{t-}\big(-\int_0^\cdot \frac{1}{1+\Delta R_s}\,\mathrm{d}R_s\big) &\overset{(41)}{=} \frac{1}{1+\Delta R_t}\exp\big(-R_{t-}\big)\prod_{s<t}\frac{\exp\big(\Delta R_s\big)}{1+\Delta R_s} \\
&= \frac{\exp\big(-R_{t-}\big)}{1+\Delta R_t}\frac{1+\Delta R_t}{\exp\big(\Delta R_t\big)}\prod_{s\le t}\frac{\exp\big(\Delta R_s\big)}{1+\Delta R_s} = \exp\big(-R_t\big)\prod_{s\le t}\frac{\exp\big(\Delta R_s\big)}{1+\Delta R_s} \\
&\overset{(41)}{=} \mathcal{E}_t^{-1}(R).
\end{aligned}$$

Thus we have proven the first two equalities for $\mathcal{E}_t^{-1}(R)$. Verification of the other two equalities for $\mathcal{E}_t^{-1}(-R)$ can be done in an analogous manner. Here we only have to ensure that $\Delta R_s \neq -1$ holds, otherwise the process $\frac{1}{1+\Delta R_s}$ is not well defined. But this is excluded by assumption. Such a complication does not appear in the first part of the proof because R_s is monotonously increasing, ΔR_s is positive and $1 + \Delta R_s \geq 1$ holds. □

Proof of Lemma 3.3

We have

$$\begin{aligned}
Z_t &= Z_0 - \int_0^t a_s \frac{1}{2c_s}\{\cdot\}.\,\mathrm{d}R_s - \int_0^t \frac{a_s}{2c_s} M(\mathrm{d}s, Z_{s-}) \\
&= Z_0 - a\int_0^t \big(H_{x^\star} Z_{s-} + V(Z_{s-}, c_s)\big)\big(1+R_{s-}\big)^{-\alpha}\,\mathrm{d}R_s - \int_0^t \frac{a}{2c}(1+R_{s-})^{\gamma-\alpha} M(\mathrm{d}s, Z_{s-})
\end{aligned}$$

$$= Z_0 - T\int_0^t \overline{D}_s T^T Z_{s-}(1+R_{s-})^{-\alpha}\,\mathrm{d}R_s - \int_0^t \frac{a}{2c}(1+R_{s-})^{\gamma-\alpha} M(\mathrm{d}s, x^\star) - \tilde{R}_t, \qquad (43)$$

where $\overline{D}_s := \{\overline{a\lambda_i}^s \delta_{ij}\}_{ij}$. Now we show that

$$Z_t = T\phi_t\left(T^T Z_0 - \frac{a}{2c}\int_0^t \phi_s^{-1} T^T (1+R_{s-})^{\gamma-\alpha} M(\mathrm{d}s, x^\star) - \int_0^t \phi_s^{-1} T^T \mathrm{d}\tilde{R}_s\right) \qquad (44)$$

solves the stochastic integral equation (43). We will show this by applying Z_t as given in (44) on the left hand side and on the right hand side of (43). Then we will verify the equality of the expressions for every coordinate $i \in \{1, \dots, d\}$. To do so, we use the i-th coordinate

$$\mathrm{d}Z_t^i = -\sum_{j,k} T_{ij} T_{jk}^T \overline{a\lambda_j}^s Z_{s-}^k (1+R_{s-})^{-\alpha}\,\mathrm{d}R_s - \frac{a}{2c}(1+R_{s-})^{\gamma-\alpha} M_i(\mathrm{d}s, x^\star) - \mathrm{d}\tilde{R}_t^i$$

from (43) and the i-th coordinate

$$\begin{aligned}\mathrm{d}(\phi_t^{-1} T^T Z_t)_i &= -\frac{a}{2c}\sum_k T_{ik}^T \phi_t^{ii^{-1}} (1+R_{t-})^{\gamma-\alpha} M_k(\mathrm{d}t, x^\star) - \sum_k T_{ik}^T \phi_t^{ii^{-1}} \mathrm{d}\tilde{R}_t^k \\ &= \sum_k T_{ik}^T \phi_t^{ii^{-1}}\left(-\frac{a}{2c}(1+R_{t-})^{\gamma-\alpha} M_k(\mathrm{d}t, x^\star) - \mathrm{d}\tilde{R}_t^k\right)\end{aligned}$$

which we easily get from (44). The i-th coordinate of the left hand side of (43) obeys

$$\begin{aligned}
&\mathrm{d}Z_t^i = \mathrm{d}(T\phi_t\phi_t^{-1}T^T Z_t)_i = \sum_{j,k} T_{ij}T_{jk}^T \mathrm{d}(\phi_t^{jj}\phi_t^{jj^{-1}} Z_t^k) \\
&= \sum_{j,k} T_{ij}T_{jk}^T\left(\phi_{t-}^{jj}\mathrm{d}(\phi_t^{jj^{-1}} Z_t^k) + (\phi_{t-}^{jj^{-1}} Z_{t-}^k)\mathrm{d}\phi_t^{jj} + \mathrm{d}[\phi^{jj}, \phi^{jj^{-1}} Z^k]_t\right) \\
&\overset{(\star)}{=} \sum_{j,k} T_{ij}T_{jk}^T\left(\phi_t^{jj}\mathrm{d}(\phi_t^{jj^{-1}} Z_t^k) - \Delta\phi_t^{jj}\mathrm{d}(\phi_t^{jj^{-1}} Z_t^k) + (\phi_{t-}^{jj^{-1}} Z_{t-}^k)\mathrm{d}\phi_t^{jj} + \Delta\phi_t^{jj}\mathrm{d}(\phi_t^{jj^{-1}} Z_t^k)\right) \\
&= \sum_{j,k} T_{ij}T_{jk}^T\phi_t^{jj}\mathrm{d}(\phi_t^{jj^{-1}} Z_t^k) + \sum_{j,k} T_{ij}T_{jk}^T\phi_{t-}^{jj^{-1}} Z_{t-}^k\mathrm{d}\phi_t^{jj} \\
&= \sum_j T_{ij}\phi_t^{jj}\underbrace{\mathrm{d}(\phi_t^{jj^{-1}}\sum_k T_{jk}^T Z_t^k)}_{=\mathrm{d}(\phi_t^{-1}T^T Z_t)_j} + \sum_{j,k} T_{ij}T_{jk}^T\phi_{t-}^{jj^{-1}} Z_{t-}^k\mathrm{d}\phi_t^{jj} \\
&= \sum_k \underbrace{\sum_j T_{ij}T_{jk}^T}_{=(TT^T)_{ik} = I_{ik} = \delta_{ik}}\left(-\frac{a}{2c}(1+R_{t-})^{\gamma-\alpha}M_k(\mathrm{d}t, x^\star) - \mathrm{d}\tilde{R}_t^k\right) + \sum_{j,k} T_{ij}T_{jk}^T\phi_{t-}^{jj^{-1}} Z_{t-}^k\mathrm{d}\phi_t^{jj} \\
&= -\frac{a}{2c}(1+R_{t-})^{\gamma-\alpha}M_i(\mathrm{d}t, x^\star) - \mathrm{d}\tilde{R}_t^i + \sum_{j,k} T_{ij}T_{jk}^T\phi_{t-}^{jj^{-1}} Z_{t-}^k\mathrm{d}\phi_t^{jj}.
\end{aligned}$$

To show ($\star$), that is, $\mathrm{d}[\phi^{jj}, \phi^{jj^{-1}} Z^k]_t = \Delta\phi_t^{jj} \mathrm{d}(\phi_t^{jj^{-1}} Z_t^k)$, we recall that ϕ_t^{jj} is a process of finite variation. The i-th coordinate on the right hand side of (43) gives

$$\mathrm{d}Z_t^i = -\sum_{j,k} T_{ij} T_{jk}^T Z_{t-}^k \underbrace{\overline{a\lambda_j}^t (1+R_{t-})^{-\alpha}\,\mathrm{d}R_t}_{= -\phi_{t-}^{jj^{-1}}(-\phi_{t-}^{jj}\overline{a\lambda_j}^t(1+R_{t-})^{-\alpha})\,\mathrm{d}R_t = -\phi_{t-}^{jj^{-1}}\,\mathrm{d}\phi_t^{jj}} - \frac{a}{2c}(1+R_{t-})^{\gamma-\alpha} M_i(\mathrm{d}t, x^\star) - \mathrm{d}\tilde{R}_t^i$$

$$= \sum_{j,k} T_{ij} T_{jk}^T Z_{t-}^k \phi_{t-}^{jj^{-1}}\,\mathrm{d}\phi_t^{jj} - \frac{a}{2c}(1+R_{t-})^{\gamma-\alpha} M_i(\mathrm{d}t, x^\star) - \mathrm{d}\tilde{R}_t^i.$$

So, using Z_t from (44), the left- and the right-hand side of (43) are equal. Thus, Z_t given by (44) solves the stochastic integral equation (43) and the lemma is proven. $\square$

3.2.2 Asymptotic Normality of the Process Using Standard Weights

Here we consider the process that is given by (5) or, equivalently, by (40), with the choices $a_s := a(1+R_{s-})^{-1}$ and $c_s := c(1+R_{s-})^{-\gamma}$. Furthermore, we make the restriction that the process R_t is deterministic. To prove asymptotic normality this process requires a restriction on the smallest eigenvalue of the Hessian $H_{x^\star}$ at $x^\star$, which is usually unknown. We will discuss and eliminate this problem using slowly decaying weights and the averaged process in the next chapter.

Theorem 3.2. *Let the assumptions of Theorem 3.1 with $\alpha = 1$ be valid. Further we assume $\sum_{0\le s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} < \infty$ for all $i \in \{1,\dots,d\}$,*

$$\lim_{\substack{s\to\infty\\ x\to x^\star}} h_s(x, x^\star) = \lim_{\substack{s\to\infty\\ x\to x^\star}} h_s(x) = h(x^\star) \quad \textit{where} \quad h_s(x,y) := \frac{\mathrm{d}\left[\int_0^\cdot M(\mathrm{d}t, x), \int_0^\cdot M(\mathrm{d}t, y)\right]_s}{\mathrm{d}R_s}$$

and, for all $\epsilon \in (0,1]$,

$$\frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{G_{s,t}^\epsilon} x^T x\ \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x)}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \xrightarrow{\mathbf{P}} 0 \ \ (t\to\infty),$$

where $M_t^\star := \int_0^t M(\mathrm{d}s, x^\star)$, $G_{s,t}^\epsilon := \left\{ x\in\mathbb{R}^d \,\middle|\, \|x\| > \epsilon \frac{(1+R_t)^{\gamma-\frac{1}{2}+a\lambda_{\min}}}{(1+R_{s-})^{\gamma-1}(1+R_s)^{a\lambda_{\min}}} \right\}$, and $\nu^{M^\star}$ is the compensator of the jump-measure $\mu^{M^\star}$ of the local martingale $M_t^\star$.

(a) If f is twice differentiable at $x^\star$ and $\gamma = \frac{1}{4}$, we get

$$(1+R_t)^{\frac{1}{4}}(Z_t - x^\star) \xrightarrow{\mathcal{D}} N(0, \Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma := TUT^T$, and

$$U_{kv} := \frac{a^2 \left(T^T h(x^\star) T\right)_{kv}}{4c^2(a(\lambda_k+\lambda_v) - \frac{1}{2})}.$$

(b) If f is three times differentiable at $x^\star$ and $\gamma = \frac{1}{6}$, we get

$$(1+R_t)^{\frac{1}{3}}(Z_t - x^\star) \xrightarrow{\mathcal{D}} N(\mu, \Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma := TUT^T$,

$$\mu := -\frac{ac^2}{6}(aH_{x^\star} - \frac{1}{3}I)^{-1}\left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star)\right)_{i\in\{1,\dots,d\}}, \text{ and } U_{kv} := \frac{a^2\left(T^T h(x^\star)T\right)_{kv}}{4c^2(a(\lambda_k+\lambda_v)-\frac{2}{3})}.$$

To ensure asymptotic normality in Theorem 3.2, we use a condition on the jumps of the local martingale of a Lindeberg type. The measure $\nu^{M^\star}$ plays a major role as compensator of the jump-measure $\mu^{M^\star}$ of the process $M_t^\star$. Assumptions of this type are well-known in the theory of limit theorems of martingales and semimartingales, in particular for continuous-time processes (see, e.g., [19]). Roughly speaking it is sufficient to verify the "classical" condition of uniform asymptotic negligibility of jumps and the convergence in probability of the quadratic variation process to ensure convergence in distribution of a local martingale or a normalized local martingale to a Brownian motion. But the detection of convergence in probability of the predictable quadratic variation and the "classical" condition is not sufficient (see, for example, [12]). In our case, it is impossible to obtain convergence only on the basis of $\lceil M \rceil_t = \int_0^t h_s \, \mathrm{d}R_s$ and not-too-restrictive assumptions on h and R.

Subsequent to the theorem above, which includes a condition of the Lindeberg type, we present a lemma that does not use the jump measure or its compensator explicitly. Conditions $(S1)$ and $(S2)$ in Lemma 3.4 are easier to interpret than the condition of the Lindeberg type.

Lemma 3.4. *If we replace the condition*

$$\frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{G^\epsilon_{s,t}} x^T x \; \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x)}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \xrightarrow{\mathbf{P}} 0 \;\; (t\to\infty),$$

which appears in Theorem 3.2, with the two conditions

$$(S1) \qquad \frac{\mathbf{E}\sup_{0\le s\le t}(1+R_{s-})^{2a\lambda_{\min}+2\gamma-2}\|\Delta M_s^\star\|^2}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \xrightarrow{t\to\infty} 0$$

$$(S2) \qquad \frac{\mathbf{E}\sum_{s\le t}(\Delta M_s^{\star^m}\Delta M_s^{\star^n})^2(1+R_{s-})^{4a\lambda_{\min}+4\gamma-4}}{(1+R_t)^{4a\lambda_{\min}+4\gamma-2}} \xrightarrow{t\to\infty} 0,$$

then the conclusion of Theorem 3.2 also holds.

Proof of Theorem 3.2

Without loss of generality we assume $x^\star = 0$. If we use the representation of Z_t presented in Lemma 3.3, it is easy to see that we can split the proof into three parts.

$$\textbf{(I)} \qquad (1+R_t)^{\frac{1}{2}-\gamma} T\phi_t T^T Z_0 \xrightarrow{\mathbf{P}} \vec{0}$$

$$\textbf{(II)} \quad \begin{cases} (a) & (1+R_t)^{\frac{1}{4}} T\phi_t \int_0^t \phi_s^{-1} T^T \,\mathrm{d}\tilde{R}_s \xrightarrow{\mathbf{P}} \vec{0} \\ (b) & (1+R_t)^{\frac{1}{3}} T\phi_t \int_0^t \phi_s^{-1} T^T \,\mathrm{d}\tilde{R}_s \xrightarrow{\mathbf{P}} \frac{ac^2}{6}(aH_0 - \frac{1}{3}I)^{-1} \left(\frac{\partial^3 f}{(\partial x_i)^3}(0)\right)_{i\in\{1,\ldots,d\}} \end{cases}$$

$$\textbf{(III)} \qquad (1+R_t)^{\frac{1}{2}-\gamma} T\phi_t \frac{a}{2c} \int_0^t (1+R_{s-})^{\gamma-1} \phi_s^{-1} T^T M(\mathrm{d}s, 0) \xrightarrow{\mathcal{D}} N(0,\Sigma)$$

Using Slutsky's theorem we get the desired result.

The matrix-valued process ϕ_s appears in all parts, therefore we consider it more closely. First, it will be helpful to examine the integral that appears in the course of ϕ_s.

$$\begin{aligned} -\int_0^t \frac{\overline{a\lambda_i}^s}{(1+R_{s-})^\alpha} \,\mathrm{d}R_s &= -a\lambda_i \int_0^t (1+R_{s-})^{-\alpha} \mathbb{1}_{\{a\lambda_i \Delta R_s \neq (1+R_{s-})^\alpha\}} \,\mathrm{d}R_s \\ &= -a\lambda_i \int_0^t (1+R_{s-})^{-\alpha} \,\mathrm{d}R_s + a\lambda_i \int_0^t (1+R_{s-})^{-\alpha} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})^\alpha\}} \,\mathrm{d}R_s \\ &= -a\lambda_i \int_0^t (1+R_{s-})^{-\alpha} \,\mathrm{d}R_s + \sum_{s\le t} \frac{a\lambda_i \Delta R_s}{(1+R_{s-})^\alpha} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})^\alpha\}} \\ &= -a\lambda_i \int_0^t (1+R_{s-})^{-\alpha} \,\mathrm{d}R_s + \sum_{s\le t} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})^\alpha\}} \end{aligned} \tag{45}$$

In the case $\alpha = 1$, using Itô's formula with $f'(x) := (1+x)^{-1}$, we have

$$\begin{aligned} &- a\lambda_i \int_0^t (1+R_{s-})^{-1} \,\mathrm{d}R_s \\ &\qquad = -a\lambda_i \ln(1+R_t) + a\lambda_i \sum_{s\le t} \left\{ \ln(1+R_s) - \ln(1+R_{s-}) - \frac{\Delta R_s}{1+R_{s-}} \right\} \\ &\qquad = -a\lambda_i \ln(1+R_t) + a\lambda_i \sum_{s\le t} \left\{ \ln\left(\frac{1+R_s}{1+R_{s-}}\right) - \frac{\Delta R_s}{1+R_{s-}} \right\} \\ &\qquad = -a\lambda_i \ln(1+R_t) + a\lambda_i \sum_{s\le t} \left\{ \ln\left(1 + \frac{\Delta R_s}{1+R_{s-}}\right) - \frac{\Delta R_s}{1+R_{s-}} \right\} \end{aligned}$$

and by (45)

$$\begin{aligned} -\Delta \int_0^t \frac{\overline{a\lambda_i}^s}{(1+R_{s-})} \,\mathrm{d}R_s &= -a\lambda_i \Delta \int_0^t (1+R_{s-})^{-1} \,\mathrm{d}R_s + \Delta \sum_{s\le t} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} \\ &= -\frac{a\lambda_i \Delta R_t}{1+R_{t-}} + \mathbb{1}_{\{a\lambda_i \Delta R_t = (1+R_{t-})\}}. \end{aligned}$$

Thus the formula $\mathcal{E}_t(X) = e^{X_t - X_0 - 0.5[X,X]_t^c} \prod_{s \le t}(1+\Delta X_s)e^{-\Delta X_s}$ for the stochastic exponential leads to

$$\begin{aligned}
\phi_t^{ii} &= \mathcal{E}_t(-\int_0^{\cdot} \frac{\overline{a\lambda_i}^s}{(1+R_{s-})}\,\mathrm{d}R_s) \\
&= \exp\left(-a\lambda_i \ln(1+R_t) + a\lambda_i \sum_{s\le t}\left\{\ln\left(1+\frac{\Delta R_s}{1+R_{s-}}\right) - \frac{\Delta R_s}{1+R_{s-}}\right\} + \sum_{s\le t} \mathbb{1}_{\{\frac{a\lambda_i \Delta R_s}{(1+R_{s-})}=1\}}\right) \\
&\quad \prod_{s\le t}\left(1-\frac{a\lambda_i\Delta R_s}{1+R_{s-}} + \mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}}\right)\exp\left(\frac{a\lambda_i\Delta R_s}{1+R_{s-}} - \mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}}\right) \\
&= (1+R_t)^{-a\lambda_i} \underbrace{\prod_{0<s\le t}\left((1-\frac{a\lambda_i\Delta R_s}{1+R_{s-}} + \mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}})(1+\frac{\Delta R_s}{1+R_{s-}})^{a\lambda_i}\right)}_{=:\,\Pi_{0,t}^{a\lambda_i}} \\
&= (1+R_t)^{-a\lambda_i}\,\Pi_{0,t}^{a\lambda_i}.
\end{aligned}$$

To show that $\Pi_{0,t}^{a\lambda_i}$ converges, we split the product and take the logarithm to rewrite one part of the product as a sum. To split the product, we use the time

$$\tau^1 := \min\left\{\, t \in \mathbb{R}_+ \;:\; \frac{\Delta R_s}{1+R_{s-}} < \frac{1}{a\lambda_{\max}+1} \;\; \forall_{s>t} \right\}.$$

From the assumption, that is $\frac{\Delta R_s}{1+R_{s-}} \to 0$, follows that $\tau^1 < \infty$. Then we use a Taylor expansion to obtain

$$\begin{aligned}
&\left(\ln(1-\frac{a\lambda_i\Delta R_s}{1+R_{s-}}) + a\lambda_i \ln(1+\frac{\Delta R_s}{1+R_{s-}})\right) \\
&\qquad = \left(-\frac{a\lambda_i}{2} - \frac{a^2\lambda_i^2}{2} + \rho_s\right)\left(\frac{\Delta R_s}{1+R_{s-}}\right)^2 \quad \text{where } \rho_s \to 0.
\end{aligned}$$

Now we define the time

$$\tau := \min\left\{\, t \ge \tau^1 \;:\; |\rho_s| < 1 \;\; \forall_{s>t} \right\}$$

to split the product. By $\rho_s \to 0$ it follows $\tau < \infty$. We write

$$\Pi_{0,t}^{a\lambda_i} = \Pi_{0,\tau}^{a\lambda_i}\,\Pi_{\tau,t}^{a\lambda_i}.$$

Thus it is sufficient to verify $|\Pi_{0,\tau}^{a\lambda_i}| < \infty$ and convergence of $\Pi_{\tau,t}^{a\lambda_i}$ (where, in the product $\Pi_{\tau,t}^{a\lambda_i}$, only positive terms appear). To verify convergence of $\Pi_{\tau,t}^{a\lambda_i}$, as mentioned above, it is reasonable to rewrite the product into a sum using logarithms

$$\ln(\Pi_{\tau,t}^{a\lambda_i}) = \sum_{\tau<s\le t} \ln\left((1-\frac{a\lambda_i\Delta R_s}{1+R_{s-}} + \mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}})(1+\frac{\Delta R_s}{1+R_{s-}})^{a\lambda_i}\right)$$

$$
\begin{aligned}
&= \sum_{\tau<s\leq t} \ln\left((1-\frac{a\lambda_i \Delta R_s}{1+R_{s-}})(1+\frac{\Delta R_s}{1+R_{s-}})^{a\lambda_i}\right)\\
&= \sum_{\tau<s\leq t} \left(\ln(1-\frac{a\lambda_i \Delta R_s}{1+R_{s-}}) + a\lambda_i \ln(1+\frac{\Delta R_s}{1+R_{s-}})\right)\\
&= \sum_{\tau<s\leq t} \left(-\frac{a\lambda_i}{2} - \frac{a^2\lambda_i^2}{2} + \rho_s\right)\left(\frac{\Delta R_s}{1+R_{s-}}\right)^2.
\end{aligned}
$$

Convergence of $\ln(\Pi_{\tau,t}^{a\lambda_i})$ follows from $\rho_s \to 0$, $|\rho_s| < 1$,

$$
\sum_{0<s\leq\infty} \left(\frac{\Delta R_s}{1+R_{s-}}\right)^2 = \frac{1}{a^2} \sum_{0<s\leq\infty} a_s^2 \left(\Delta R_s\right)^2 = \frac{1}{a^2}\int_0^\infty a_s^2 \Delta R_s \,\mathrm{d}R_s^d < \infty
$$

and thus $\Pi_{\tau,t}^{a\lambda_i}$ converges, too.

Now we verify boundedness of $|\Pi_{0,\tau}^{a\lambda_i}|$. We write

$$
\Pi_{0,\tau}^{a\lambda_i} = \left(\prod_{0<s\leq\tau}(1-\frac{a\lambda_i \Delta R_s}{1+R_{s-}} + \mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}})\right)\left(\prod_{0<s\leq\tau}(1+\frac{\Delta R_s}{1+R_{s-}})\right)^{a\lambda_i}
$$

and seek a bound of both factors. The last factor satisfies

$$
\begin{aligned}
\ln\left(\prod_{0<s\leq\tau}(1+\frac{\Delta R_s}{1+R_{s-}})\right) &= \sum_{0<s\leq\tau} \ln\left(1+\frac{\Delta R_s}{1+R_{s-}}\right) \leq \sum_{0<s\leq\tau}\frac{\Delta R_s}{1+R_{s-}} \leq \sum_{0<s\leq\tau}\Delta R_s\\
&\leq R_\tau^d < \infty
\end{aligned}
$$

and thus it is bounded. Now we show that the first factor is bounded. From the assumptions it follows that the set

$$
\left\{\ s \in \mathbb{R}_+ \ :\ a\lambda_i \Delta R_s = (1+R_{s-})\ \right\}
$$

is finite. Hence we can neglect the term $\mathbb{1}_{\{a\lambda_i\Delta R_s=(1+R_{s-})\}}$. Furthermore, we split the product into two products. The first one contains all factors with a norm smaller than one and thus the norm of this product is bounded by one. The last step will show that the second product has only a finite number of factors. We have

$$
\left|1-\frac{a\lambda_i \Delta R_s}{1+R_{s-}}\right| > 1 \qquad \Longrightarrow \qquad \Delta R_s \geq \frac{2}{a\lambda_i}.
$$

The finiteness of the set

$$
\left\{\ s \in \mathbb{R}_+ \ :\ s \leq \tau \ \wedge\ \Delta R_s \geq \frac{2}{a\lambda_i}\ \right\}
$$

follows from $R^d_\tau < \infty$. Hence convergence of $\Pi^{a\lambda_i}_{0,t}$ is proven. In the following we use

$$\phi^{ii}_t = \mathcal{E}_t(-\int_0^\cdot \frac{\overline{a\lambda_i}^s}{(1+R_{s-})}\,\mathrm{d}R_s) = (1+R_t)^{-a\lambda_i}\Pi^{a\lambda_i}_{0,t} \qquad \text{where} \quad \Pi^{a\lambda_i}_{0,t} \overset{t\to\infty}{\longrightarrow} \Pi^{a\lambda_i}_\infty.$$

Part (I): Obviously we have

$$(1+R_t)^{\frac{1}{2}-\gamma}(1+R_t)^{-a\lambda_{\min}} \to 0 \qquad \Longrightarrow \qquad (1+R_t)^{\frac{1}{2}-\gamma}T\phi_t T^T Z_0 \overset{\mathbf{P}}{\longrightarrow} \vec{0}.$$

Since $R_t \uparrow \infty$ holds, the left hand side follows from $\lambda_{\min} > \frac{1-2\gamma}{2a}$, which in turn is a consequence of the assumptions.

Part (II): *(a) Case: f is two times differentiable at* 0. We mention that we have $\gamma = \frac{1}{4}$, and hence $\frac{1}{2} - \gamma = \frac{1}{4}$. But here we use the expression $\frac{1}{2} - \gamma$ in order to reuse the arguments and steps also in part (b). Using

$$\phi_t = \left(\frac{\Pi^{a\lambda_i}_{0,t}}{(1+R_t)^{a\lambda_i}}\delta_{ij}\right)_{ij},$$

it remains to verify $(1+R_t)^{\frac{1}{2}-\gamma}T\phi_t \int_0^t \phi_s^{-1}T^T\,\mathrm{d}\tilde{R}_s \overset{\mathbf{P}}{\to} \vec{0}$ or

$$\left\{\sum_{j,k=1}^d (1+R_t)^{\frac{1}{2}-\gamma-a\lambda_k}T_{ik}T^T_{kj}\int_0^t \frac{\Pi^{a\lambda_k}_{0,t}}{\Pi^{a\lambda_k}_{0,s}}(1+R_s)^{a\lambda_k}\,\mathrm{d}\tilde{R}_s^{\,j}\right\}_{i\in\{1,\dots,d\}} \overset{\mathbf{P}}{\to} \vec{0}.$$

Again this can be simplified by the following arguments. The assumption $\lambda_{\min} > \frac{1-2\gamma}{2a}$ implies $(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \to 0$. Furthermore, for all $i \in \{1,\dots,d\}$ we find

$$|\Pi^{a\lambda_i}_{0,t}| \le \mathcal{C} < \infty \quad \wedge \quad \Pi^{a\lambda_i}_{0,t} \to \Pi^{a\lambda_i}_\infty \qquad \Longrightarrow \qquad \frac{\Pi^{a\lambda_i}_{0,t}}{\Pi^{a\lambda_i}_{0,s}} = \Pi^{a\lambda_i}_{s,t} = (1+\mathrm{o_b}(1)).$$

Since we have

$$\begin{aligned}\mathbf{P}\Big[\|X_t\| > \epsilon\Big] &= \mathbf{P}\Big[\sum_{i=1}^d X_t^{i^2} > \epsilon^2\Big] \le \mathbf{P}\Big(\bigcup_{i=1}^d [X_t^{i^2}(\omega) > \frac{\epsilon^2}{d}]\Big) \le \sum_{i=1}^d \mathbf{P}\Big[X_t^{i^2}(\omega) > \frac{\epsilon^2}{d}\Big] \\ &= \sum_{i=1}^d \mathbf{P}\Big[|X_t^i| > \epsilon^\star\Big],\end{aligned}$$

it is sufficient to consider only convergence in probability of each coordinate. We have to prove, for all $i, j \in \{1,\dots,d\}$, that

$$(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i}\int_0^t (1+R_s)^{a\lambda_i}\,\mathrm{d}\tilde{R}_s^{\,j} \simeq H^1_t + H^2_t + H^3_t \overset{\mathbf{P}}{\to} 0,$$

where

$$H_t^1 := a(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^t (1+R_{s-})^{a\lambda_i-1} V^j(Z_{s-}, c_s)\, \mathrm{d}R_s$$
$$H_t^2 := \frac{a}{2c}(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^t (1+R_{s-})^{a\lambda_i+\gamma-1} (M_j(\mathrm{d}s, Z_{s-}) - M_j(\mathrm{d}s, 0))$$
$$H_t^3 := \sum_{l,k} (1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \sum_{s\le t} (1+R_{s-})^{a\lambda_i} T_{jl} T_{lk}^T Z_{s-}^k \mathbb{1}_{\{\lambda_l \Delta R_s = (1+R_{s-})\}}.$$

The above expressions are asymptotically equal only because on the right side we used the integrands $(1+R_{s-})^{a\lambda_i}$ instead of $(1+R_s)^{a\lambda_i}$. But this substitution does not have an effect on the asymptotic results as $a\lambda_{\min} > \frac{1}{2} - \gamma$ and thus $(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \overset{t\to\infty}{\longrightarrow} 0$ holds, and by $\frac{\Delta R_s}{1+R_{s-}} \to 0$ it follows

$$(1+R_s)^{a\lambda_i} = \left(\frac{1+R_s}{1+R_{s-}}\right)^{a\lambda_i} (1+R_{s-})^{a\lambda_i} = \underbrace{\left(1 + \frac{\Delta R_s}{1+R_{s-}}\right)^{a\lambda_i}}_{\overset{s\to\infty}{\longrightarrow} 1} (1+R_{s-})^{a\lambda_i}. \tag{46}$$

Finally, we have to verify

$$(i)\;\; H_t^1 \overset{\mathbf{P}}{\to} 0 \qquad (ii)\;\; H_t^2 \overset{\mathbf{P}}{\to} 0 \qquad (iii)\;\; H_t^3 \overset{\mathbf{P}}{\to} 0.$$

Proof of (i): We will show

$$\underset{\epsilon_1>0}{\forall}\ \underset{\epsilon_2>0}{\forall}\ \underset{t_0}{\exists}\ \underset{t\ge t_0}{\forall} \quad \mathbf{P}\left[\left|\int_0^t \frac{(1+R_{s-})^{a\lambda_i-1}}{(1+R_t)^{\gamma+a\lambda_i-\frac{1}{2}}} V^j(Z_{s-}, c_s)\, \mathrm{d}R_s\right| > \epsilon_1\right] \le \epsilon_2. \tag{47}$$

Let $\epsilon_1, \epsilon_2 > 0$ be arbitrary but fixed. Application of Theorem 3.1 with $\epsilon := \frac{\epsilon_2}{8}$ and $\gamma := \frac{1}{4}$ (or $\gamma := \frac{1}{6}$ in (b)) results in a T_1 such that, for all $t \ge T_1$, we have

$$\mathbf{E}\|Y_t\|^2 \le K(1+R_t)^{-p} \quad \text{and} \quad \mathbf{P}\left[\forall_{t\ge 0}\ Y_t = Z_t\right] \ge 1 - \frac{\epsilon_2}{8}$$

for $p := \frac{1}{2}$ (or $p := \frac{2}{3}$ in (b)). From the assumption on the differentiability of the function f and (31) it follows $V^j(x,b) = \mathrm{o}(\|x\|) + \mathrm{o}(b)$, and hence we have

$$\underset{\rho>0}{\exists}\ \underset{\|x\|,b\le\rho}{\forall} \quad |V^j(x,b)| \le \frac{\epsilon_1\epsilon_2(a\lambda_i - \frac{p}{2})}{64\sqrt{K}}\|x\| + \frac{\epsilon_1(a\lambda_i - \gamma)}{8c} b.$$

Furthermore, Lemma 3.1, with the choices $\epsilon := \frac{\epsilon_2}{4}$ and $\delta := \rho$, ensures the existence of a T_2 such that

$$\mathbf{P}\left[\sup_{t\ge T_2} \|Z_t\| \le \rho\right] \ge 1 - \frac{\epsilon_2}{4}$$

holds. The existence of a $T_3 < \infty$ such that, for all $t > T_3$, we have $c_t \le \rho$ follows from $c_s \to 0$. We choose $T := \max\{T_1, T_2, T_3\}$. Consequently T is deterministic and $T < \infty$ holds. To verify (47), it suffices to show

$$\begin{aligned}
\mathbf{P}&\left[\left|\int_0^T \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s + \int_{T+}^t \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| > \epsilon_1\right] \\
&\le \mathbf{P}\left(\left[\left|\int_0^T \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \ge \frac{\epsilon_1}{2}\right] \cup \left[\left|\int_{T+}^t \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \ge \frac{\epsilon_1}{2}\right]\right) \\
&\le \underbrace{\mathbf{P}\left[\left|\int_0^T \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \ge \frac{\epsilon_1}{2}\right]}_{\le \frac{\epsilon_2}{2}} + \underbrace{\mathbf{P}\left[\left|\int_{T+}^t \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \ge \frac{\epsilon_1}{2}\right]}_{\le \frac{\epsilon_2}{2}}.
\end{aligned}$$

Here we used $\div$ instead of the real fraction to shorten notations. To prove the first bound "$\le \frac{\epsilon_2}{2}$", we need an inequality for $V^j(\cdot)$. Since $\nabla f(x)$ is Lipschitz-continuous we have

$$\begin{aligned}
|V^j(x, c)| &= |\frac{1}{2}\int_{-1}^1 \nabla_j f(x + tce_j)\,\mathrm{d}t - (H_0 x)_j| \\
&\le \frac{1}{2}\int_{-1}^1 \|\nabla f(x + tce_j) - \nabla f(0)\|\,\mathrm{d}t + \|H_0\|\|x\| \\
&\le L\|x\| + \frac{Lc}{2} + \lambda_{\max}\|x\| = (L + \lambda_{\max})\|x\| + \frac{Lc}{2}.
\end{aligned}$$

Given that Z_s is a strong solution of the stochastic integral equation on $[0, \infty)$, we know that no explosion times exist and we further know that $Z_s \to x^\star = 0$ holds. By combining these two statements, we get $\|Z_s\| \le C(\omega) < \infty$. Furthermore, we have $c_s = c(1 + R_{s-})^{-\gamma} \le c$ and we find

$$|V^j(Z_{s-}, c_s)| \le (L + \lambda_{\max})\|Z_{s-}\| + \frac{L}{2}c_s \le (L + \lambda_{\max})C(\omega) + \frac{Lc}{2}.$$

Therefore we get

$$\begin{aligned}
|(1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_i} &\int_0^T (1 + R_{s-})^{a\lambda_i - 1} V^j(Z_{s-}, c_s)\,\mathrm{d}R_s| \\
&\le (1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^T (1 + R_{s-})^{a\lambda_i - 1}\left((L + \lambda_{\max})C(\omega) + \frac{Lc}{2}\right)\mathrm{d}R_s \\
&\le \underbrace{(1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_i}}_{\to 0} \underbrace{\left((L + \lambda_{\max})C(\omega) + \frac{Lc}{2}\right)}_{< \infty} \underbrace{\int_0^T (1 + R_{s-})^{a\lambda_i - 1}\,\mathrm{d}R_s}_{< \infty},
\end{aligned}$$

yielding almost sure convergence, thus convergence in probability of the left hand side. Then, we have

$$\underset{t_0^1}{\exists}\ \underset{t \ge t_0^1}{\forall}\quad \mathbf{P}\left[\left|(1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^T (1 + R_{s-})^{a\lambda_i - 1} V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \ge \frac{\epsilon_1}{2}\right] \le \frac{\epsilon_2}{2}$$

which ensures the first bound above. To verify the second bound "$\leq \frac{\epsilon_2}{2}$", we consider

$$
\begin{aligned}
&\mathbf{P}\left[\left|\int_{T+}^{t} \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \geq \frac{\epsilon_1}{2}\right] \\
&\leq \mathbf{P}\left(\left[\left|\int_{T+}^{t} \div V^j(Z_{s-}, c_s)\,\mathrm{d}R_s\right| \geq \frac{\epsilon_1}{2}\right] \cap \left[\sup_{t\geq T}\|Z_t\| < \rho\right]\right) + \mathbf{P}\left[\sup_{t\geq T}\|Z_t\| \geq \rho\right] \\
&\leq \mathbf{P}\left[\left|\int_{T+}^{t} \div \left(\frac{\epsilon_1\epsilon_2(a\lambda_i - \frac{p}{2})}{64\sqrt{K}}\|Z_{s-}\| + \frac{\epsilon_1(a\lambda_i - \gamma)}{8c}c_s\right)\mathrm{d}R_s\right| \geq \frac{\epsilon_1}{2}\right] + \mathbf{P}\left[\sup_{t\geq T}\|Z_t\| \geq \rho\right] \\
&\leq \mathbf{P}\left(\left[\left|\int_{T+}^{t} \div \frac{\epsilon_1\epsilon_2(a\lambda_i - \frac{p}{2})}{64\sqrt{K}}\|Z_{s-}\|\,\mathrm{d}R_s\right| \geq \frac{\epsilon_1}{4}\right] \cup \left[\left|\int_{T+}^{t} \div \frac{\epsilon_1(a\lambda_i - \gamma)}{8c}c_s\,\mathrm{d}R_s\right| \geq \frac{\epsilon_1}{4}\right]\right) \\
&\quad + \mathbf{P}\left[\sup_{t\geq T}\|Z_t\| \geq \rho\right] \\
&\leq \underbrace{\mathbf{P}\left[\int_{T+}^{t} \div \|Z_{s-}\|\,\mathrm{d}R_s \geq \frac{16\sqrt{K}}{\epsilon_2(a\lambda_i - \frac{p}{2})}\right]}_{\overset{(A1)}{\leq} \frac{\epsilon_2}{4}} + \underbrace{\mathbf{P}\left[\int_{T+}^{t} \div c_s\,\mathrm{d}R_s \geq \frac{2c}{a\lambda_i - \gamma}\right]}_{\overset{(A2)}{=} 0} + \frac{\epsilon_2}{4}.
\end{aligned}
$$

We need to show (A1) and (A2).
To (A1): We have

$$
\begin{aligned}
&\mathbf{P}\left[\left|\int_{T+}^{t} \div \|Z_{s-}\|\,\mathrm{d}R_s\right| \geq \frac{16\sqrt{K}}{\epsilon_2(a\lambda_i - \frac{p}{2})}\right] \\
&\leq \mathbf{P}\left[\underset{t\geq 0}{\forall} Z_t = Y_t\right]^C + \mathbf{P}\left(\left[\left|\int_{T+}^{t} \div \|Z_{s-}\|\,\mathrm{d}R_s\right| \geq \frac{16\sqrt{K}}{\epsilon_2(a\lambda_i - \frac{p}{2})}\right] \cap \left[\underset{t\geq 0}{\forall} Z_t = Y_t\right]\right) \\
&\leq \underbrace{\mathbf{P}\left[\underset{t\geq 0}{\forall} Z_t = Y_t\right]^C}_{\leq \frac{\epsilon_2}{8}} + \underbrace{\mathbf{P}\left[\left|\int_{T+}^{t} \div \|Y_{s-}\|\,\mathrm{d}R_s\right| \geq \frac{16\sqrt{K}}{\epsilon_2(a\lambda_i - \frac{p}{2})}\right]}_{\overset{(\star)}{\leq} \frac{\epsilon_2}{8}} \leq \frac{\epsilon_2}{4}.
\end{aligned}
$$

Now we verify $(\star)$. Using

$$
\mathbf{E}\|Y_t\|^2 \leq K(1+R_t)^{-p} \qquad \Longrightarrow \qquad \mathbf{E}\|Y_t\| \leq \sqrt{K}(1+R_t)^{-\frac{p}{2}}
$$

and Markov's inequality we get

$$
\begin{aligned}
\mathbf{P}\left[\left|\int_{T+}^{t} \div \|Y_{s-}\|\,\mathrm{d}R_s\right| \geq \frac{16\sqrt{K}}{\epsilon_2(a\lambda_i - \frac{p}{2})}\right] &\leq \frac{\epsilon_2(a\lambda_i - \frac{p}{2})}{16\sqrt{K}} \int_{T+}^{t} \div \underbrace{\mathbf{E}\|Y_{s-}\|}_{\leq \sqrt{K}(1+R_{s-})^{-\frac{p}{2}}}\,\mathrm{d}R_s \\
&\leq \frac{\epsilon_2(a\lambda_i - \frac{p}{2})}{16}(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i}\int_0^t (1+R_{s-})^{a\lambda_i - 1 - \frac{p}{2}}\,\mathrm{d}R_s
\end{aligned}
$$

for $p := \frac{1}{2}$ (or $p := \frac{2}{3}$ in (b)). With Itô's formula and an obvious Taylor expansion around R_{s-}, there exists some $\vartheta_s \in (0,1)$ with

$$\begin{aligned}
&\int_0^t (1+R_{s-})^{a\lambda_i-1-\frac{p}{2}}\,\mathrm{d}R_s \\
&= \frac{(1+R_t)^{a\lambda_i-\frac{p}{2}}-1}{a\lambda_i-\frac{p}{2}} - \sum_{s\le t}\left\{\frac{\Delta(1+R_s)^{a\lambda_i-\frac{p}{2}}}{a\lambda_i-\frac{p}{2}} - (1+R_{s-})^{a\lambda_i-1-\frac{p}{2}}\Delta R_s\right\} \\
&= \frac{(1+R_t)^{a\lambda_i-\frac{p}{2}}-1}{a\lambda_i-\frac{p}{2}} - (a\lambda_i-\frac{p}{2}-1)\sum_{s\le t}(1+R_{s-}+\vartheta_s\Delta R_s)^{a\lambda_i-\frac{p}{2}-2}(\Delta R_s)^2 \\
&\le \frac{(1+R_t)^{a\lambda_i-\frac{p}{2}}-1}{a\lambda_i-\frac{p}{2}} + \mathcal{C}\sum_{s\le t}(\Delta R_s)^2(1+R_{s-}+\vartheta_s\Delta R_s)^{-2-\frac{p}{2}}(1+R_{s-}+\vartheta_s\Delta R_s)^{a\lambda_i} \\
&\le \frac{(1+R_t)^{a\lambda_i-\frac{p}{2}}-1}{a\lambda_i-\frac{p}{2}} + \mathcal{C}\sum_{s\le t}(\Delta R_s)^2(1+R_{s-})^{-2-\frac{p}{2}}(1+R_s)^{a\lambda_i} \\
&= \frac{(1+R_t)^{a\lambda_i-\frac{p}{2}}-1}{a\lambda_i-\frac{p}{2}} + \mathcal{C}\sum_{s\le t}\Bigg(\underbrace{\frac{1+R_s}{1+R_{s-}}}_{=1+\frac{\Delta R_s}{1+R_{s-}}=1+\mathrm{o_b}(1)}\Bigg)^{a\lambda_i}(\Delta R_s)^2(1+R_{s-})^{a\lambda_i-2-\frac{p}{2}}. \qquad (48)
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
&\frac{\epsilon_2(a\lambda_i-\frac{p}{2})}{16}\int_0^t \frac{(1+R_{s-})^{a\lambda_i-1-\frac{p}{2}}}{(1+R_t)^{-\frac{1}{2}+\gamma+a\lambda_i}}\,\mathrm{d}R_s \\
&\qquad \le \frac{\epsilon_2}{16}\underbrace{(1+R_t)^{\frac{1}{2}-\frac{p}{2}-\gamma}}_{=1,\text{ as }\frac{p}{2}+\gamma=\frac{1}{2}} + \underbrace{\mathcal{C}\sum_{s\le t}\frac{(1+R_{s-})^{a\lambda_i-2-\frac{p}{2}}}{(1+R_t)^{-\frac{1}{2}+\gamma+a\lambda_i}}(\Delta R_s)^2}_{\to 0} \longrightarrow \frac{\epsilon_2}{16}.
\end{aligned}$$

If the implied convergence of the second summand above holds, then the bound $(\star)$ is valid for any sufficiently large t. To prove convergence of the sum, we use Kronecker's lemma and obtain with $\frac{p}{2}+\gamma=\frac{1}{2}$

$$\sum_{0<s}\frac{(1+R_{s-})^{a\lambda_i-2-\frac{p}{2}}}{(1+R_s)^{-\frac{1}{2}+\gamma+a\lambda_i}}(\Delta R_s)^2 \le \mathcal{C}\int_0^\infty a_s^2\Delta R_s\,\mathrm{d}R_s^d < \infty.$$

To (A2): For a sufficiently large t we have

$$(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i}\int_{T+}^t c(1+R_{s-})^{a\lambda_i-1-\gamma}\mathrm{d}R_s$$

$$\leq \frac{c}{a\lambda_i - \gamma} \underbrace{(1+R_t)^{\frac{1}{2}-2\gamma}}_{=1,\text{ as } \frac{1}{2}-2\gamma=0} + \mathcal{C} \underbrace{\sum_{s\leq t} (\Delta R_s)^2 \frac{(1+R_{s-})^{a\lambda_i-2-\gamma}}{(1+R_t)^{a\lambda_i+\gamma-\frac{1}{2}}}}_{\to 0 \text{ see above}}$$

$$< \frac{2c}{a\lambda_i - \gamma}.$$

Hence the desired bound (A2) is justified.

Verification of (ii): Now we prove

$$(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^t (1+R_{s-})^{a\lambda_i+\gamma-1} (M_j(\mathrm{d}s, Z_{s-}) - M_j(\mathrm{d}s, 0)) \xrightarrow{\mathbf{P}} 0.$$

We detect convergence in probability by using the Lenglart-Rebolledo inequality (compare [19], p. 66, Theorem 3). For the processes X_t and Y_t, which appear in the theorem, we choose

$$X_t := \int_0^t \frac{M_j(\mathrm{d}s, Z_{s-}) - M_j(\mathrm{d}s, 0)}{(1+R_{s-})^{1-a\lambda_i-\gamma}} \qquad \text{and} \qquad Y_t := \lceil X \rceil_t .$$

Y_t represents the predictable compensator of X_t^2, thus $X_t^2 - Y_t \in \mathcal{M}_{\mathrm{loc}}$. By Theorem 3 in [19] on p. 33 we have $\mathbf{E}X_\tau = \mathbf{E}Y_\tau$ for every stopping time τ. Now we apply the Lenglart-Rebolledo inequality to the processes X_t and Y_t. For arbitrary $\epsilon_1, \epsilon_2 > 0$ and sufficiently large t, we obtain

$$\mathbf{P}\left[\left|(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \int_0^t (1+R_{s-})^{a\lambda_i+\gamma-1} (M_j(\mathrm{d}s, Z_{s-}) - M_j(\mathrm{d}s, 0))\right| > \epsilon_1\right]$$
$$= \mathbf{P}\left[X_t^2 > \epsilon_1^2 (1+R_t)^{2a\lambda_i+2\gamma-1}\right] \leq \mathbf{P}\left[\sup_{s\leq t} X_s^2 > \epsilon_1^2 (1+R_t)^{2a\lambda_i+2\gamma-1}\right]$$
$$\leq \frac{b}{\epsilon_1^2 (1+R_t)^{2a\lambda_i+2\gamma-1}} + \mathbf{P}\left[Y_t \geq b\right] \overset{(\star)}{=} \frac{\epsilon_2}{2} + \frac{\epsilon_2}{2} = \epsilon_2.$$

Here we get $(\star)$ for the choice $b := \frac{\epsilon_1^2 \epsilon_2}{2} (1+R_t)^{2a\lambda_i+2\gamma-1} > 0$. The implication

$$(h_s^{jj}(Z_{s-}) - 2h_s^{jj}(Z_{s-}, 0) + h_s^{jj}(0)) \to 0 \quad \Longrightarrow \quad \frac{\int_0^t \frac{h_s^{jj}(Z_{s-}) - 2h_s^{jj}(Z_{s-},0) + h_s^{jj}(0)}{(1+R_{s-})^{2-2a\lambda_i-2\gamma}} \, \mathrm{d}R_s}{(1+R_t)^{2\gamma+2a\lambda_i-1}} \to 0$$

follows from Toeplitz's lemma.

Verification of (iii): The assumption $\sum_{0\leq s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} < \infty$ ensures

$$|H_t^3| \leq \mathcal{C} (1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i} \sum_l \sum_{s\leq t} \|Z_{s-}\|^2 \mathbb{1}_{\{a\lambda_l \Delta R_s = (1+R_{s-})\}}$$

$$\leq \underbrace{\mathcal{C}_\omega}_{<\infty} \underbrace{(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_i}}_{\to 0} \underbrace{\sum_l \sum_{0\leq s} \mathbb{1}_{\{a\lambda_l \Delta R_s = (1+R_{s-})\}}}_{<\infty},$$

and hence the required convergence.

(b) Case: f is three times differentiable at 0. We reconsider the proof of (a) to obtain the desired result. In part (a) we have $V^j(x,b) = \mathrm{o}(\|x\|) + \mathrm{o}(b)$ and hence

$$\underset{\rho>0}{\exists} \underset{\|x\|,b\leq\rho}{\forall} \quad |V^j(x,b)| \leq \frac{\epsilon_1\epsilon_2(a\lambda_i - \frac{p}{2})}{64K}\|x\| + \frac{\epsilon_1(a\lambda_i - \gamma)}{8c}b.$$

We see that the major difference is the process $V(Z_{s-}, c_s)$. Therefore we will consider this process in the case of three times differentiable f more closely. Afterwards we will use an analogous argument as in part (a). We seek a handy representation of $V(x,c)$, where $V(x,c)$ is given by

$$\left\{\frac{f(x+ce_i) - f(x-ce_i)}{2c}\right\}_{i\in\{1,\ldots,d\}} = H_0 x + V(x,c).$$

Using Taylor's formula, in which the remainder is represented as an integral, we find

$$\nabla f(x) = \nabla f(0) + H_0 x + \frac{1}{2} D^3 f(0)[x,x] + r(0,x) \qquad \text{with } r(0,x) = \mathrm{o}(\|x\|^2),$$

since ∇f is still two times differentiable at 0. The formula above results in

$$\begin{aligned}
\frac{f(x+ce_i) - f(x-ce_i)}{2c} &= \frac{1}{2c}\int_{-1}^{1} \langle ce_i, \nabla f(x+tce_i)\rangle \,\mathrm{d}t = \frac{1}{2}\int_{-1}^{1} \nabla_i f(x+tce_i)\,\mathrm{d}t \\
&= \frac{1}{2}\int_{-1}^{1} \Big(\nabla f(0) + H_0(x+tce_i) + \frac{1}{2}D^3 f(0)[(x+tce_i),(x+tce_i)] + r(0,x+tce_i)\Big)_i \mathrm{d}t \\
&= \frac{1}{2}\int_{-1}^{1} \left(H_0 x + \frac{1}{2}\sum_{l,k} \frac{\partial^2 \nabla f}{\partial x_l \partial x_k}(0)\left(x_l x_k + t^2 c^2 \delta_{il}\delta_{ik}\right) + r(0,x+tce_i)\right)_i \mathrm{d}t
\end{aligned}$$

and

$$\begin{aligned}
&\left\{\frac{f(x+ce_i) - f(x-ce_i)}{2c}\right\}_{i\in\{1,\ldots,d\}} \\
&\qquad = H_0 x + \frac{1}{2}\sum_{l,k} \frac{\partial^2 \nabla f}{\partial x_l \partial x_k}(0) x_l x_k + \frac{1}{6}c^2 \left(\frac{\partial^3 f}{(\partial x_i)^3}(0)\right)_{i\in\{1,\ldots,d\}} + \mathrm{o}(\|x\|^2) + \mathrm{o}(c^2).
\end{aligned}$$

With $A_3 := \left(\frac{\partial^3 f}{(\partial x_i)^3}(0)\right)_{i\in\{1,\ldots,d\}}$ or $A_3^i = \frac{\partial^3 f}{(\partial x_i)^3}(0)$, we obtain for $V(Z_{s-}, c_s)$

$$V(Z_{s-}, c_s) = \frac{1}{6}A_3 c_s^2 + \mathcal{O}(\|Z_{s-}\|^2) + \mathrm{o}(c_s^2) = \frac{1}{6}A_3 c_s^2 + \mathrm{o}(\|Z_{s-}\|) + \mathrm{o}(c_s^2) \tag{49}$$

$$= \frac{1}{6} A_3 c_s^2 + \tilde{V}(Z_{s-}, c_s).$$

Thus, compared to $\int_0^t \vdots V^j(Z_{s-}, c_s)\,\mathrm{d}R_s \xrightarrow{\mathbf{P}} 0$ in part (a), we have to prove

$$\int_0^t \vdots \tilde{V}^j(Z_{s-}, c_s)\,\mathrm{d}R_s \xrightarrow{\mathbf{P}} 0$$

and

$$a \sum_{j,k}^{d} (1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_k} T_{ik} T_{kj}^T \int_0^t (1 + R_{s-})^{a\lambda_k - 1} \frac{1}{6} A_3^j c_s^2\,\mathrm{d}R_s \xrightarrow{\mathbf{P}} \left(\frac{ac^2}{6} (aH_0 - \frac{1}{3} I)^{-1} A_3 \right)_i .$$

The first part follows from the work we have done in part (a). Here in the case where f is three times differentiable, we obtain the desired speed in (A1) since p equals $\frac{2}{3}$ instead of $\frac{1}{2}$ and in (A2) because we have c_s^2 instead of c_s. We note here that it is important to have $\mathrm{o}(c_s^2)$ and not only $\mathcal{O}(c_s^2)$. That is the reason why we separate the term including $\frac{1}{6} A_3 c_s^2$ above. Now we consider the second part and get, as we are interested in the asymptotic behavior,

$$\begin{aligned}(1 + R_t)^{\frac{1}{2}-\gamma-a\lambda_k} \int_0^t (1 + R_{s-})^{a\lambda_k - 1} c_s^2\,\mathrm{d}R_s &= c^2 (1 + R_t)^{\frac{1}{3}-a\lambda_k} \int_0^t (1 + R_{s-})^{a\lambda_k - \frac{4}{3}}\,\mathrm{d}R_s \\ \simeq \frac{c^2}{a\lambda_k - \frac{1}{3}} (1 + R_t)^{\frac{1}{3}-a\lambda_k} (1 + R_t)^{a\lambda_k - \frac{1}{3}} &= \frac{c^2}{a\lambda_k - \frac{1}{3}}.\end{aligned}$$

Furthermore we get

$$\frac{a}{6} \sum_{j,k}^{d} T_{ik} T_{kj}^T \frac{\int_0^t (1 + R_{s-})^{a\lambda_k - 1} c_s^2\,\mathrm{d}R_s}{(1 + R_t)^{a\lambda_k + \gamma - \frac{1}{2}}} A_3^j \xrightarrow{\mathbf{P}} \frac{ac^2}{6} \sum_{j,k}^{d} T_{ik} (aD_{kk} - \frac{1}{3} I_{kk})^{-1} T_{kj}^T A_3^j.$$

Here the limit is

$$\begin{aligned}\frac{ac^2}{6} \sum_{j,k} T_{ik} \left(aD_{kk} - \frac{1}{3} I_{kk} \right)^{-1} T_{kj}^T A_3^j &= \frac{ac^2}{6} \left(T (aD - \frac{1}{3} I)^{-1} T^T A_3 \right)_i \\ &= \frac{ac^2}{6} \left((aH_0 - \frac{1}{3} I)^{-1} A_3 \right)_i .\end{aligned}$$

Part (III): We will prove

$$(1 + R_t)^{\frac{1}{2}-\gamma} T \phi_t \frac{a}{2c} \int_0^t (1 + R_{s-})^{\gamma - 1} \phi_s^{-1} T^T M(\mathrm{d}s, 0) \xrightarrow{\mathcal{D}} N(0, \Sigma).$$

We restrict ourselves to sequences to prove the statement. We will show that, for an arbitrarily increasing sequence t_n with $t_n \uparrow \infty$, we have

$$(1 + R_{t_n})^{\frac{1}{2}-\gamma} T \phi_{t_n} \frac{a}{2c} \int_0^{t_n} (1 + R_{s-})^{\gamma - 1} \phi_s^{-1} T^T M(\mathrm{d}s, 0) \xrightarrow{\mathcal{D}} N(0, \Sigma).$$

The limit distribution does not depend on the choice of the explicit sequence t_n. Let t_n be arbitrary but fixed. To apply a central limit theorem, we consider the sequence M_s^n

$$M_s^n := (1+R_{t_n})^{\frac{1}{2}-\gamma} T\phi_{t_n} \frac{a}{2c} \int_0^{st_n} (1+R_{r-})^{\gamma-1}\phi_r^{-1}T^T \, M(\mathrm{d}r, 0)$$

which represents a sequence of locally square integrable martingales. We have to show

$$M_1^n \xrightarrow{\mathcal{D}} M \quad \text{where} \quad M \sim N(0,\Sigma).$$

For this purpose we use the vector-valued version of Theorem 4 on page 435 in [19]:

Let (X_t^n) be a sequence of locally square integrable martingales and S a non-empty set of $\mathbb{R}_+$. Further let the assumptions

$$(i) \qquad \underset{\delta\in(0,1]}{\forall} \ \underset{t\in S}{\forall} \quad x^T x \mathbb{1}_{[\|x\|>\delta]} * \nu_t^n \xrightarrow{\mathbf{P}} 0,$$

$$(ii) \qquad \underset{t\in S}{\forall} \qquad \lceil X^n \rceil_t \xrightarrow{\mathbf{P}} \lceil X \rceil_t,$$

then

$$X_t^n \xrightarrow{\mathcal{D}_f(S)} X \quad (n\to\infty).$$

Now we choose $S=\{1\}$ and prove assumptions (i) and (ii).

Verification of (i): ν^n (that is, here, ν^{M^n}) is the compensator of the jump measure μ^n of the process M_s^n. Since the assumptions of Theorem 3.2 only apply to the local martingale $M_t^\star := \int_0^t M(\mathrm{d}s,0)$, which appears in the stochastic integral equation, we should include only its compensator $\nu^{M^\star}$. To obtain this, we consider $x^T x$ and $\mathbb{1}_{[\|x\|>\delta]}$, which corresponds to $(\Delta M_s^n)^T(\Delta M_s^n)$ and $\mathbb{1}_{[\|\Delta M_s^n\|>\delta]}$ here. First to $(\Delta M_s^n)^T(\Delta M_s^n)$. We have

$$\Delta M_s^n = \frac{a}{2c}\frac{(1+R_{t_n})^{\frac{1}{2}-\gamma}}{(1+R_{st_n-})^{1-\gamma}} T\phi_{t_n}\phi_{st_n}^{-1}T^T\Delta M_{st_n}^\star,$$

and therefore

$$\begin{aligned}(\Delta M_s^n)^T(\Delta M_s^n) = \|\Delta M_s^n\|^2 &= \frac{a^2}{4c^2}\frac{(1+R_{t_n})^{1-2\gamma}}{(1+R_{st_n-})^{2-2\gamma}}\|T\phi_{t_n}\phi_{st_n}^{-1}T^T\Delta M_{st_n}^\star\|^2 \\ &\le \frac{a^2}{4c^2}\frac{(1+R_{t_n})^{1-2\gamma}}{(1+R_{st_n-})^{2-2\gamma}}\|T\phi_{t_n}\phi_{st_n}^{-1}T^T\|^2\|\Delta M_{st_n}^\star\|^2 \\ &\overset{(\star)}{\le} \mathcal{C}\,\frac{(1+R_{st_n-})^{2\gamma-2}(1+R_{st_n})^{2a\lambda_{\min}}}{(1+R_{t_n})^{2\gamma+2a\lambda_{\min}-1}}\Delta M_{st_n}^{\star T}\Delta M_{st_n}^\star.\end{aligned}$$

Here, $(\star)$ follows from $a\lambda_{\min} > 0$ and $\left(\frac{1+R_{st_n}}{1+R_{t_n}}\right)^{a\lambda_i} \leq \left(\frac{1+R_{st_n}}{1+R_{t_n}}\right)^{a\lambda_{\min}}$. Furthermore, we have

$$\begin{aligned}
[\|\Delta M^n_s\| > \delta] &= \left[\left\| \frac{a}{2c}\frac{(1+R_{t_n})^{\frac{1}{2}-\gamma}}{(1+R_{st_n-})^{1-\gamma}} T\phi_{t_n}\phi^{-1}_{st_n}T^T\Delta M^\star_{st_n}\right\| > \delta\right] \\
&= \left[\|T\phi_{t_n}\phi^{-1}_{st_n}T^T\Delta M^\star_{st_n}\| > \frac{2c\delta}{a}\frac{(1+R_{t_n})^{\gamma-\frac{1}{2}}}{(1+R_{st_n-})^{\gamma-1}}\right] \\
&\subset \left[\|\phi_{t_n}\phi^{-1}_{st_n}\|\|\Delta M^\star_{st_n}\| > \frac{2c\delta}{a\|T\|^2}\frac{(1+R_{t_n})^{\gamma-\frac{1}{2}}}{(1+R_{st_n-})^{\gamma-1}}\right] \\
&\subset \left[\|\Delta M^\star_{st_n}\| > \delta\mathcal{C}\,\frac{(1+R_{t_n})^{\gamma-\frac{1}{2}+a\lambda_{\min}}}{(1+R_{st_n-})^{\gamma-1}(1+R_{st_n})^{a\lambda_{\min}}}\right].
\end{aligned}$$

Thus it is sufficient to consider—instead of $[\|\Delta M^n_s\| > \delta]$ for all $\delta \in (0,1]$—the set

$$G^\epsilon_{s,t_n} := \left\{x \in \mathbb{R}^d \,\middle|\, \|x\| > \epsilon\frac{(1+R_{t_n})^{\gamma-\frac{1}{2}+a\lambda_{\min}}}{(1+R_{st_n-})^{\gamma-1}(1+R_{st_n})^{a\lambda_{\min}}}\right\}$$

for all $\epsilon \in (0,1]$. We find with the inequalities above, $S = \{1\}$, and $t = 1$

$$\begin{aligned}
x^Tx\mathbb{1}_{[\|x\|>\delta]} * \nu^{M^n}_t &= x^Tx\mathbb{1}_{[\|x\|>\delta]} * \nu^{M^n}_1 = \int_0^1\int_{\mathbb{R}^d} x^Tx\mathbb{1}_{[\|x\|>\delta]}\ \nu^{M^n}(\mathrm{d}s,\mathrm{d}x) \\
&\leq \mathcal{C}\int_0^{t_n}\int_{G^\epsilon_{s,t_n}} \frac{(1+R_{s-})^{2\gamma-2}(1+R_s)^{2a\lambda_{\min}}}{(1+R_{t_n})^{2\gamma+2a\lambda_{\min}-1}} x^Tx\ \nu^{M^\star}(\mathrm{d}s,\mathrm{d}x) \\
&\leq \mathcal{C}\,\frac{\int_0^{t_n}\frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}}\int_{G^\epsilon_{s,t_n}} x^Tx\ \nu^{M^\star}(\mathrm{d}s,\mathrm{d}x)}{(1+R_{t_n})^{2a\lambda_{\min}+2\gamma-1}} \xrightarrow{\mathbf{P}} 0 \qquad (n\to\infty).
\end{aligned}$$

Verification of (ii): We have to show $\lceil M^n\rceil_1 \xrightarrow{\mathbf{P}} \lceil M\rceil_1$, which is equivalent to

$$\lceil M^{n^i}, M^{n^j}\rceil_1 \xrightarrow{\mathbf{P}} \lceil M^i, M^j\rceil_1$$

for all $i,j \in \{1,\dots,d\}$. Because ϕ_t is a diagonal matrix, we obtain

$$\begin{aligned}
&\lceil\left(T\phi_{t_n}\phi_r^{-1}T^T\,M(\mathrm{d}r,0)\right)_i, \left(T\phi_{t_n}\phi_r^{-1}T^T\,M(\mathrm{d}r,0)\right)_j\rceil_s \\
&\quad= \lceil\sum_{l,k} T_{ik}\phi^{kk}_{t_n}\phi^{kk^{-1}}_s T^T_{kl}\,M_l(\mathrm{d}s,0), \sum_{u,v} T_{jv}\phi^{vv}_{t_n}\phi^{vv^{-1}}_s T^T_{vu}\,M_u(\mathrm{d}s,0)\rceil \\
&\quad= \sum_{l,k,v,u} T_{ik}\phi^{kk}_{t_n}\phi^{kk^{-1}}_s T^T_{kl}T_{jv}\phi^{vv}_{t_n}\phi^{vv^{-1}}_s T^T_{vu}\lceil\, M_l(\mathrm{d}s,0)\,,\, M_u(\mathrm{d}s,0)\,\rceil
\end{aligned}$$

$$= \sum_{l,k,v,u} T_{ik} \frac{\Pi_{0,t_n}^{a\lambda_k} \Pi_{0,t_n}^{a\lambda_v}}{(1+R_{t_n})^{a(\lambda_k+\lambda_v)}} T_{kl}^T T_{jv} \frac{(1+R_s)^{a(\lambda_v+\lambda_k)}}{\Pi_{0,s}^{a\lambda_v} \Pi_{0,s}^{a\lambda_k}} T_{vu}^T h_s^{lu}(0)\,\mathrm{d}R_s$$
$$= \sum_{k,v} T_{ik} T_{vj}^T \Pi_{s,t_n}^{a\lambda_k} \Pi_{s,t_n}^{a\lambda_v} \frac{(1+R_s)^{a(\lambda_v+\lambda_k)}}{(1+R_{t_n})^{a(\lambda_k+\lambda_v)}} \left(T^T h_s(0) T\right)_{kv} \mathrm{d}R_s$$

and

$$\lceil M^{n^i}, M^{n^j} \rceil_1$$
$$= \frac{a^2}{4c^2} \int_0^{t_n} \frac{(1+R_{s-})^{2\gamma-2}}{(1+R_{t_n})^{2\gamma-1}} \mathrm{d}\lceil \left(T\phi_{t_n}\phi_r^{-1} T^T M(\mathrm{d}r,0)\right)_i, \left(T\phi_{t_n}\phi_r^{-1} T^T M(\mathrm{d}r,0)\right)_j \rceil_s$$
$$= \frac{a^2}{4c^2} \sum_{k,v} T_{ik} T_{vj}^T \left(T^T h(0) T\right)_{kv} \frac{\int_0^{t_n} (1+\mathrm{o_b}(1))(1+R_{s-})^{2\gamma-2+a(\lambda_v+\lambda_k)}\,\mathrm{d}R_s}{(1+R_{t_n})^{2\gamma-1+a(\lambda_v+\lambda_k)}}$$
$$\simeq \frac{a^2}{4c^2} \sum_{k,v} T_{ik} T_{vj}^T \left(T^T h(0) T\right)_{kv} \underbrace{\frac{\int_0^{t_n} (1+R_{s-})^{2\gamma-2+a(\lambda_v+\lambda_k)}\,\mathrm{d}R_s}{(1+R_{t_n})^{2\gamma-1+a(\lambda_k+\lambda_v)}}}_{\xrightarrow{n\to\infty} \frac{1}{2\gamma+a(\lambda_k+\lambda_v)-1}}$$
$$\longrightarrow \sum_{k,v} T_{ik} \underbrace{\frac{a^2\left(T^T h(0) T\right)_{kv}}{4c^2(2\gamma+a(\lambda_k+\lambda_v)-1)}}_{:= U_{kv}} T_{vj}^T = \sum_{k,v} T_{ik} U_{kv} T_{vj}^T$$
$$= \left(T\,U\,T^T\right)_{ij} = \Sigma_{ij}$$

where we used $\simeq$ to indicate that the expressions on both sides of this sign are asymptotically identical. □

Proof of Lemma 3.4

Considering the assumptions and the conclusion of Lemma 3.4, we immediately see that we have to change only **Part (III)** in the proof of Theorem 3.2. To prove convergence, we use a recent result of Crimaldi and Pratelli. Their Theorem 2.2 in [2] says:

On a probability space $(\Omega, \mathcal{A}, \mathbf{P})$, endowed with a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions, let $\tilde{M} = (\tilde{M}_t)_{t\geq 0}$ be a (right-continuous with limits from the left) d-dimensional martingale. Further let $(\tilde{a}_t)_{t\geq 0}$ be a family of invertible $d \times d$-matrices. Let us suppose that the following conditions hold (as $t \to \infty$):

(a) $\quad \|\tilde{a}_t\| \to 0$

(b) $\quad \mathbf{E}\left(\sup_{0\leq s\leq t} \|\tilde{a}_t \Delta \tilde{M}_s\|\right) \to 0$

(c) $\qquad \tilde{a}_t[\tilde{M}]_t\tilde{a}_t^T \xrightarrow{\mathbf{P}} \Sigma.$

Then the random vector $\tilde{a}_t M_t$ converges $\mathcal{A}$-stably to the Gaussian kernel $N(0,\Sigma)$.

It is an easy task to generalize Theorem 2.2 to the case of a local martingale. To apply the theorem above, we consider

$$\tilde{a}_t := \frac{a}{2c}(1+R_t)^{\frac{1}{2}-\gamma}T\phi_t \qquad \tilde{M}_t := \int_0^t (1+R_{s-})^{\gamma-1}\phi_s^{-1}T^T M(\mathrm{d}s,0)$$

and verify the assumptions (a), (b), and (c) of the theorem of Crimaldi and Pratelli. Then, Lemma 3.4 follows from this theorem.

Verification of (a): According to the definition of $\tilde{a}_t$, the already calculated ϕ_t, and $\lambda_{\min} > \frac{1-2\gamma}{2a}$, we have

$$\|\tilde{a}_t\|^2 \le \mathcal{C}(1+R_t)^{1-2\gamma}\|\phi_t\|^2 \le \mathcal{C}(1+R_t)^{1-2\gamma}(1+R_t)^{-2a\lambda_{\min}} \to 0$$

and hence $\|\tilde{a}_t\| \to 0$.

Verification of (b): We consider the squared process and obtain

$$\begin{aligned}
&\left(\mathbf{E}\sup_{0\le s\le t}\|\tilde{a}_t\Delta\tilde{M}_s\|\right)^2 \le \mathbf{E}\sup_{0\le s\le t}\|\tilde{a}_t\Delta\tilde{M}_s\|^2 \le \mathbf{E}\sup_{0\le s\le t}\sum_{i=1}^d\left(\sum_{j=1}^d \tilde{a}_t^{ij}\Delta\tilde{M}_s^j\right)^2 \\
&\le \mathbf{E}\sup_{0\le s\le t}\sum_{i=1}^d\left(\sum_{j=1}^d \underbrace{\left(\frac{1+R_s}{1+R_t}\right)^{a\lambda_j}}_{\le \left(\frac{1+R_s}{1+R_t}\right)^{a\lambda_{\min}},\ \text{as } \frac{1+R_s}{1+R_t}\in(0,1]} \frac{\Pi_{0,t}^{a\lambda_j}}{\Pi_{0,s}^{a\lambda_j}} T_{ij}\frac{(1+R_t)^{\frac{1}{2}-\gamma}}{(1+R_{s-})^{1-\gamma}}\sum_{l=1}^d T_{jl}^T|\Delta M_s^{\star^l}|\right)^2 \\
&\le \mathbf{E}\sup_{0\le s\le t}\sum_{i=1}^d\left(\sum_{j=1}^d \underbrace{\Pi_{s,t}^{a\lambda_j}\left(1+\frac{\Delta R_s}{1+R_{s-}}\right)^{a\lambda_{\min}}}_{\le \mathcal{C}} T_{ij}\frac{(1+R_t)^{\frac{1}{2}-\gamma-a\lambda_{\min}}}{(1+R_{s-})^{1-\gamma-a\lambda_{\min}}}\sum_{l=1}^d T_{jl}^T|\Delta M_s^{\star^l}|\right)^2 \\
&\le \mathcal{C}\,\mathbf{E}\sup_{0\le s\le t}\sum_{i=1}^d\left(\frac{(1+R_{s-})^{a\lambda_{\min}+\gamma-1}}{(1+R_t)^{a\lambda_{\min}+\gamma-\frac{1}{2}}}\sum_{l=1}^d \underbrace{\sum_{j=1}^d T_{ij}T_{jl}^T}_{=(TT^T)_{il}=I_{il}=\delta_{il}}|\Delta M_s^{\star^l}|\right)^2 \\
&\le \mathcal{C}\,\frac{\mathbf{E}\sup_{0\le s\le t}(1+R_{s-})^{2a\lambda_{\min}+2\gamma-2}\|\Delta M_s^{\star}\|^2}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \xrightarrow{t\to\infty} 0.
\end{aligned}$$

Verification of (c): We will show

$$(i)\ \ \tilde{a}_t\lceil\tilde{M}\rceil_t\tilde{a}_t^T \xrightarrow{\mathbf{P}} \Sigma \qquad \text{and} \qquad (ii)\ \ \tilde{a}_t\left([\tilde{M}]_t - \lceil\tilde{M}\rceil_t\right)\tilde{a}_t^T \xrightarrow{\mathbf{P}} 0.$$

According to (a) we have $\|\tilde{a}_t\| \to 0$ and hence we can use the argumentation (46). To this end substitution of the process $(1+R_s)^{a\lambda_i}$ by the process $(1+R_{s-})^{a\lambda_i}$ does not effect the asymptotic result.
To (i): Let us now examine the matrix-valued process $\lceil \tilde{M} \rceil_t$

$$\begin{aligned}
\lceil \tilde{M} \rceil_t^{ij} &= \int_0^t (1+R_{s-})^{2\gamma-2} \lceil \left(\phi^{-1} T^T M(\mathrm{d}r,0)\right)_i, \left(\phi^{-1} T^T M(\mathrm{d}r,0)\right)_j \rceil_s \\
&= \int_0^t \frac{(1+R_{s-})^{2\gamma-2+a\lambda_i+a\lambda_j}}{\Pi_{0,s}^{a\lambda_i}\Pi_{0,s}^{a\lambda_j}} \sum_{l,k} T^T_{ik} h_s^{kl}(0) T_{lj}\,\mathrm{d}R_s \\
&= \int_0^t \frac{(1+R_{s-})^{2\gamma-2+a\lambda_i+a\lambda_j}}{\Pi_{0,s}^{a\lambda_i}\Pi_{0,s}^{a\lambda_j}} (T^T h_s(0) T)_{ij}\,\mathrm{d}R_s. \qquad (50)
\end{aligned}$$

Thus, because ϕ is a diagonal matrix, we obtain

$$\begin{aligned}
\frac{a^2}{4c^2}(1+R_t)^{1-2\gamma}\left(\phi_t \lceil \tilde{M} \rceil_t \phi_t^T\right)_{ij} &= \frac{a^2}{4c^2}(1+R_t)^{1-2\gamma}\phi_t^{ii} \lceil \tilde{M} \rceil_t^{ij} \phi_t^{jj} \\
&= \frac{a^2}{4c^2} \frac{\int_0^t \Pi_{s,t}^{a\lambda_i}\Pi_{s,t}^{a\lambda_j}(1+R_{s-})^{2\gamma-2+a\lambda_i+a\lambda_j}(T^T h_s(0) T)_{ij}\,\mathrm{d}R_s}{(1+R_t)^{2\gamma-1+a\lambda_i+a\lambda_j}} \\
&= \frac{a^2}{4c^2}(T^T(0)T)_{ij} \underbrace{\frac{\int_0^t (1+\mathrm{o_b}(1))(1+R_{s-})^{2\gamma-2+a\lambda_i+a\lambda_j}\,\mathrm{d}R_s}{(1+R_t)^{2\gamma-1+a\lambda_i+a\lambda_j}}}_{\to \frac{1}{a\lambda_i+a\lambda_j+2\gamma-1}}.
\end{aligned}$$

Consequently we have

$$\tilde{a}_t \lceil \tilde{M} \rceil_t \tilde{a}_t^T \xrightarrow{\mathbf{P}} \Sigma \quad \text{where } U_{ij} := \frac{a^2 (T^T h(0) T)_{ij}}{4c^2(a(\lambda_i+\lambda_j)+2\gamma-1)} \quad \text{and} \quad \Sigma = TUT^T.$$

To (ii): Using the calculation of $\lceil \tilde{M} \rceil_t^{ij}$ and proceeding with $[\tilde{M}]_t^{ij}$ analogously, we get, for all $i,j \in \{1,\dots,n\}$,

$$\begin{aligned}
&\left(\tilde{a}_t \left([\tilde{M}]_t - \lceil \tilde{M} \rceil_t\right) \tilde{a}_t^T\right)_{ij} \\
&= \sum_{l,k,m,n} T^T_{lm} T^T_{kn} T_{il} T^T_{kj} \int_0^t (1+\mathrm{o_b}(1)) \frac{(1+R_t)^{1-2\gamma-a(\lambda_l+\lambda_k)}}{(1+R_{s-})^{2-2\gamma-a(\lambda_l+\lambda_k)}}\,\mathrm{d}\Big(\overbrace{[M^\star]_s^{mn} - \lceil M^\star \rceil_s^{mn}}^{=:\, L_s^{mn}}\Big).
\end{aligned}$$

Then, it is sufficient to verify

$$(1+R_t)^{1-2\gamma-a(\lambda_l+\lambda_k)} \int_0^t (1+R_{s-})^{2\gamma-2+a(\lambda_l+\lambda_k)}\,\mathrm{d}L_s^{mn} \xrightarrow{\mathbf{P}} 0, \qquad (51)$$

because the term $\mathrm{o_b}(1)$, appearing above in $(1+\mathrm{o_b}(1))$, does not effect the asymptotic result. Using the definition of a compensator, we get $L_t^{mn} \in \mathcal{M}_{\mathrm{loc}}$, because $\lceil M^\star \rceil_t$ is the compensator of $[M^\star]_t$. Now we use the Davis inequality ([19], p. 70, Theorem 6), which states

$$c\,\mathbf{E}\sqrt{[M]_T} \le \mathbf{E}\sup_{s\le T}|M_s| \le C\,\mathbf{E}\sqrt{[M]_T}. \tag{52}$$

To verify (51) it suffices to show

$$(1+R_t)^{1-2\gamma-a(\lambda_l+\lambda_k)}\mathbf{E}\sup_{s\le t}\left|\int_0^s (1+R_{r-})^{2\gamma-2+a(\lambda_l+\lambda_k)}\,\mathrm{d}L_r^{mn}\right| \to 0.$$

To this end we apply (52) in a first step and find

$$\begin{aligned}
&(1+R_t)^{1-2\gamma-a(\lambda_l+\lambda_k)}\,\mathbf{E}\sup_{s\le t}\left|\int_0^s (1+R_{r-})^{a(\lambda_l+\lambda_k)-2+2\gamma}\,\mathrm{d}L_r^{mn}\right| \\
&\overset{(52)}{\le} \mathcal{C}\,(1+R_t)^{1-2\gamma-a(\lambda_l+\lambda_k)}\mathbf{E}\sqrt{\int_0^t (1+R_{s-})^{2a(\lambda_l+\lambda_k)-4+4\gamma}\,\mathrm{d}\,[L]_s^{mn}} \\
&\overset{(\star)}{\le} \mathcal{C}\,\frac{\mathbf{E}\sqrt{\sum_{s\le t}(\Delta M_s^{\star^m}\Delta M_s^{\star^n})^2(1+R_{s-})^{2a(\lambda_l+\lambda_k)+4\gamma-4}}}{(1+R_t)^{a(\lambda_l+\lambda_k)+2\gamma-1}} \\
&\quad + \mathcal{C}\sqrt{\frac{\int_0^t(1+R_{s-})^{2a(\lambda_l+\lambda_k)+4\gamma-4}h_s^{mn}(0)^2\Delta R_s\,\mathrm{d}R_s}{(1+R_t)^{2a(\lambda_l+\lambda_k)+4\gamma-2}}}.
\end{aligned} \tag{53}$$

Here we get $(\star)$ by $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$ (if $a, b \ge 0$) and

$$\begin{aligned}
[L]_t^{mn} &= \sum_{s\le t}(\Delta L_s^{mn})^2 = \sum_{s\le t}(\Delta[M^\star]_s^{mn} - \Delta\lceil M^\star\rceil_s^{mn})^2 \\
&= \sum_{s\le t}\left(\Delta M_s^{\star^m}\Delta M_s^{\star^n} - \Delta\int_0^s h_r^{mn}(0)\,\mathrm{d}R_r\right)^2 \\
&= \sum_{s\le t}\left(\Delta M_s^{\star^m}\Delta M_s^{\star^n} - h_s^{mn}(0)\Delta R_s\right)^2 \le 2\sum_{s\le t}\left((\Delta M_s^{\star^m}\Delta M_s^{\star^n})^2 + (h_s^{mn}(0)\Delta R_s)^2\right) \\
&= 2\sum_{s\le t}(\Delta M_s^{\star^m}\Delta M_s^{\star^n})^2 + 2\int_0^t h_s^{mn}(0)^2\Delta R_s\,\mathrm{d}R_s.
\end{aligned}$$

Now we utilize a generalization of Kronecker's lemma ([19], Lemma 3, p. 141) to verify the convergence of the second summand in (53). This version of Kronecker's lemma says:

Let X be a semimartingale and L be a predictable increasing process, then we have

$$\left\{L_\infty = \infty\right\} \cap \left\{\int \frac{1}{1+L_s}\,\mathrm{d}X_s \to\right\} \subseteq \left\{\frac{X}{L} \to 0\right\}.$$

From the lemma and

$$\int_0^t \frac{(1+R_{s-})^{2a(\lambda_l+\lambda_k)+4\gamma-4}h_s^{mn}(0)^2\Delta R_s}{1+(1+R_{s-})^{2a(\lambda_l+\lambda_k)+4\gamma-2}}\,\mathrm{d}R_s \le \mathcal{C}\int_0^t \frac{\Delta R_s}{(1+R_{s-})^2}\,\mathrm{d}R_s$$

$$= \mathcal{C} \int_0^t a_s^2 \Delta R_s \, \mathrm{d}R_s < \infty$$

the desired convergence result follows. To verify convergence of the first summand in (53), we utilize Jensen's inequality and obtain

$$\left(\frac{\mathbf{E}\sqrt{\sum_{s\le t}(\Delta M_s^{\star m}\Delta M_s^{\star n})^2(1+R_{s-})^{2a(\lambda_l+\lambda_k)+4\gamma-4}}}{(1+R_t)^{a(\lambda_l+\lambda_k)+2\gamma-1}} \right)^2$$

$$\le \frac{\mathbf{E}\sum_{s\le t}(\Delta M_s^{\star m}\Delta M_s^{\star n})^2(1+R_{s-})^{4a\lambda_{\min}+4\gamma-4}}{(1+R_t)^{4a\lambda_{\min}+4\gamma-2}} \xrightarrow{t\to\infty} 0.$$

□

3.3 Discussion of Special Settings

Here we specialize Theorem 3.2 to the discrete-time setting and the Itô setting. It turns out that one result, presented here as a corollary, coincides with a known result in the literature dealing with the discrete-time setting. In the Itô setting, we obtain a new result.

Corollary 3.1. *We consider the stochastic integral equation* (29) *of the Itô type as given in Corollary 2.1 and 2.3, respectively, and assume that the conditions* (A), (B), (C), (E'),

$$\sigma_s^{ij}(x) \le C(1+\|x\|), \qquad \text{and} \qquad \lim_{\substack{s\to\infty \\ x\to x^\star}} \sigma_s^{ij}(x) = \sigma^{ij}(x^\star)$$

hold for all $i,j \in \{1,\dots,d\}$.

(a) If f is twice differentiable at $x^\star$, Hessian is continuous around $x^\star$, $\gamma = \frac{1}{4}$, and $\lambda_{\min} > \frac{1}{4a}$ holds, we get

$$(1+t)^{\frac{1}{4}}(Z_t - x^\star) \xrightarrow{\mathcal{D}} N(0,\Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma = TUT^T$, and

$$U_{kv} := \frac{a^2 \left(T^T\sigma(x^\star)\sigma^T(x^\star)T\right)_{kv}}{4c^2(a(\lambda_k+\lambda_v)-\frac{1}{2})}.$$

(b) If f is three times differentiable at $x^\star$, Hessian is continuous around $x^\star$, $\gamma = \frac{1}{6}$, and $\lambda_{\min} > \frac{1}{3a}$ holds, we get

$$(1+t)^{\frac{1}{3}}(Z_t - x^\star) \xrightarrow{\mathcal{D}} N(\mu,\Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma = TUT^T$,

$$\mu := -\frac{ac^2}{6}(aH_{x^\star} - \frac{1}{3}I)^{-1}\left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star)\right)_{i\in\{1,\ldots,d\}},$$

and

$$U_{kv} := \frac{a^2\left(T^T\sigma(x^\star)\sigma^T(x^\star)T\right)_{kv}}{4c^2(a(\lambda_k+\lambda_v)-\frac{2}{3})}.$$

Figures 3 and 4 show the density of the limit distribution of the processes $(1+t)^{1/4}(Z_t - x^\star)$ and $(1+t)^{1/3}(Z_t - x^\star)$ considered in part (a) and part (b) of Corollary 3.1. In both figures, the histogram is constructed by an evaluation of paths of the corresponding processes at time $t = 250$. The density of the theoretical limit distribution as given in Corollary 3.1 are presented as well. We choose $f(x) = 0.5x^T x$, $x \in \mathbb{R}^2$, $a_s = (1+s)^{-1}$, $\sigma_s(\cdot) = I$, $c_s = (1+s)^{-1/4}$ in Figure 3, $c_s = (1+s)^{-1/6}$ in Figure 4 and use the Milstein scheme (see, e.g., [14]) to simulate paths of the process Z_t. The starting points are taken from a uniform distribution on $[-4,4]\times[-4,4]$.

Corollary 3.2. *We consider the recursion* (30) *of Corollary 2.2 and 2.4, respectively, and assume that the conditions* (A), (B), (C), (E''),

(E''') $\qquad \sup_n \mathbf{E}\left(\|V_n\|^2|\mathcal{F}_{n-1}\right) < \infty,$

(F) $\qquad \mathbf{E}\left(V_n^i V_n^j|\mathcal{F}_{n-1}\right) \overset{n\to\infty}{\longrightarrow} h^{ij} \quad (1 \le i,j \le d),$

and the Lyapunov-type condition

(L) $\qquad \sup_n \mathbf{E}\|V_n\|^{2+\delta} < \infty$ *for a* $\delta > 0$

hold.

(a) If f *is twice differentiable at* $x^\star$*, Hessian is continuous around* $x^\star$*,* $\gamma = \frac{1}{4}$*, and* $\lambda_{\min} > \frac{1}{4a}$ *holds, we get*

$$n^{\frac{1}{4}}(Z_n - x^\star) \overset{\mathcal{D}}{\longrightarrow} N(0,\Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma = TUT^T$, *and*

$$U_{kv} := \frac{a^2(T^ThT)_{kv}}{a(\lambda_k+\lambda_v)-\frac{1}{2})}.$$

(b) If f is three times differentiable at $x^\star$, Hessian is continuous around $x^\star$, $\gamma = \frac{1}{6}$, and $\lambda_{\min} > \frac{1}{3a}$ holds, we get

$$n^{\frac{1}{3}}(Z_n - x^\star) \xrightarrow{\mathcal{D}} N(\mu, \Sigma),$$

where $T^T H_{x^\star} T = D$, $\Sigma = TUT^T$,

$$\mu := -\frac{ac^2}{6}(aH_{x^\star} - \frac{1}{3}I)^{-1}\left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star)\right)_{i\in\{1,\ldots,d\}}, \text{ and } U_{kv} := \frac{a^2(T^T hT)_{kv}}{4c^2(a(\lambda_k + \lambda_v) - \frac{2}{3})}.$$

We mention here that condition (L) is stronger than the first part of condition (E'') appeared in Section 2.3.

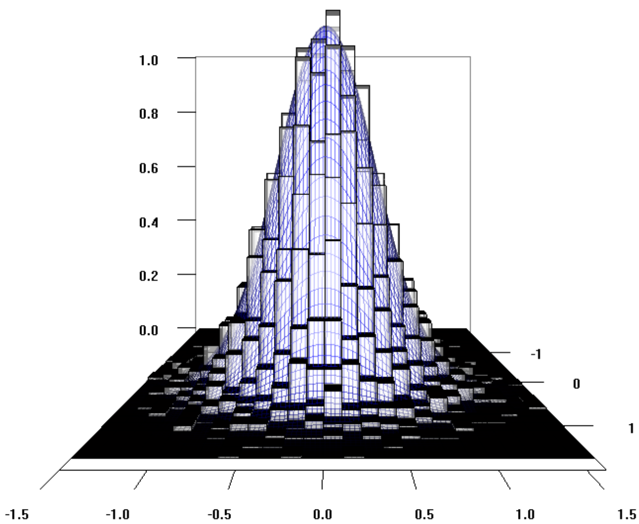

Figure 3: This plot shows the histogram as an empirical approximation of the density of the limit distribution of the process $(1+t)^{1/4}Z_t$ considered in Corollary 3.1 (a) and the density of the theoretical asymptotic distribution overlaid in blue.

Remark 3.4. We compare part (a) of Theorem 3.2 with Theorem 5.17. and part (b) with Remark 5.18 on page 38 in [20]. There $\frac{a^2\rho^2}{c^2(2af''(x^\star)-\frac{1}{2})}$ and $\frac{a^2\rho^2}{c^2(2af''(x^\star)-\frac{2}{3})}$ are given as variance, 0 and $-\frac{ac^2 f'''(x^\star)}{6af''(x^\star)-2}$ as bias of the limit distribution, respectively. It is important to mention that there an algorithm including the term "$-2\frac{a_n}{2c_n}V_n$" is considered. Consequently, to compare the variances we have to multiply them by the factor $\frac{1}{4}$.

Concerning the 1-dimensional situation we have $T = 1$ and $\lambda_k + \lambda_v = 2f''(x^\star)$ in the corollary above. Therefore, we have shown that our corollary coincides with known results in the 1-dimensional situation.

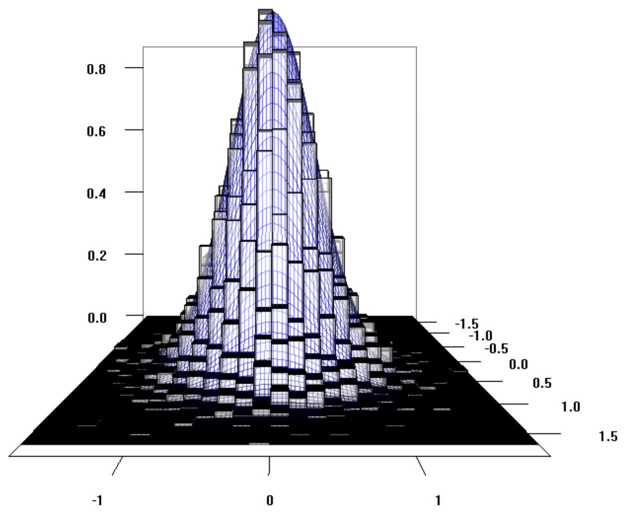

Figure 4: This plot shows the histogram as an empirical approximation of the density of the limit distribution of the process $(1+t)^{1/3}Z_t$ considered in Corollary 3.1 (b) and the density of the theoretical asymptotic distribution overlaid in blue.

Proof of Corollary 3.1

We will verify the assumptions of Theorem 3.2. Since (A), (B), (C), and (E') are valid, the assumptions of Corollary 2.3 are also valid. This ensures the existence of a strictly positive monotonously increasing process S_t with $S_t \uparrow \infty$ and $S_t\|Z_t\| \to 0$. We reuse the notation of Corollary 2.1, Corollary 2.3 and their proofs. Since the Brownian motion is continuous, it certainly fulfills the Lindeberg condition, because

$$\mu^W([0,t]\times\Gamma) = \sum_{0<s\leq t} \mathbb{1}_{\{\Delta W_s \in \Gamma\}} = \sum_{0<s\leq t} \mathbb{1}_{\{0\in\Gamma\}} = \sum_{0<s\leq t} 0 = 0 \quad \Gamma \in \mathcal{B}_d(\mathbb{R}\setminus\{0\}) \quad t \in \mathbb{R}_+,$$

which results in $\nu^W([0,t]\times\Gamma) = 0$ for all $\Gamma \in \mathcal{B}_d(\mathbb{R}\setminus\{0\})$ and $t \in \mathbb{R}_+$. Since $R_s := s$ is continuous, we have

$$\Delta R_s = \Delta s = 0 \quad \Longrightarrow \quad \sum_{0\leq s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} < \infty.$$

Considering

$$\lceil \int_0^\cdot M_i(\mathrm{d}s, J_s), \int_0^\cdot M_j(\mathrm{d}s, J_s) \rceil_t = \int_0^t \sum_{l,k} \sigma_s^{il}(J_s)\sigma_s^{jk}(J_s)\,\mathrm{d}\lceil W^l, W^k \rceil_s$$
$$= \int_0^t \sum_{l,k} \sigma_s^{il}(J_s)\sigma_s^{jk}(J_s)\,\delta_{lk}\,\mathrm{d}s = \int_0^t \sum_{k=1}^d \sigma_s^{ik}(J_s)\sigma_s^{jk}(J_s)\,\mathrm{d}s$$

we obtain $h_s^{ij}(x) = \sum_{k=1}^d \sigma_s^{ik}(x)\sigma_s^{jk}(x)$ and thus

$$\lim_{\substack{s\to\infty\\ x\to x^\star}} h_s^{ij}(x) = \sum_{k=1}^d \lim_{\substack{s\to\infty\\ x\to x^\star}} \left(\sigma_s^{ik}(x)\sigma_s^{jk}(x)\right) = \sum_{k=1}^d \sigma^{ik}(x^\star)\sigma^{jk}(x^\star)$$
$$= \left(\sigma(x^\star)\sigma^T(x^\star)\right)_{ij} =: h^{ij}(x^\star).$$

Furthermore by $\|x\| \leq S$ we have

$$|h_s^{ij}(x)| = |\sum_{k=1}^d \sigma_s^{ik}(x)\sigma_s^{jk}(x)| \leq \sum_{k=1}^d |\sigma_s^{ik}(x)||\sigma_s^{jk}(x)| \leq dC^2(1+\|x\|)^2 \leq \mathcal{C}.$$

□

Proof of Corollary 3.2

We will verify the assumptions of Theorem 3.2. Since (A), (B), (C), and (E'') are valid, the assumptions of Corollary 2.4 are also valid. This ensures the existence of a strictly positive monotonously increasing process S_t with $S_t \uparrow \infty$ and $S_t\|Z_t\| \to 0$. We reuse the notation from Corollary 2.2, Corollary 2.4 and their proofs. The quadratic variation gives

$$\left[\int_0^\cdot M_i(\mathrm{d}s, x)\right]_t = \left[\int_0^\cdot \tilde{V}_s^i\,\mathrm{d}R_s\right]_t = \sum_{\substack{n\leq t\\ n\in\mathbb{N}}} V_n^{i^2}(\Delta R_n)^2 = \sum_{\substack{n\leq t\\ n\in\mathbb{N}}} V_n^{i^2}$$

and hence the predictable quadratic variation satisfies

$$\left\lceil\int_0^\cdot M_i(\mathrm{d}s, x)\right\rceil_t = \sum_{\substack{n\leq t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2}|\mathcal{F}_{n-1}\right) = \sum_{\substack{n\leq t\\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2}|\mathcal{F}_{n-1}\right).$$

We have $h_n^{ij} := \mathbf{E}\left(V_n^i V_n^j|\mathcal{F}_{n-1}\right)$ and we obtain convergence and boundedness of h_n by

$$h_n^{ij} = \mathbf{E}\left(V_n^i V_n^j|\mathcal{F}_{n-1}\right) \overset{n\to\infty}{\longrightarrow} h^{ij} \quad\text{and}\quad \sup_n h_n^{ii} = \sup_n \mathbf{E}\left(V_n^{i^2}|\mathcal{F}_{n-1}\right) < \infty.$$

Since no argument occurs at h_n^{ii}, we get

$$\mathbf{E}\left(h_s^{ii}(Z_{s-}) - 2h^{ii}(Z_{s-},0) + h_s^{ii}(0)\right) = \mathbf{E}\left(h_s^{ii} - 2h_s^{ii} + h_s^{ii}\right) = 0.$$

Furthermore, we have

$$\mathbf{E}\left[\int_0^\cdot \frac{a_s}{c_s} M_i(\mathrm{d}s, x)\right]_t \le \mathbf{E} \sum_{\substack{n\le t\\ n\in\mathbb{N}}} V_n^{i2} = \sum_{\substack{n\le t\\ n\in\mathbb{N}}} \mathbf{E} V_n^{i2} \le \left(\sup_n \mathbf{E}\|V_n\|^2\right) \lfloor t\rfloor < \infty$$

and

$$\sum_{0\le s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} = \sum_{j=1}^{\infty} \mathbb{1}_{\{a\lambda_i = j\}} \le \sum_{j=1}^{\lfloor a\lambda_{\max}\rfloor} \le a\lambda_{\max} < \infty.$$

Now we show that the condition of the Lindeberg type holds. Considering a triangle schemata, it is known that the Lyapunov-type condition implies the Lindeberg-type condition. Now we prove the Lindeberg-type condition using assumption (L), which is a Lyapunov-type condition. We have

$$\begin{aligned}
&\frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{G_{s,t}^\epsilon} x^T x\ \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x)}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \le \frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{\mathbb{R}^d} x^T x\ \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x)}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \\
&= \frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2\, N_s(\omega, \mathrm{d}x)\, \mathrm{d}C_s}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \le \mathcal{C} \frac{\int_0^t \frac{(1+R_s)^{2a\lambda_{\min}}}{(1+R_{s-})^{2-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2\, N_s(\omega, \mathrm{d}x)\, \mathrm{d}R_s}{(1+R_t)^{2a\lambda_{\min}+2\gamma-1}} \\
&\le \mathcal{C} \frac{\sum_{i=1}^{\lfloor t\rfloor} \frac{(1+i)^{2a\lambda_{\min}}}{i^{2-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2\ N_i(\omega, \mathrm{d}x)}{(1+\lfloor t\rfloor)^{2a\lambda_{\min}+2\gamma-1}} \le \mathcal{C} \frac{\sum_{i=1}^{\lfloor t\rfloor} \frac{i^{2a\lambda_{\min}}}{i^{2-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2\ \mathbf{P}_{V_i|\mathcal{F}_{i-1}}(\mathrm{d}x)}{\lfloor t\rfloor^{2a\lambda_{\min}+2\gamma-1}} \\
&\le \mathcal{C} \frac{\sum_{i=1}^{\lfloor t\rfloor} i^{2a\lambda_{\min}+2\gamma-2} \mathbf{E}\left(\|V_i\|^2 | \mathcal{F}_{i-1}\right)}{\lfloor t\rfloor^{2a\lambda_{\min}+2\gamma-1}} = \mathcal{C} \frac{\sum_{i=1}^{n} i^{2a\lambda_{\min}+2\gamma-2} \mathbf{E}\left(\|V_i\|^2 | \mathcal{F}_{i-1}\right)}{n^{2a\lambda_{\min}+2\gamma-1}}
\end{aligned} \tag{54}$$

for $t = n \in \mathbb{N}$. Here we used

$$\nu^{M^\star}(\omega, \mathrm{d}t, \mathrm{d}x) = N_t(\omega, \mathrm{d}x)\, \mathrm{d}C_t, \quad \text{where } \ C_t = \sum_{i=1}^{d} \lceil M^{\star^i} \rceil_t$$

and $N_i(\omega, A) = \mathbf{P}(V_i \in A | \mathcal{F}_{i-1})$. More detailed information on this can be found, for example, in [12], [37] or [38]. To finally ensure that expression (54) converges to zero in probability, we observe

$$\begin{aligned}
&\left(\mathbf{E} \frac{\sum_{i=1}^n i^{2a\lambda_{\min}+2\gamma-2} \mathbf{E}\left(\|V_i\|^2 | \mathcal{F}_{i-1}\right)}{n^{2a\lambda_{\min}+2\gamma-1}}\right)^{1+\delta/2} \\
&\le \mathbf{E} \frac{\sum_{i=1}^n i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)} \mathbf{E}\left(\|V_i\|^{2+\delta} | \mathcal{F}_{i-1}\right)}{n^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} = \frac{\sum_{i=1}^n i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)} \mathbf{E}\|V_i\|^{2+\delta}}{n^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} \\
&\le \left(\sup_j \mathbf{E}\|V_j\|^{2+\delta}\right) \frac{\sum_{i=1}^n i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)}}{n^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} \le \mathcal{C} \frac{\sum_{i=1}^n i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)}}{n^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} \overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

We obtain the last convergence claim by Kronecker's lemma

$$\sum_{i=1}^{\infty} \frac{i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)}}{i^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} = \sum_{i=1}^{\infty} i^{-(1+\delta/2)} < \infty \implies \lim_{n\to\infty} \frac{\sum_{i=1}^{n} i^{(2a\lambda_{\min}+2\gamma-2)(1+\delta/2)}}{n^{(2a\lambda_{\min}+2\gamma-1)(1+\delta/2)}} = 0.$$

□

4 Asymptotic Properties of the Averaged Process Using Slowly Decaying Weights

Below we deal with the averaged process

$$\bar{Z}_{\delta,t} := \frac{1+\delta}{(1+R_t)^{1+\delta}} \int_0^t (1+R_{s-})^\delta Z_s \,\mathrm{d}R_s. \tag{55}$$

Here we presuppose that the process R_t is deterministic, $a_s := \frac{a}{(1+R_{s-})^\alpha}$, and $c_s := \frac{c}{(1+R_{s-})^\gamma}$ with $0 < \gamma, \alpha < 1$. We study the asymptotic behavior of $\bar{Z}_{\delta,t}$ in detail to compare it with the behavior of Z_t. It should be noted that the integral which appears in (55) is a Riemann-Stieltjes integral and not a stochastic integral in its narrow sense. Therefore we can use Z_s instead of Z_{s-} at this point. Let us consider the discrete-time version of $\bar{Z}_{\delta,t}$ to get a better understanding. If we use $R_t := \lfloor t \rfloor$, the process results in

$$\begin{aligned}\bar{Z}_{\delta,t} &= \frac{1+\delta}{(1+R_t)^{1+\delta}} \int_0^t (1+R_{s-})^\delta Z_s \,\mathrm{d}R_s = \frac{1+\delta}{(1+\lfloor t \rfloor)^{1+\delta}} \sum_{i=0}^{\lfloor t \rfloor} i^\delta Z_i \Delta R_i \\ &= \frac{1+\delta}{(1+n)^{1+\delta}} \sum_{i=1}^{n} i^\delta Z_i \quad \text{for } t = n \in \mathbb{N}\end{aligned}$$

allowing a comparison with the averaged process

$$\tilde{X}_{\delta,n} := \frac{1+\delta}{n^{1+\delta}} \sum_{i=1}^{n} i^\delta X_i$$

which was considered in [3] and [4]. Thus it is clear that we can interpret the averaged process as a weighted mean.

4.1 Consistency

We want to investigate which one of the two processes, $\bar{Z}_{\delta,t}$ and Z_t, has preferable asymptotic properties. We come to the first crucial question: Is the averaged process consistent? It turns out that consistency of the averaged process follows directly from the consistency of the process Z_t.

Theorem 4.1. *Let the process Z_t be the unique strong solution of the stochastic integral equation* (5) *and let $\bar{Z}_{\delta,t}$ be the corresponding averaged process* (55). *If Z_t is consistent, that is, $Z_t \to x^\star$ almost surely, then the corresponding averaged process $\bar{Z}_{\delta,t}$ is also consistent for $\delta > -1$, that is*

$$\bar{Z}_{\delta,t} \overset{t\to\infty}{\longrightarrow} x^\star \quad a.s..$$

In the theorem above we assure the consistency of the process Z_t. Requirements that assure consistency of the process are given in Theorem 2.1.

Figure 5 compares paths of the processes Z_t of (5) and $\bar{Z}_{\delta,t}$ of (55) in the Itô setting. The path of the process $\bar{Z}_{\delta,t}$ is based on the presented path of the process Z_t. We choose $f(x) = 0.5x^T x$, $x \in \mathbb{R}^2$, $Z_0 = (5,6)^T$, $a = c = 1$, $\alpha = 0.75$, $\gamma = 0.25$, $\delta = 1$, $\sigma_s(\cdot) = I$ and use the Milstein scheme (see, e.g., [14]) to simulate a path of the process Z_t in the special situation of (29). Furthermore Figure 6 illustrates how the processes $f(Z_t)$ and $f(\bar{Z}_{\delta,t})$ behave on the surface of the function f by showing a path of each process. Here we take the same choices and especially the same paths of Z_t and $\bar{Z}_{\delta,t}$ as in Figure 5.

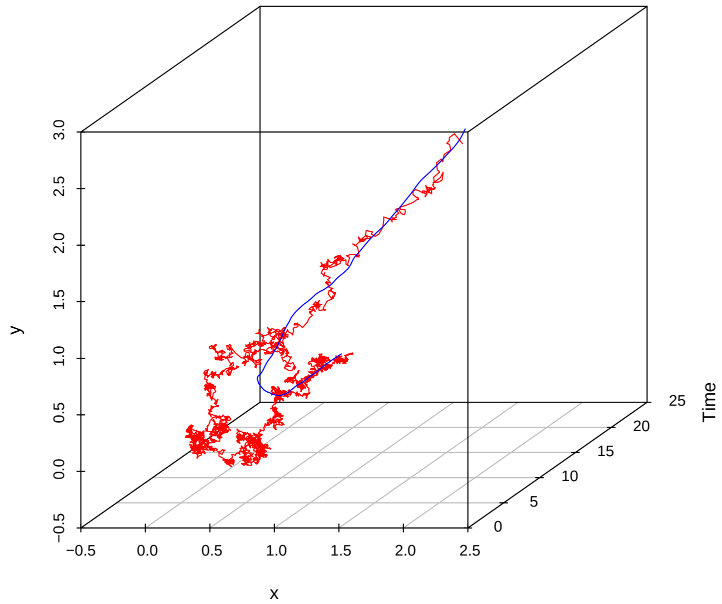

Figure 5: Simulation of a path of the process Z_t from (29) (Itô type) in red. The corresponding averaged process $\bar{Z}_{\delta,t}$ from (55) of the same path is overlaid in blue.

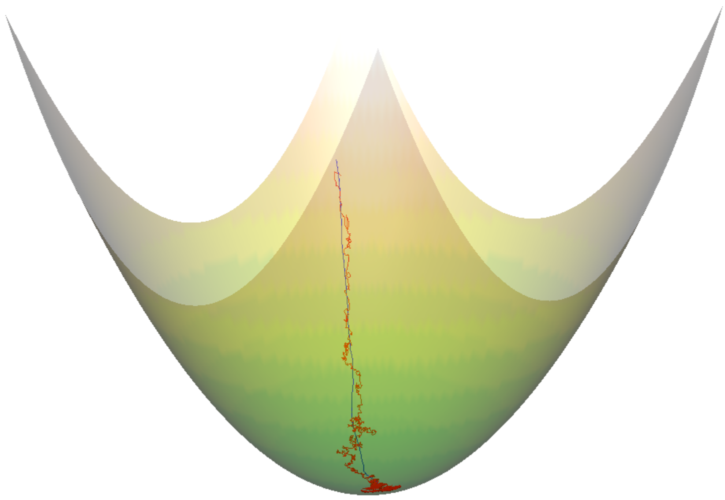

Figure 6: Simulation of a path of the process $f(Z_t)$ in red and of the process $f(\bar{Z}_{1,t})$ in blue. The paths of both processes are based on the simulated paths of Figure 5.

Proof of Theorem 4.1

We rewrite the process $\bar{Z}_{\delta,t}$ slightly and use the Toeplitz lemma to prove the theorem. First, we rewrite the denominator of $\bar{Z}_{\delta,t}$ a little by using Itô's formula and a Taylor expansion

$$\begin{aligned}
\frac{(1+R_t)^{1+\delta}-1}{1+\delta} &= \int_0^t (1+R_{s-})^\delta \,\mathrm{d}R_s + \delta \sum_{0<s\le t} (1+R_{s-}+\xi_s\Delta R_s)^{\delta-1}(\Delta R_s)^2 \\
&= \int_0^t (1+R_{s-})^\delta \,\mathrm{d}R_s + \delta \sum_{0<s\le t} (1+R_{s-})^\delta \left(1+\frac{\xi_s\Delta R_s}{1+R_{s-}}\right)^{\delta-1} \frac{(\Delta R_s)^2}{1+R_{s-}} \\
&= \int_0^t (1+R_{s-})^\delta \,\mathrm{d}R_s + \sum_{0<s\le t} (1+R_{s-})^\delta \left(1+\mathrm{o_b}(1)\right)^{\delta-1} \mathrm{o_b}(1)\Delta R_s \\
&= \int_0^t (1+R_{s-})^\delta \,\mathrm{d}R_s + \sum_{0<s\le t} (1+R_{s-})^\delta \mathrm{o_b}(1)\Delta R_s.
\end{aligned}$$

All the above processes are deterministic and we are only interested in the asymptotic behavior. Hence we have

$$\begin{aligned}
\lim_{t\to\infty} \bar{Z}_{\delta,t} &= (1+\delta)\lim_{t\to\infty} \frac{\int_0^t (1+R_{s-})^\delta Z_s \,\mathrm{d}R_s}{(1+R_t)^{1+\delta}} = (1+\delta)\lim_{t\to\infty} \frac{\int_0^t (1+R_{s-})^\delta Z_s \,\mathrm{d}R_s}{(1+\delta)\int_0^t (1+R_{s-})^\delta \,\mathrm{d}R_s} \\
&= \lim_{t\to\infty} \frac{(1+R_{t-})^\delta Z_t}{(1+R_{t-})^\delta} = \lim_{t\to\infty} Z_t = x^\star \text{ a.s.}
\end{aligned}$$

because $(1+R_t)^{1+\delta} \uparrow \infty$ holds for all $\delta > -1$. This proves consistency of the averaged process $\bar{Z}_{\delta,t}$. □

4.2 Asymptotic Normality

In the previous section we introduced the averaged process $\bar{Z}_{\delta,t}$ which was built with slowly decaying weights $a_t := \frac{a}{(1+R_{s-})^\alpha}$ and $c_t := \frac{c}{(1+R_{s-})^\gamma}$ with $0 < \gamma, \alpha < 1$. Then we proved consistency of the process $\bar{Z}_{\delta,t}$. Hence there are two consistent processes available. For the standard process we achieved, after careful consideration, asymptotic normality. To compare the two processes we need a closer inspection of the asymptotic behavior of the process $\bar{Z}_{\delta,t}$. Theorem 4.2 states a result for asymptotic normality of the process $\bar{Z}_{\delta,t}$. For both processes the limit distribution is normal, but with different means and variances.

Theorem 4.2. *Let f be three times differentiable at $x^\star$, Hessian continuous around $x^\star$, $\gamma = \frac{1}{6}$, and the assumptions of Theorem 3.1 with $\alpha \in (\frac{5}{6}, 1)$ be valid. Furthermore, assume $\delta > -\frac{2}{3}$, $\sum_{0<s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} < \infty$ for all $i \in \{1, \dots, d\}$,*

$$\lim_{\substack{s\to\infty \\ x\to x^\star}} h_s(x, x^\star) = \lim_{\substack{s\to\infty \\ x\to x^\star}} h_s(x) = h(x^\star) \quad \textit{where} \quad h_s(x,y) := \frac{\mathrm{d}\lceil \int_0^\cdot M(\mathrm{d}t, x), \int_0^\cdot M(\mathrm{d}t, y) \rceil_s}{\mathrm{d}R_s}$$

and, for all $\epsilon \in (0, 1]$,

$$\frac{\int_0^t \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{G^\epsilon_{s,t}} x^T x \; \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x)}{(1+R_t)^{1+2\gamma+2\delta}} \xrightarrow{\mathbf{P}} 0 \quad (t \to \infty),$$

where $M_t^\star := \int_0^t M(\mathrm{d}s, x^\star)$, $G^\epsilon_{s,t} := \left\{ x \in \mathbb{R}^d \;\middle|\; \|x\|^2 > \epsilon \frac{(1+R_t)^{\frac{2}{3}+\delta}}{(1+R_s)^{\delta+\alpha}(1+R_{s-})^{\gamma-\alpha}} \right\}$, and $\nu^{M^\star}$ is the compensator of the jump-measure $\mu^{M^\star}$ of the local martingale $M_t^\star$.

We get

$$(1+R_t)^{\frac{1}{3}} (\bar{Z}_{\delta,t} - x^\star) \xrightarrow{\mathcal{D}} N(\tilde{\mu}, \tilde{\Sigma}),$$

where

$$\tilde{\mu} := -\frac{c^2(1+\delta)}{4+6\delta} H_{x^\star}^{-1} \left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star) \right)_{i\in\{1,\dots,d\}} \quad \textit{and} \quad \tilde{\Sigma} := \frac{3(1+\delta)^2}{8c^2(2+3\delta)} H_{x^\star}^{-1} h(x^\star) H_{x^\star}^{-1}.$$

Remark 4.1. In Theorem 3.2(b) a bias occurred, if f is three times differentiable at $x^\star$. Considering the theorem above and in particular $\tilde{\mu}$, we find that here a bias appears, too. This is a well-known phenomenon that is due to the "basic" approximation with finite differences. Using modified differences, the bias can be eliminated. The articles [3], [4], and [9] are recommended to interested readers. Among other things, they illustrate why the bias does not occur in the case f is two times differentiable (compare Theorem 3.2(a)).

Remark 4.2. The restriction $\alpha \in (\frac{5}{6}, 1)$ in Theorem 4.2 is stronger than the restriction $\alpha \in (\frac{2}{3}, 1)$ which is generally used in the discrete-time setting. But if we consider the averaged Robbins-Monro process, the stronger condition seems natural in the continuous-time setting. There, in the continuous-time setting the restriction $\alpha \in (\frac{5}{6}, 1)$ is also needed as opposed to the discrete-time setting.

Remark 4.3. If we consider the assumptions of the Theorem 4.2, we find that the condition on the eigenvalue $\lambda_{\min} > \frac{1-2\gamma}{2a}$ is not needed. This is one of the most useful benefits of the averaged process, because considering the standard process and in particular Theorem 3.2, we recognize the following dilemma. A minimization of the variance concerning the parameter a, which appears in the limit distribution, is desirable. But to keep the condition $\lambda_{\min} > \frac{1-2\gamma}{2a}$, which guarantees stability, and simultaneously to minimize the variance concerning the parameter a is not possible without knowledge of $\lambda_{\min}$, which is generally unknown.

Remark 4.4. The condition $\lambda_{\min} > \frac{1-2\gamma}{2a}$ can be omitted since we consider a_t with $\alpha < 1$. We have already seen a similar conclusion in Theorem 3.1.

Remark 4.5. We point out that the averaged process does not perform well until the standard process is in a small neighborhood of $x^\star$. Before we reach such a region we should prefer the standard process. On the other hand, the averaged process mostly performs well if the standard process oscillates around $x^\star$, otherwise we should prefer the standard process as well.

Proof of Theorem 4.2

Without loss of generality we assume $x^\star = 0$. Frequently we write γ instead of $\frac{1}{6}$ to get a tight notation. We consider the averaged process $\bar{Z}_{\delta,t}$ using the explicit representation of Z_t, that is with (40)

$$\begin{aligned}
\bar{Z}_{\delta,t} &= \frac{1+\delta}{(1+R_t)^{1+\delta}} \int_0^t Z_s (1+R_{s-})^\delta \, \mathrm{d}R_s \\
&= \frac{1+\delta}{(1+R_t)^{1+\delta}} \int_0^t (1+R_{s-})^\delta \Big(T\phi_s \Big(T^T Z_0 \\
&\quad - \frac{a}{2c} \int_0^s (1+R_{r-})^{\gamma-\alpha} \phi_r^{-1} T^T M(\mathrm{d}r, 0) - \int_0^s \phi_r^{-1} T^T \, \mathrm{d}\tilde{R}_r \Big)\Big) \, \mathrm{d}R_s \\
&= \frac{1+\delta}{(1+R_t)^{1+\delta}} \Big(\int_0^t (1+R_{s-})^\delta T\phi_s T^T Z_0 \, \mathrm{d}R_s - \int_0^t (1+R_{s-})^\delta T\phi_s \int_0^s \phi_r^{-1} T^T \, \mathrm{d}\tilde{R}_r \, \mathrm{d}R_s \\
&\quad - \int_0^t (1+R_{s-})^\delta T\phi_s \frac{a}{2c} \int_0^s (1+R_{r-})^{\gamma-\alpha} \phi_r^{-1} T^T M(\mathrm{d}r, 0) \, \mathrm{d}R_s \Big) \\
&= A_t - B_t - C_t.
\end{aligned}$$

Now we show

$$\textbf{(I)} \qquad (1+R_t)^{\frac{1}{3}}\, A_t \xrightarrow{\mathbf{P}} 0$$

$$\textbf{(II)} \qquad (1+R_t)^{\frac{1}{3}}\, B_t \xrightarrow{\mathbf{P}} \frac{c^2(1+\delta)}{4+6\delta} H_0^{-1} \left(\frac{\partial^3 f}{(\partial x_i)^3}(0) \right)_{i\in\{1,\dots,d\}}$$

$$\textbf{(III)} \qquad (1+R_t)^{\frac{1}{3}}\, C_t \xrightarrow{\mathcal{D}} N(0,\tilde{\Sigma}) \quad \text{where } \tilde{\Sigma} := \frac{3(1+\delta)^2}{8c^2(2+3\delta)}\, H_0^{-1} h(0) H_0^{-1}.$$

Thus we achieve the desired result using Slutsky's Theorem.

But before we prove (I), (II), and (III), we pursue the same strategy as in the proof of the asymptotic normality of the standard process. Because the process ϕ_t appears above, we consider it in more detail as in the proof of Theorem 3.2. In the proof of Theorem 3.2 we studied the case $\alpha = 1$, now we consider the case $\alpha \in (\frac{5}{6}, 1)$. We use (45),

$$\begin{aligned} &- a\lambda_i \int_0^t (1+R_{s-})^{-\alpha}\,\mathrm{d}R_s \\ &\qquad = -\frac{a\lambda_i}{1-\alpha}\left((1+R_t)^{1-\alpha} - 1 - \sum_{s\le t}\Big(\Delta(1+R_s)^{1-\alpha} - (1-\alpha)(1+R_{s-})^{-\alpha}\Delta R_s \Big) \right) \end{aligned}$$

and with the formula $\mathcal{E}_t(X) = e^{X_t - X_0 - 0.5[X,X]_t^c} \prod_{s\le t}(1+\Delta X_s)e^{-\Delta X_s}$ we obtain

$$\phi_t^{ii} = \exp\Big(-\frac{a\lambda_i}{1-\alpha}(1+R_t)^{1-\alpha} \Big)\, \Pi_{0,t}^{a\lambda_i}$$

where

$$\Pi_{0,t}^{a\lambda_i} := e^{\frac{a\lambda_i}{1-\alpha}} \prod_{s\le t}\left(\exp\Big(\frac{a\lambda_i}{1-\alpha}\Delta(1+R_s)^{1-\alpha} \Big) \Big(1 - \frac{a\lambda_i \Delta R_s}{(1+R_{s-})^{\alpha}} + \mathbb{1}_{\{\frac{a\lambda_i\Delta R_s}{(1+R_{s-})^{\alpha}}=1\}} \Big) \right).$$

Now we show that the process $\Pi_{0,t}^{a\lambda_i}$ converges. In a first step we conclude

$$\int_0^\infty a_s\,\mathrm{d}R_s = \infty \quad \wedge \quad \int_0^\infty a_s^2 \Delta R_s\,\mathrm{d}R_s^d < \infty \qquad \Longrightarrow \qquad a_s\Delta R_s = \frac{a\Delta R_s}{(1+R_s)^{\alpha}} \to 0.$$

In a second step we use a Taylor expansion and see, for sufficiently large s, that the product

$$\begin{aligned} &\ln \prod_{s\le t}\left(\exp\Big(\frac{a\lambda_i}{1-\alpha}\Delta(1+R_t)^{1-\alpha} \Big) \Big(1 - \frac{a\lambda_i\Delta R_t}{(1+R_{t-})^{\alpha}} \Big) \right) \\ &\qquad = \sum_{s\le t}\left(\frac{a\lambda_i}{1-\alpha}\Delta(1+R_t)^{1-\alpha} + \ln\Big(1 - \frac{a\lambda_i\Delta R_t}{(1+R_{t-})^{\alpha}} \Big) \right) \end{aligned}$$

$$\leq \sum_{s\leq t}\left(\rho_t - \frac{a^2\lambda_i^2}{2}\right)\left(\frac{\Delta R_t}{(1+R_t)^\alpha}\right)^2 \quad \text{with } \rho_t \to 0$$

converges. By combining both steps, it is easy to transfer the arguments from the proof of convergence of $\Pi_{0,t}^{a\lambda_i}$ (in the case $\alpha = 1$) in Theorem 3.2. In the following, we use

$$\phi_t^{ii} = \exp\Big(-\frac{a\lambda_i}{1-\alpha}(1+R_t)^{1-\alpha}\Big)\,\Pi_{0,t}^{a\lambda_i} \qquad \text{where} \quad \Pi_{0,t}^{a\lambda_i} \to \Pi_{0,\infty}^{a\lambda_i} < \infty.$$

Part (I): To prove $(1+R_t)^{\frac{1}{3}}\,A_t \xrightarrow{\mathbf{P}} 0$, it suffices to show for all $i \in \{1,\ldots,d\}$

$$(1+\delta)(1+R_t)^{-\frac{2}{3}-\delta}\int_0^t (1+R_{s-})^\delta \phi_s^{ii}\,\mathrm{d}R_s \longrightarrow 0.$$

All the above processes are purely deterministic, although we use a generalization of Kronecker's lemma ([19], p. 141, Lemma 3), because we have already used this several times. Thus it suffices to show

$$\left[(1+R_t)^{\frac{2}{3}+\delta} \to \infty\right] \cap \left[\int_0^\infty \frac{(1+R_{s-})^\delta \phi_s^{jj}}{(1+R_s)^{\frac{2}{3}+\delta}}\,\mathrm{d}R_s < \infty\right] = \Omega.$$

Using the assumption $\delta > -\frac{2}{3}$, the first set equals Ω. The same is true for the second set, because

$$\int_0^\infty \frac{(1+R_{s-})^\delta \phi_s^{jj}}{(1+R_s)^{\frac{2}{3}+\delta}}\,\mathrm{d}R_s \leq \int_0^\infty \frac{\phi_s^{jj}}{(1+R_s)^{\frac{2}{3}}}\,\mathrm{d}R_s < \infty.$$

This completes verification of (I).

Part (II): Here the process $\tilde{R}_t$ of Lemma 3.3 appears. Therefore we have

$$\tilde{R}_t = a\int_0^t V(Z_{s-},c_s)(1+R_{s-})^{-\alpha}\,\mathrm{d}R_s + K_t^1 + K_t^2$$

where

$$K_t^1 := \frac{a}{2c}\int_0^t (1+R_{s-})^{\gamma-\alpha}\left(M(\mathrm{d}s,Z_{s-}) - M(\mathrm{d}s,0)\right)$$

$$K_t^2 := \left\{\sum_{j,k} T_{ij}T_{jk}^T \sum_{s\leq t} Z_{s-}^k\,\mathbb{1}_{\{a\lambda_j \Delta R_s = (1+R_{s-})^\alpha\}}\right\}_{i\in\{1,\ldots,d\}}.$$

As $V(\cdot,\cdot)$ appears, we use representation (49)

$$V(Z_{s-},c_s) = \frac{1}{6}A_3 c_s^2 + \mathcal{O}(\|Z_{s-}\|^2) + \mathrm{o}(c_s^2)$$

and obtain

$$a(1+R_{s-})^{-\alpha}V(Z_{s-},c_s) = (1+R_{s-})^{-\alpha}\left(\frac{a}{6}A_3c_s^2 + \mathcal{O}(\|Z_{s-}\|^2) + \mathrm{o}(c_s^2)\right).$$

Hence, it suffices to verify the following five convergences to prove (II):

$$(i) \qquad -\sum_{j,k} T_{ij}T_{jk}^T A_3^k \frac{a(1+\delta)}{6}\frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}c_r^2 \,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \to \tilde{\mu}$$

$$\text{where } \tilde{\mu} := -\frac{c^2(1+\delta)}{4+6\delta}H_0^{-1}\left(\frac{\partial^3 f}{(\partial x_i)^3}(0)\right)_{i\in\{1,\dots,d\}}$$

$$(ii) \qquad \frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}\mathcal{O}(\|Z_{r-}\|^2)\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \xrightarrow{\mathbf{P}} 0$$

$$(iii) \qquad \frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}\mathrm{o}(c_r^2)\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \to 0$$

$$(iv) \qquad \frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}}(1+R_{r-})^{\gamma-\alpha}\left(M_k(\mathrm{d}r,Z_{r-}) - M_k(\mathrm{d}r,0)\right)\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \xrightarrow{\mathbf{P}} 0$$

$$(v) \qquad \frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \sum_{r\le s} \phi_r^{jj^{-1}} Z_{r-}^k \mathbb{1}_{\{a\lambda_j\Delta R_r = (1+R_{r-})^\alpha\}}\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \xrightarrow{\mathbf{P}} 0.$$

Verification of (i): It suffices to show

$$\frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha-2\gamma}\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \longrightarrow \frac{1}{a\lambda_j(\frac{2}{3}+\delta)}.$$

Let us consider the process $(1+R_t)^\delta\phi_t^{jj^{-1}}$ for large t. We use integration by parts and obtain two processes:

$$(1+R_t)^\delta\phi_t^{jj^{-1}} = 1 + \int_0^t \phi_s^{jj^{-1}}\,\mathrm{d}(1+R_s)^\delta + \int_0^t (1+R_{s-})^\delta\,\mathrm{d}\phi_s^{jj^{-1}}.$$

The first process satisfies

$$\begin{aligned}\int_0^t \phi_s^{jj^{-1}}\,\mathrm{d}(1+R_s)^\delta &= \mathcal{C}\int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{\delta-1}\,\mathrm{d}R_s + \int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{\delta-1}\mathrm{o_b}(1)\,\mathrm{d}R_s^d\\ &\le \mathcal{C}\int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{\delta-1}\,\mathrm{d}R_s\end{aligned}$$

and the second process satisfies

$$\int_0^t (1+R_{s-})^\delta\,\mathrm{d}\phi_s^{jj^{-1}} = \mathcal{C}\int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{\delta-\alpha}\,\mathrm{d}R_s.$$

It is obvious that the first process is of a lower order. Since we are only interested in the asymptotic behavior, we will neglect processes of lower order in the following. We use the Toeplitz lemma and get

$$\begin{aligned}
&\Big(\frac{2}{3}+\delta\Big)\lim_{t\to\infty}\frac{\int_0^t(1+R_{s-})^{\delta}\phi_s^{jj}\int_0^s\phi_s^{jj^{-1}}(1+R_{r-})^{-\alpha-2\gamma}\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}}\\
&\quad=\Big(\frac{2}{3}+\delta\Big)\lim_{t\to\infty}\frac{\int_0^t(1+R_{s-})^{\delta}\phi_s^{jj}\int_0^s\phi_s^{jj^{-1}}(1+R_{r-})^{-\alpha-\frac{1}{3}}\,\mathrm{d}R_r\,\mathrm{d}R_s}{(\frac{2}{3}+\delta)\int_0^t(1+R_{s-})^{\delta-\frac{1}{3}}\,\mathrm{d}R_s}\\
&\quad=\lim_{t\to\infty}\frac{(1+R_{t-})^{\delta}\phi_t^{jj}\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha-\frac{1}{3}}\,\mathrm{d}R_s}{(1+R_{t-})^{\delta-\frac{1}{3}}}\\
&\quad=\lim_{t\to\infty}\frac{\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha-\frac{1}{3}}\,\mathrm{d}R_s}{\phi_t^{jj^{-1}}(1+R_{t-})^{-\frac{1}{3}}}=\lim_{t\to\infty}\frac{\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha-\frac{1}{3}}\,\mathrm{d}R_s}{\int_0^t(1+R_{s-})^{-\frac{1}{3}}\,\mathrm{d}\phi_s^{jj^{-1}}}\\
&\quad=\lim_{t\to\infty}\frac{\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha-\frac{1}{3}}\,\mathrm{d}R_s}{\int_0^t(1+R_{s-})^{-\frac{1}{3}-\alpha}\overline{a\lambda_j}^{s}\phi_s^{jj^{-1}}\,\mathrm{d}R_s}=\lim_{t\to\infty}\frac{\phi_t^{jj^{-1}}(1+R_{t-})^{-\alpha-\frac{1}{3}}}{a\lambda_j(1+R_{t-})^{-\frac{1}{3}-\alpha}\phi_t^{jj^{-1}}}\\
&\quad=\frac{1}{a\lambda_j}.
\end{aligned}$$

Verification of (ii): We consider L^1-convergence as we are interested in the convergence in probability and neglect processes of lower order. We use Fubini's theorem and Theorem 3.1. It remains to prove

$$\frac{\int_0^t(1+R_{s-})^{\delta}\phi_s^{jj}\int_0^s\phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}\mathbf{E}\|Y_{r-}\|^2\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}}\to 0$$

where Y_t follows from Theorem 3.1. In a first step we assume

$$\frac{\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s}{\phi_t^{jj^{-1}}(1+R_{t-})^{-\frac{1}{3}}}\longrightarrow 0 \tag{56}$$

and with the Toeplitz lemma we obtain

$$\begin{aligned}
&\lim_t\frac{\int_0^t(1+R_{s-})^{\delta}\phi_s^{jj}\int_0^s\phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}\mathbf{E}\|Y_{r-}\|^2\,\mathrm{d}R_r\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}}\\
&\quad=\lim_t\frac{\int_0^t(1+R_{s-})^{\delta}\phi_s^{jj}\int_0^s\phi_r^{jj^{-1}}(1+R_{r-})^{-\alpha}\mathbf{E}\|Y_{r-}\|^2\,\mathrm{d}R_r\,\mathrm{d}R_s}{\int_0^t(1+R_{s-})^{\delta-\frac{1}{3}}\,\mathrm{d}R_s}\\
&\quad=\lim_t\frac{\int_0^t\phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s}{\phi_t^{jj^{-1}}(1+R_{t-})^{-\frac{1}{3}}}\\
&\quad=0.
\end{aligned}$$

Finally we need to prove (56). From Theorem 3.1 with $\gamma = \frac{1}{6}$ and $\alpha \in (\frac{5}{6}, 1)$ the L^2-convergence rate of the process of Y_t is given by

$$\mathbf{E}\|Y_t\|^2 \le \mathcal{C}(1+R_t)^{\frac{4}{3}-2\alpha}$$

and (56) follows from

$$\begin{aligned}\lim_t \frac{\int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{-\alpha}\mathbf{E}\|Y_{s-}\|^2\,\mathrm{d}R_s}{\phi_t^{jj^{-1}}(1+R_{t-})^{-\frac{1}{3}}} &\le \mathcal{C}\lim_t \frac{\int_0^t \phi_s^{jj^{-1}}(1+R_{s-})^{\frac{4}{3}-3\alpha}\,\mathrm{d}R_s}{\phi_t^{jj^{-1}}(1+R_{t-})^{-\frac{1}{3}}} \\ &= \mathcal{C}\lim_t \frac{\int_0^t (1+R_{s-})^{\frac{4}{3}-2\alpha}\,\mathrm{d}\phi_s^{jj^{-1}}}{\int_0^t (1+R_{s-})^{-\frac{1}{3}}\,\mathrm{d}\phi_s^{jj^{-1}}} = \mathcal{C}\lim_t (1+R_{s-})^{\frac{5}{3}-2\alpha} = 0 \quad \alpha \in \Big(\frac{5}{6}, 1\Big).\end{aligned}$$

Verification of (iii): Since $A_3 c_s^2$ appears with the convergence to $\tilde{\mu}$ in (i) and $\mathrm{o}(c_s^2)$ with the convergence to 0 in (iii), (iii) follows from (i).

Verification of (iv): Follows immediately from the proof of part (III), the assumption

$$\lim_{\substack{s\to\infty \\ x\to x^\star}} h_s(x, x^\star) = \lim_{\substack{s\to\infty \\ x\to x^\star}} h_s(x) = h(x^\star),$$

and the fact that $(M(\mathrm{d}r, Z_{r-}) - M(\mathrm{d}r, 0))$ appears as integrator instead of $M(\mathrm{d}r, 0)$.

Verification of (v): We will show

$$\frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \sum_{r\le s} \phi_r^{jj^{-1}} Z_{r-}^k \mathbb{1}_{\{a\lambda_j \Delta R_r = (1+R_{r-})^\alpha\}}\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \xrightarrow{\mathbf{P}} 0.$$

Using

$$\sum_{0\le s} \mathbb{1}_{\{a\lambda_j \Delta R_s = (1+R_{s-})^\alpha\}} < \infty \quad \Longrightarrow \quad \sum_{0\le r} \phi_r^{jj^{-1}} Z_{r-}^k \mathbb{1}_{\{a\lambda_j \Delta R_r = (1+R_{r-})^\alpha\}} \le \mathcal{C}_\omega < \infty,$$

we easily get

$$\frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj} \sum_{r\le s} \phi_r^{jj^{-1}} Z_{r-}^k \mathbb{1}_{\{a\lambda_j \Delta R_r = (1+R_{r-})^\alpha\}}\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}} \le \mathcal{C}_\omega \underbrace{\frac{\int_0^t (1+R_{s-})^\delta \phi_s^{jj}\,\mathrm{d}R_s}{(1+R_t)^{\frac{2}{3}+\delta}}}_{\to 0}$$

by $\phi_s^{jj} = \exp\left(-\frac{a\lambda_i}{1-\alpha}(1+R_s)^{1-\alpha}\right)\Pi_{0,s}^{a\lambda_i} \to 0$.

Part (III): Our next goal is the proof of part (III), that is, verification of $(1+R_t)^{\frac{1}{3}} C_t \xrightarrow{\mathcal{D}} N(0, \tilde{\Sigma})$ and, more precisely,

$$\left\{\frac{a(1+\delta)}{2c} \sum_{j,k=1}^{d} T_{ij} T_{jk}^T (1+R_t)^{-\frac{2}{3}-\delta} X_t^k \right\}_{i\in\{1,\ldots,d\}} \xrightarrow{\mathcal{D}} N(0, \tilde{\Sigma})$$

where

$$X_t^k := \int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s (1+R_{r-})^{\gamma-\alpha} \phi_r^{jj^{-1}} \, M_k(\mathrm{d}r,0)\,\mathrm{d}R_s.$$

We investigate the process X_t^k using $L_t := \int_0^t (1+R_{r-})^{\gamma-\alpha}\phi_r^{jj^{-1}}\, M_k(\mathrm{d}r,0)$ and $N_t := \int_0^t (1+R_{s-})^\delta \phi_s^{jj}\,\mathrm{d}R_s$. Then we have

$$\begin{aligned}
X_t^k &= \int_0^t (1+R_{s-})^\delta \phi_s^{jj} \int_0^s (1+R_{r-})^{\gamma-\alpha}\phi_r^{jj^{-1}}\, M_k(\mathrm{d}r,0)\,\mathrm{d}R_s = \int_0^t L_s\,\mathrm{d}N_s \\
&= \int_0^t L_{s-}\,\mathrm{d}N_s + \int_0^t \Delta L_s\,\mathrm{d}N_s = L_t N_t - \int_0^t N_{s-}\,\mathrm{d}L_s \\
&= N_t \int_0^t \mathrm{d}L_s - \int_0^t N_{s-}\,\mathrm{d}L_s = \int_0^t N_t\,\mathrm{d}L_s - \int_0^t N_{s-}\,\mathrm{d}L_s \\
&= \int_0^t (N_t - N_{s-})\,\mathrm{d}L_s.
\end{aligned}$$

To use the representation above effectively, we show that $\int_0^t \phi_{s-}^{jj}(1+R_{s-})^\delta\,\mathrm{d}R_s$ can be used instead of N_t:

$$\begin{aligned}
N_t &= \int_0^t (1+R_{s-})^\delta \phi_s^{jj}\,\mathrm{d}R_s = \int_0^t (1+R_{s-})^\delta (\phi_{s-}^{jj} + \Delta\phi_s^{jj})\,\mathrm{d}R_s \\
&= \int_0^t (1+R_{s-})^\delta \Big(\phi_{s-}^{jj} - \Delta \int_0^s \overline{a\lambda_j}^s \phi_{s-}^{jj}(1+R_{s-})^{-\alpha}\,\mathrm{d}R_s \Big)\,\mathrm{d}R_s \\
&= \int_0^t \phi_{s-}^{jj} \Big(1 - \frac{\overline{a\lambda_j}^s \Delta R_s}{(1+R_{s-})^\alpha} \Big)(1+R_{s-})^\delta\,\mathrm{d}R_s \\
&= \int_0^t \phi_{s-}^{jj}(1+R_{s-})^\delta\,\mathrm{d}R_s + \int_0^t \phi_{s-}^{jj}\mathrm{o_b}(1)(1+R_{s-})^\delta\,\mathrm{d}R_s.
\end{aligned}$$

We use the crucial trick

$$\begin{aligned}
N_t - N_{s-} &= \int_s^t \phi_{r-}^{jj}(1+R_{r-})^\delta\,\mathrm{d}R_r \overset{(\star)}{=} -\frac{1}{a\lambda_j}\int_s^t (1+R_{r-})^{\delta+\alpha}\,\mathrm{d}\phi_r^{jj} \\
&= -\frac{1}{a\lambda_j}\left((1+R_t)^{\delta+\alpha}\phi_t^{jj} - (1+R_s)^{\delta+\alpha}\phi_s^{jj} - \int_s^t \phi_r^{jj}\,\mathrm{d}(1+R_r)^{\delta+\alpha} \right).
\end{aligned}$$

Here $(\star)$, that is $\overline{a\lambda_j}^s = a\lambda_j \mathbb{1}_{\{a\lambda_j\Delta R_s \neq (1+R_{s-})^\alpha\}} = a\lambda_j$, holds for sufficiently large s, since $\frac{\Delta R_s}{(1+R_{s-})^\alpha} \to 0$. We have

$$\begin{aligned}
\lim_{t\to\infty} \frac{a\lambda_j X_t^k}{(1+R_t)^{\frac{1}{2}+\gamma+\delta}} &= \lim_{t\to\infty} \frac{a\lambda_j \int_0^t (N_t - N_{s-})\,\mathrm{d}L_s}{(1+R_t)^{\frac{1}{2}+\gamma+\delta}} \\
&= \lim_{t\to\infty} M_{t,1}^k - \lim_{t\to\infty} M_{t,2}^k + \lim_{t\to\infty} M_{t,3}^k,
\end{aligned}$$

where

$$M_{t,1}^k := \frac{\int_0^t (1+R_s)^{\delta+\alpha}(1+R_{s-})^{\gamma-\alpha}\, M_k(\mathrm{d}s,0)}{(1+R_t)^{\frac{1}{2}+\gamma+\delta}}$$

$$M_{t,2}^k := \frac{\int_0^t (1+R_{s-})^{\gamma-\alpha} \phi_s^{jj^{-1}} \, M_k(\mathrm{d}s, 0)}{(1+R_t)^{\frac{1}{2}+\gamma-\alpha} \phi_t^{jj^{-1}}}$$
$$M_{t,3}^k := \frac{\int_0^t (1+R_{s-})^{\gamma-\alpha} \phi_s^{jj^{-1}} \int_s^t \phi_r^{jj} \, \mathrm{d}(1+R_r)^{\alpha+\delta} \, M_k(\mathrm{d}s, 0)}{(1+R_t)^{\frac{1}{2}+\gamma+\delta}}.$$

Now we show

$$M_{t,1} \xrightarrow{\mathcal{D}} N\left(0, \frac{h(0)}{1+2\gamma+2\delta}\right) \qquad \text{and} \qquad M_{t,2},\ M_{t,3} \xrightarrow{\mathbf{P}} 0.$$

If this is true, then

$$\begin{aligned} \left\{ \sum_{j,k=1}^{d} \frac{a(1+\delta)}{2ac\lambda_j} T_{ij} T_{jk}^T M_{t,1}^k \right\}_{i \in \{1,\ldots,d\}} &= \frac{(1+\delta)}{2c} \left(T D^{-1} T^T \right) M_{t,1} \\ &= \frac{1+\delta}{2c} H_0^{-1} M_{t,1} \xrightarrow{\mathcal{D}} N(0, \tilde{\Sigma}) \end{aligned}$$

is also true and by Slutsky's theorem it follows

$$(1+R_t)^{\frac{1}{2}-\gamma} C_t \xrightarrow{\mathcal{D}} N(0, \tilde{\Sigma}).$$

To show $M_{t,2} \xrightarrow{\mathbf{P}} 0$ we use

$$\mathbf{E}\|M_{t,2}\|^2 \longrightarrow 0 \quad \Longrightarrow \quad M_{t,2} \xrightarrow{\mathbf{P}} 0.$$

Using the Itô isometry, we find

$$\begin{aligned} \mathbf{E}\left(M_{t,2}^k\right)^2 &= \frac{\mathbf{E}\left(\int_0^t (1+R_{s-})^{\gamma-\alpha} \phi_s^{jj^{-1}} \, M_k(\mathrm{d}s,0)\right)^2}{(1+R_t)^{1+2\gamma-2\alpha} \phi_t^{jj^{-2}}} = \frac{\mathbf{E}\lceil \int_0^t (1+R_{s-})^{\gamma-\alpha} \phi_s^{jj^{-1}} \, M_k(\mathrm{d}s,0) \rceil}{(1+R_t)^{1+2\gamma-2\alpha} \phi_t^{jj^{-2}}} \\ &\leq \frac{\phi_t^{jj^2} \int_0^t (1+R_{s-})^{2\gamma-2\alpha} \phi_s^{jj^{-2}} h_s^{kk}(0) \, \mathrm{d}R_s}{(1+R_t)^{1+2\gamma-2\alpha}} \\ &\leq \mathcal{C}\,(1+R_t)^{-1-2\gamma+2\alpha} \phi_t^{jj^2} \underbrace{\int_0^t (1+R_{s-})^{2\gamma-\alpha} \phi_s^{jj^{-1}} \, \mathrm{d}\phi_s^{jj^{-1}}}_{\leq (1+R_t)^{2\gamma-\alpha} \phi_t^{jj^{-1}} (\phi_t^{jj^{-1}} - \phi_0^{jj^{-1}})} \\ &\leq \mathcal{C}\,(1+R_t)^{\alpha-1} \to 0. \end{aligned}$$

To verify $M_{t,3} \xrightarrow{\mathbf{P}} 0$, we proceed as in the investigation of $M_{t,2}$. We obtain

$$\mathbf{E}\left(M_{t,3}^k\right)^2 \leq \mathcal{C}\, \frac{\int_0^t (1+R_{s-})^{2\gamma-2\alpha} \phi_s^{jj^{-2}} \left(\int_s^t \phi_r^{jj} \, \mathrm{d}(1+R_r)^{\alpha+\delta} \right)^2 \mathrm{d}R_s}{(1+R_t)^{1+2\gamma+2\delta}}.$$

We use ϕ^{jj}_{s-} instead of ϕ^{jj}_{s} in the inner integral as $\phi^{jj}_{s} = \phi^{jj}_{s-}\left(1 + \frac{\overline{a\lambda_j}^{s}\Delta R_s}{(1+R_{s-})^{\alpha}}\right)$ holds. If s and t are sufficiently large, we get

$$\begin{aligned}
&\int_s^t \phi^{jj}_{r-}\,\mathrm{d}(1+R_r)^{\alpha+\delta} \\
&\qquad \leq \int_s^t \phi^{jj}_{r-}(1+R_{r-})^{\alpha+\delta-1}\,\mathrm{d}R_r + \mathcal{C}\sum_{s<r\leq t}\phi^{jj}_{r-}(1+R_{r-}+\xi_r\Delta R_r)^{\alpha+\delta-2}(\Delta R_r)^2 \\
&\qquad \leq \int_s^t \phi^{jj}_{r-}(1+R_{r-})^{\alpha+\delta-1}\,\mathrm{d}R_r + \mathcal{C}\sum_{s<r\leq t}\phi^{jj}_{r-}(1+R_{r-})^{\alpha+\delta-1}\mathrm{o_b}(1)\Delta R_r \\
&\qquad \leq \mathcal{C}\int_s^t \phi^{jj}_{r-}(1+R_{r-})^{\alpha+\delta-1}\,\mathrm{d}R_r = \mathcal{C}\,|\int_s^t \frac{1}{\overline{a\lambda_j}^{r}}(1+R_{r-})^{2\alpha+\delta-1}\,\mathrm{d}\phi^{jj}_{r}| \\
&\qquad \leq \mathcal{C}\,\frac{1}{a\lambda_j}|\int_s^t (1+R_{r-})^{2\alpha+\delta-1}\,\mathrm{d}\phi^{jj}_{r}| \\
&\qquad \overset{(\star)}{\leq} \mathcal{C}\left((1+R_{s-})^{2\alpha+\delta-1} + (1+R_t)^{2\alpha+\delta-1}\right)\phi^{jj}_{s}
\end{aligned}$$

with $(\star)$ because of $2\alpha+\delta-1 \geq 0$. Consequently, we have

$$\mathbf{E}\left(M^k_{t,3}\right)^2 \leq \mathcal{C}\underbrace{\frac{\int_0^t (1+R_{s-})^{2\gamma+2\alpha+2\delta-2}\,\mathrm{d}R_s}{(1+R_t)^{1+2\gamma+2\delta}}}_{\to 0} + \mathcal{C}\underbrace{\frac{\int_0^t (1+R_{s-})^{2\gamma-2\alpha}\,\mathrm{d}R_s}{(1+R_t)^{3-4\alpha+2\gamma}}}_{\to 0}.$$

Convergence results of this type often appeared in Chapter 3. Essentially, here we obtain them by

$$\infty > \int_0^\infty (1+R_{s-})^{2\alpha-3}\,\mathrm{d}R_s = \begin{cases} \int_0^\infty \frac{(1+R_{s-})^{2\gamma+2\alpha+2\delta-2}}{(1+R_{s-})^{1+2\gamma+2\delta}}\,\mathrm{d}R_s \\ \int_0^\infty \frac{(1+R_{s-})^{2\gamma-2\alpha}}{(1+R_{s-})^{3-4\alpha+2\gamma}}\,\mathrm{d}R_s \end{cases}.$$

Now we show

$$M_{t,1} = \frac{\int_0^t (1+R_s)^{\delta+\alpha}(1+R_{s-})^{\gamma-\alpha}\,M(\mathrm{d}s,0)}{(1+R_t)^{\frac{1}{2}+\gamma+\delta}} \xrightarrow{\mathcal{D}} N\left(0, \frac{h(0)}{1+2\gamma+2\delta}\right).$$

We use the same arguments as in the proof of asymptotic normality of the standard process, that is, we use sequences. Let t_n be an increasing, arbitrary but fixed sequence with $t_n \uparrow \infty$. To use a central limit theorem, we consider the sequence M^n_s

$$M^n_s := (1+R_{t_n})^{-\frac{1}{2}-\gamma-\delta}\int_0^{st_n}(1+R_r)^{\delta+\alpha}(1+R_{r-})^{\gamma-\alpha}\,M(\mathrm{d}r,0)$$

which is a sequence of locally square integrable martingales. Thus, we have to show

$$M^n_1 \xrightarrow{\mathcal{D}} M \quad \text{where} \quad M \sim N\left(0, \frac{h(0)}{1+2\gamma+2\delta}\right).$$

For this purpose we use, as in the proof of Theorem 3.2, the vector-valued version of Theorem 4 on page 435 in [19]. To apply this theorem, we have to prove the assumptions

$$(i) \qquad \underset{\delta\in(0,1]}{\forall} \quad x^T x \mathbb{1}_{[\|x\|>\delta]} * \nu_1^{M^n} \xrightarrow{\mathbf{P}} 0,$$

$$(ii) \qquad \lceil M^n \rceil_1 \xrightarrow{\mathbf{P}} \lceil M \rceil_1,$$

of the theorem.

Verification of (i): Here the measure ν^n (that is ν^{M^n}) occurs, which is the compensator of the jump measure μ^n of the process M_s^n. Because the assumptions of Theorem 4.2 are related to the local martingale $M_t^\star := \int_0^t M(\mathrm{d}s, 0)$, which occurs in the stochastic integral equation, we should only use its associated compensator $\nu^{M^\star}$. To obtain this, we consider $x^T x$ and $\mathbb{1}_{[\|x\|>\epsilon]}$, which correspond to the examination of $(\Delta M_s^n)^T (\Delta M_s^n)$ and $\mathbb{1}_{[\|\Delta M_s^n\|>\epsilon]}$. We have

$$\Delta M_s^n = \frac{(1+R_{st_n})^{\delta+\alpha}(1+R_{st_n-})^{\gamma-\alpha}}{(1+R_{t_n})^{\frac{1}{2}+\gamma+\delta}} \, \Delta M_{st_n}^\star$$

and an easy calculation shows

$$(\Delta M_s^n)^T (\Delta M_s^n) = \frac{(1+R_{st_n})^{2\delta+2\alpha}(1+R_{st_n-})^{2\gamma-2\alpha}}{(1+R_{t_n})^{1+2\gamma+2\delta}} \, \Delta M_{st_n}^{\star T} \Delta M_{st_n}^\star.$$

Furthermore, it follows that

$$\begin{aligned}[\|\Delta M_s^n\| > \epsilon] &= \left[\| \frac{(1+R_{st_n})^{\delta+\alpha}(1+R_{st_n-})^{\gamma-\alpha}}{(1+R_{t_n})^{\frac{1}{2}+\gamma+\delta}} \, \Delta M_{st_n}^\star \| > \epsilon \right] \\ &= \left[\|\Delta M_{st_n}^\star\| > \frac{\epsilon (1+R_{t_n})^{\frac{1}{2}+\gamma+\delta}}{(1+R_{st_n})^{\delta+\alpha}(1+R_{st_n-})^{\gamma-\alpha}} \right].\end{aligned}$$

Hence it suffices to use, for all $\epsilon \in (0,1]$,

$$\left\{ x \in \mathbb{R}^d \,\middle|\, \|x\| > \epsilon \frac{(1+R_{t_n})^{\frac{1}{2}+\gamma+\delta}}{(1+R_{st_n})^{\delta+\alpha}(1+R_{st_n-})^{\gamma-\alpha}} \right\}$$

instead of $[\|\Delta M_s^n\| > \epsilon]$ for all $\epsilon \in (0,1]$. Summarizing the above inequalities, choosing $S = \{1\}$ and thus $t = 1$, we get

$$\begin{aligned} x^T x \mathbb{1}_{[\|x\|>\delta]} * \nu_t^{M^n} &= x^T x \mathbb{1}_{[\|x\|>\delta]} * \nu_1^{M^n} = \int_0^1 \int_{\mathbb{R}^d} x^T x \mathbb{1}_{[\|x\|>\delta]} \; \nu^{M^n}(\mathrm{d}s, \mathrm{d}x) \\ &\le \int_0^{t_n} \int_{G_{s,t_n}^\epsilon} \frac{(1+R_s)^{2\delta+2\alpha}(1+R_{s-})^{2\gamma-2\alpha}}{(1+R_{t_n})^{1+2\gamma+2\delta}} x^T x \; \nu^{M^\star}(\mathrm{d}s, \mathrm{d}x) \end{aligned}$$

$$\leq \frac{\int_0^{t_n} \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{G^{\epsilon}_{s,t_n}} x^T x\ \nu^{M^\star}(\mathrm{d}s,\mathrm{d}x)}{(1+R_{t_n})^{1+2\gamma+2\delta}} \xrightarrow{\mathbf{P}} 0 \qquad (n\to\infty).$$

Verification of (ii): We have to show

$$\lceil M^{n^i}, M^{n^j}\rceil_1 \xrightarrow{\mathbf{P}} \lceil M^i, M^j\rceil_1$$

for all $i,j\in\{1,\dots,d\}$. We have

$$\begin{aligned}
&\lceil M^{n^i}, M^{n^j}\rceil_1 \\
&\quad = (1+R_{t_n})^{-1-2\gamma-2\delta}\int_0^{t_n}(1+R_r)^{2\delta+2\alpha}(1+R_{r-})^{2\gamma-2\alpha}\lceil M_i(\mathrm{d}r,0), M_j(\mathrm{d}r,0)\rceil_s \\
&\quad = (1+R_{t_n})^{-1-2\gamma-2\delta}\int_0^{t_n}(1+R_r)^{2\delta+2\alpha}(1+R_{r-})^{2\gamma-2\alpha}h_s^{ij}(0)\,\mathrm{d}R_s \\
&\quad = (1+R_{t_n})^{-1-2\gamma-2\delta}\int_0^{t_n}\Big(1+\frac{\Delta R_r}{1+R_{r-}}\Big)^{2\delta+2\alpha}(1+R_{r-})^{2\gamma-2\alpha+2\delta+2\alpha}h_s^{ij}(0)\,\mathrm{d}R_s \\
&\quad \simeq h^{ij}(0)\frac{\int_0^{t_n}(1+\mathrm{o_b}(1))^{2\delta+2\alpha}(1+R_{r-})^{2\gamma+2\delta}(1+\mathrm{o_b}(1))\,\mathrm{d}R_s}{(1+R_{t_n})^{1+2\gamma+2\delta}} \longrightarrow \frac{h^{ij}(0)}{1+2\gamma+2\delta},
\end{aligned}$$

since $h_s(0)\to h(0)$, $1+2\gamma+2\delta>0$, and $\frac{\int_0^{t_n}(1+R_{r-})^{2\gamma+2\delta}\mathrm{d}R_s}{(1+R_{t_n})^{1+2\gamma+2\delta}} \longrightarrow \frac{1}{1+2\gamma+2\delta}$ hold. □

4.3 Discussion of Special Settings

We consider Theorem 4.2 in the discrete-time and in the Itô setting. It turns out that the result in the discrete-time setting presented here as a corollary coincides with a known result in the literature. In the Itô setting, we obtain a new result.

Corollary 4.1. *We consider the stochastic integral equation* (29) *of the Itô type of Corollary 2.1 with* $\alpha\in(\frac{5}{6},1)$. *Let the Hessian be positive definite at* $x^\star$, *continuous around* $x^\star$ *and the conditions* (*A*), (*B*), (*C*), *and* (*E*′) *be valid. We further assume, for all* $i,j\in\{1,\dots,d\}$,

$$\sum_{j=1}^d \sigma_s^{ij}(x) \leq C(1+\|x\|) \quad \text{and} \quad \lim_{\substack{s\to\infty\\ x\to x^\star}} \sigma_s^{ij}(x) = \sigma^{ij}(x^\star).$$

If f *is three times differentiable at* $x^\star$, $\gamma=\frac{1}{6}$, *and* $\delta>-\frac{2}{3}$ *holds, we get*

$$(1+t)^{\frac{1}{3}}(\bar{Z}_{\delta,t}-x^\star) \xrightarrow{\mathcal{D}} N(\tilde{\mu},\tilde{\Sigma}),$$

where

$$\tilde{\mu} := -\frac{c^2(1+\delta)}{4+6\delta}H_{x^\star}^{-1}\left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star)\right)_{i\in\{1,\dots,d\}} \quad \text{and} \quad \tilde{\Sigma} := \frac{3(1+\delta)^2}{8c^2(2+3\delta)}H_{x^\star}^{-1}h(x^\star)H_{x^\star}^{-1}.$$

Figure 7 shows the density of the limit distribution of the process $(1+t)^{1/3}(Z_t - x^\star)$ considered in Corollary 4.1. The histogram is constructed by evaluation of paths of the corresponding process at time $t = 250$. Also the density of the theoretical limit distribution as given in Corollary 4.1 is presented. We choose $f(x) = 0.5x^T x$, $x \in \mathbb{R}^2$, $a = 3$, $c = 1$, $\alpha = 11/12$, $\gamma = 1/6$, $\sigma = I$, $\delta = -1/3$ and use the Milstein scheme (see, e.g., [14]) to simulate paths of the process Z_t. The starting points are taken from a uniform distribution on $[-4, 4] \times [-4, 4]$.

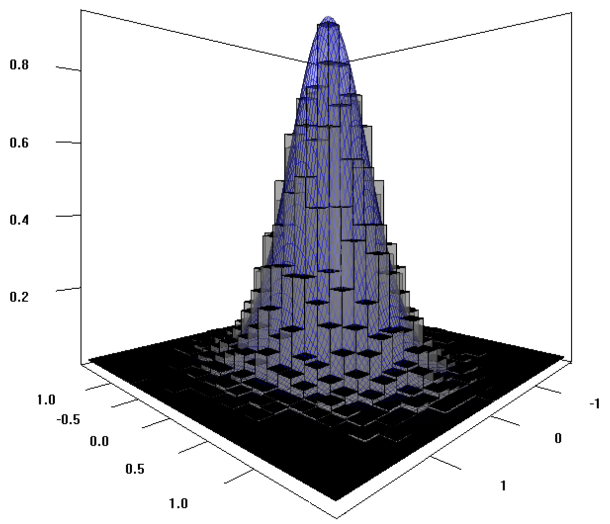

Figure 7: This plot shows the histogram as an empirical approximation of the density of the limit distribution of the process $(1+t)^{1/3}Z_t$ considered in the example below Corollary 4.1 and the density of the theoretical asymptotic distribution overlaid in blue.

Corollary 4.2. *We consider recursion* (30) *of Corollary 2.2 with* $\alpha \in (\frac{5}{6}, 1)$. *Let the Hessian be positive definite at* $x^\star$, *continuous around* $x^\star$, *and conditions* (A), (B), (C), *and* (E'') *be valid. Furthermore, we assume for all* $i, j \in \{1, \dots, d\}$

$$(E''') \qquad \sup_n \mathbf{E}\left(\|V_n\|^2 | \mathcal{F}_{n-1}\right) < \infty,$$

$$(F) \qquad \mathbf{E}\left(V_n^i V_n^j | \mathcal{F}_{n-1}\right) \stackrel{n\to\infty}{\longrightarrow} h^{ij} \quad (1 \le i, j \le d),$$

and the Lyapunov-type condition

$$(L) \qquad \sup_n \mathbf{E}\|V_n\|^{2+\kappa} < \infty \ \text{ for a } \kappa > 0.$$

If f is three times differentiable at $x^\star$, $\gamma = \frac{1}{6}$, and $\delta > -\frac{2}{3}$ hold, we get

$$(1+t)^{\frac{1}{3}}(\bar{Z}_{\delta,n} - x^\star) \xrightarrow{\mathcal{D}} N(\tilde{\mu}, \tilde{\Sigma}),$$

where

$$\tilde{\mu} := -\frac{c^2(1+\delta)}{4+6\delta} H_{x^\star}^{-1} \left(\frac{\partial^3 f}{(\partial x_i)^3}(x^\star) \right)_{i\in\{1,\ldots,d\}} \quad \text{and } \tilde{\Sigma} := \frac{3(1+\delta)^2}{8c^2(2+3\delta)} H_{x^\star}^{-1} h H_{x^\star}^{-1}.$$

Proof of Corollary 4.1

We reuse the notation of Corollary 2.1, Corollary 2.3, and their proofs. Since the Brownian motion is continuous, it certainly fulfills the Lindeberg-type condition, because

$$\mu^W([0,t]\times\Gamma) = \sum_{0<s\le t} \mathbb{1}_{\{\Delta W_s \in \Gamma\}} = \sum_{0<s\le t} \mathbb{1}_{\{0\in\Gamma\}} = \sum_{0<s\le t} 0 = 0 \quad \Gamma \in \mathcal{B}_d(\mathbb{R}\setminus\{0\}) \quad t \in \mathbb{R}_+,$$

which results in $\nu^W([0,t]\times\Gamma) = 0$ for all $\Gamma \in \mathcal{B}_d(\mathbb{R}\setminus\{0\})$ and $t \in \mathbb{R}_+$. Since $R_s := s$ is continuous, we have

$$\Delta R_s = \Delta s = 0 \quad \Longrightarrow \quad \sum_{0\le s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} < \infty.$$

Since $h_s^{ij}(x) = \sum_{k=1}^d \sigma_s^{ik}(x)\sigma_s^{jk}(x)$ holds,

$$\lim_{\substack{s\to\infty \\ x\to x^\star}} h_s^{ij}(x) = \sum_{k=1}^d \lim_{\substack{s\to\infty \\ x\to x^\star}} \left(\sigma_s^{ik}(x)\sigma_s^{jk}(x)\right) = \sum_{k=1}^d \sigma^{ik}(x^\star)\sigma^{jk}(x^\star) = \left(\sigma(x^\star)\sigma^T(x^\star)\right)_{ij}$$
$$=: h^{ij}(x^\star)$$

follows. The assumptions ensure positive definiteness of $H_{x^\star}$ and that conditions (A), (B), (C), and (E') are valid. From a consideration of the assumptions of Corollary 2.3 and its proof, we easily obtain the existence of a $\kappa > 0$ such that $(1+t)^\kappa \|Z_t\| \to 0$ holds. □

Proof of Corollary 4.2

We will verify the assumptions of the Theorem 4.2. For the quadratic variation we get

$$\left[\int_0^\cdot M_i(\mathrm{d}s, x)\right]_t = \left[\int_0^\cdot \tilde{V}_s^i \, \mathrm{d}R_s\right]_t = \sum_{\substack{n\le t \\ n\in\mathbb{N}}} V_n^{i^2} (\Delta R_n)^2 = \sum_{\substack{n\le t \\ n\in\mathbb{N}}} V_n^{i^2},$$

and hence for the predictable quadratic variation

$$\left\lceil \int_0^\cdot M_i(\mathrm{d}s, x)\right\rceil_t = \sum_{\substack{n\le t \\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2} | \mathcal{F}_{n-1}\right) = \sum_{\substack{n\le t \\ n\in\mathbb{N}}} \mathbf{E}\left(V_n^{i^2} | \mathcal{F}_{n-1}\right).$$

Therefore, we have $h_n^{ij} := \mathbf{E}\,(V_n^i V_n^j|\mathcal{F}_{n-1})$ and obtain convergence

$$h_n^{ij} = \mathbf{E}\left(V_n^i V_n^j|\mathcal{F}_{n-1}\right) \overset{n\to\infty}{\longrightarrow} h^{ij}.$$

Furthermore,

$$\sum_{0\le s} \mathbb{1}_{\{a\lambda_i \Delta R_s = (1+R_{s-})\}} = \sum_{j=1}^{\infty} \mathbb{1}_{\{a\lambda_i = j\}} \le a\lambda_{\max} < \infty$$

holds. The assumptions ensure positive definiteness of $H_{x^\star}$ and validity of conditions (A), (B), (C), and (E''). From a consideration of the assumptions of Corollary 2.4 and its proof, we easily obtain the existence of a $\kappa > 0$ such that $n^\kappa \|Z_n\| \to 0$ holds.

Let us now turn to the proof of the condition of Lindeberg type. Considering a triangle scheme, it is well-known that the Lyapunov condition implies the Lindeberg condition. We prove validity of the Lindeberg condition using assumption (L), which corresponds to a Lyapunov condition. We have

$$\begin{aligned}
&\frac{\int_0^t \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{G_{s,t}^\epsilon} \|x\|^2 \ \nu^{M^\star}(\mathrm{d}s,\mathrm{d}x)}{(1+R_t)^{1+2\delta+2\gamma}} \le \frac{\int_0^t \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2 \ \nu^{M^\star}(\mathrm{d}s,\mathrm{d}x)}{(1+R_t)^{1+2\delta+2\gamma}} \\
&= \frac{\int_0^t \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2 N_s(\omega,\mathrm{d}x)\,\mathrm{d}C_s}{(1+R_t)^{1+2\delta+2\gamma}} \le \mathcal{C}\, \frac{\int_0^t \frac{(1+R_s)^{2\delta+2\alpha}}{(1+R_{s-})^{2\alpha-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2 N_s(\omega,\mathrm{d}x)\,\mathrm{d}R_s}{(1+R_t)^{1+2\delta+2\gamma}} \\
&\le \mathcal{C}\, \frac{\sum_{i=1}^{\lfloor t\rfloor} \frac{(1+i)^{2\delta+2\alpha}}{i^{2\alpha-2\gamma}} \int_{\mathbb{R}^d} \|x\|^2 \ N_i(\omega,\mathrm{d}x)}{(1+\lfloor t\rfloor)^{1+2\delta+2\gamma}} \le \mathcal{C}\, \frac{\sum_{i=1}^{\lfloor t\rfloor} i^{2\delta+2\alpha-2\alpha+2\gamma} \int_{\mathbb{R}^d} \|x\|^2 \ \mathbf{P}_{V_i|\mathcal{F}_{i-1}}(\mathrm{d}x)}{\lfloor t\rfloor^{1+2\delta+2\gamma}} \\
&\le \mathcal{C}\, \frac{\sum_{i=1}^{\lfloor t\rfloor}(i+1)^{2\delta+2\gamma}\mathbf{E}\left(\|V_i\|^2|\mathcal{F}_{i-1}\right)}{\lfloor t\rfloor^{1+2\delta+2\gamma}} = \mathcal{C}\, \frac{\sum_{i=1}^{n} i^{2\delta+2\gamma}\mathbf{E}\left(\|V_i\|^2|\mathcal{F}_{i-1}\right)}{n^{1+2\delta+2\gamma}}
\end{aligned} \tag{57}$$

with $t = n \in \mathbb{N}$. Here we used

$$\nu^{M^\star}(\omega,\mathrm{d}t,\mathrm{d}x) = N_t(\omega,\mathrm{d}x)\,\mathrm{d}C_t, \quad \text{where} \ \ C_t = \sum_{i=1}^{d} \lceil M^{\star i} \rceil_t$$

and $N_i(\omega, A) = \mathbf{P}(V_i \in A|\mathcal{F}_{i-1})$. More detailed information on this can be found, for example, in [12], [37], or [38]. To finally ensure that expression (57) converges to zero in probability, we observe

$$\begin{aligned}
\left(\mathbf{E}\frac{\sum_{i=1}^{n} i^{2\delta+2\gamma}\mathbf{E}\left(\|V_i\|^2|\mathcal{F}_{i-1}\right)}{n^{1+2\delta+2\gamma}}\right)^{1+\kappa/2} &\le \mathbf{E}\frac{\sum_{i=1}^{n} i^{(2\delta+2\gamma)(1+\kappa/2)}\mathbf{E}\left(\|V_i\|^{2+\delta}|\mathcal{F}_{i-1}\right)}{n^{(1+2\delta+2\gamma)(1+\kappa/2)}} \\
&= \frac{\sum_{i=1}^{n} i^{(2\delta+2\gamma)(1+\kappa/2)}\mathbf{E}\|V_i\|^{2+\delta}}{n^{(1+2\delta+2\gamma)(1+\kappa/2)}} \le \left(\sup_j \mathbf{E}\|V_j\|^{2+\kappa}\right) \frac{\sum_{i=1}^{n} i^{(2\delta+2\gamma)(1+\kappa/2)}}{n^{(1+2\delta+2\gamma)(1+\kappa/2)}} \\
&\le \mathcal{C}\, \frac{\sum_{i=1}^{n} i^{(2\delta+2\gamma)(1+\kappa/2)}}{n^{(1+2\delta+2\gamma)(1+\kappa/2)}} \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

which follows by Kronecker's lemma and

$$\sum_{i=1}^{\infty} \frac{i^{(2\delta+2\gamma)(1+\kappa/2)}}{i^{(1+2\delta+2\gamma)(1+\kappa/2)}} = \sum_{i=1}^{\infty} i^{-(1+\kappa/2)} < \infty \quad \Longrightarrow \quad \lim_{n\to\infty} \frac{\sum_{i=1}^{n} i^{(2\delta+2\gamma)(1+\kappa/2)}}{n^{(1+2\delta+2\gamma)(1+\kappa/2)}} = 0.$$

□

5 Asymptotic Properties of the Averaged Process Using Constant Weights

5.1 Asymptotic Properties of the Averaged Robbins-Monro Process in the Linear Case

In this section we discuss what happens in the case of a constant weight $a_s := a$. To study this more precisely, we consider a continuous-time version of a linear Robbins-Monro process that is the solution of the stochastic integral equation

$$Z_t = Z_0 - a \int_0^t (AZ_{s-} - b)\,\mathrm{d}R_s - aM_t. \tag{58}$$

As before we assume the existence of a strong solution on $[0, \infty)$. We call this solution the linear Robbins-Monro process with constant weights.

5.1.1 Unavailable Consistency

For weights of the form $a_s := a(1+R_{s-})^\alpha$ where $0 < \alpha \leq 1$ instead of a constant weight a, the solution converges, under some assumptions, almost surely to the solution $x^\star = A^{-1}b$. But what happens in the case of constant weights? To answer this question, we consider the process under simple conditions. In doing so, we study the one-dimensional version with a matrix A and a vector b that are degenerated to the value 1 and 0, respectively. Furthermore, let Z_0 be deterministic, $R_s := \lfloor s \rfloor$, and $M_s := \sum_{i=1}^{\lfloor s \rfloor} X_i$ with i.i.d sequence (X_i), which satisfies $\mathbf{P}[X_i = 1] = \frac{1}{2}$ and $\mathbf{P}[X_i = -1] = \frac{1}{2}$. We have

$$Z_n = Z_0 - a\sum_{i=1}^{n} Z_{i-1} - aM_n = Z_0 - a\sum_{i=0}^{n-1} Z_i - a\sum_{i=0}^{n-1} X_i.$$

Using

$$Z_{n+1} - Z_n = -aZ_n - aX_n \quad \Longrightarrow \quad Z_{n+1} = (1-a)Z_n - aX_n,$$

the recursion can be solved by

$$Z_n = (1-a)^n Z_0 - a\sum_{i=1}^{n}(1-a)^{i-1} X_{n-i}.$$

Since we are interested in the asymptotic behavior and in the consistency in particular, we consider the first two moments of Z_n instead of almost sure convergence. We have

$$\mathbf{E}Z_n = (1-a)^n Z_0 - a\sum_{i=1}^{n}(1-a)^{i-1}\mathbf{E}X_{n-i} = (1-a)^n Z_0 \overset{n\to\infty}{\longrightarrow} 0,$$

if $|1-a|<1$ holds. Furthermore, we find

$$\begin{aligned}\mathbf{E}(Z_n-\mathbf{E}Z_n)^2 &= \mathbf{E}\left(-a\sum_{i=1}^{n}(1-a)^{i-1}X_{n-i}\right)^2\\ &= a^2\sum_{i,j=1}^{n}(1-a)^{i+j-2}\underbrace{\mathbf{E}\left(X_{n-i}X_{n-j}\right)}_{=0\ \text{for}\ i\neq j} = a^2\sum_{i=1}^{n}(1-a)^{2(i-1)}\mathbf{E}X_{n-i}^2\\ &= a^2\sum_{i=0}^{n-1}(1-a)^{2i} = a^2\frac{1-(1-a)^{2n}}{2a-a^2}\to\frac{a}{2-a}\neq 0.\end{aligned}$$

Therefore, Z_n is an asymptotically unbiased estimator for $x^\star$, which equals 0 in our example. However, the variance of Z_n does not converge to 0 and hence the process cannot be a consistent estimator for $x^\star$. More precisely, the process fluctuates around the point we are looking for. The same behavior is visible in the continuous-time setting. At this point, we content ourselves to Figure 8. This figure shows a simulated path of the process Z_t of (58) in the Itô setting. The large-scale fluctuation of the process Z_t around 0 is clearly visible. We choose $A=I$, $b=0$, $Z_0=(5,6)^T$, $a_s=1$, $M_t=W_t$ with a two-dimensional Brownian motion W_t and use the Milstein scheme (see, e.g., [14]) to simulate a path of the resulting process Z_t. As this simple version of a Robbins-Monro process is not consistent, further research in this direction is superfluous, not to mention a Kiefer-Wolfowitz version. Surprisingly, in considering the averaged process we will find some nice features. The question of the asymptotic behavior of the averaged process will be answered in the next section.

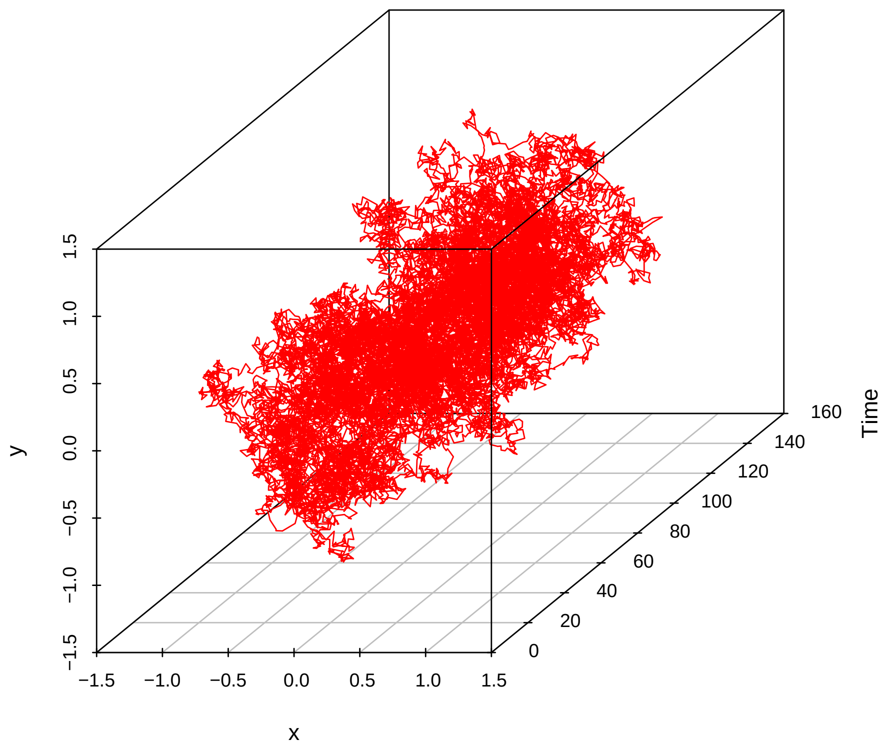

Figure 8: Simulation of a path of the Robbins-Monro process (58) in the Itô setting.

5.1.2 Asymptotic Normality

In the previous section we dwelt on the question of whether the linear Robbins-Monro process with constant weights, which is the solution of the integral equation

$$Z_t = Z_0 - a\int_0^t (AZ_{s-} - b)\,\mathrm{d}R_s - aM_t,$$

is consistent. Here, we investigate the averaged process, which we obtain using the above process. Theorem 5.1 states a result for asymptotic normality of the averaged process (59), the appearing rate described by $\sqrt{R_t}$ is the maximum achievable.

Remark 5.1. We emphasize that an advantage of using constant weights is the faster movement. If we use common weights $a_t := a(1+R_{t-})^{\alpha}$ with $0 < \alpha \leq 1$, then the condition $\int_0^\infty a_s\,\mathrm{d}R_s = \infty$ ensures that arbitrary distant points $x^\star$ can be reached. This condition is an essential to obtain consistency. But for decaying weights the process moves slowly for large t or large R_t. Hence a good choice of the starting point is important, because, if it is far away from $x^\star$, the process will still reach $x^\star$, but needs a long time. Of course, in an asymptotic consideration this may be obscured. Thus from a theoretical point of view it is not really a problem, but from a practical one it really matters. As we mentioned above, the process with constant weights has a faster movement but, as a restriction, we have to say that the parameter a, which defines the speed, cannot be chosen freely. We will see that the condition $0 < a < \frac{2}{J\lambda_{\max}}$ ensures the stability of the process if jumps exist. Here we again encounter the problem of maximizing a, while also keeping to the conditions. This problem cannot be solved without knowing $\lambda_{\max}$, which is generally unknown.

In the following, we investigate the averaged process $\bar{Z}_t$

$$\bar{Z}_t := \frac{1}{1+R_t}\int_0^t Z_s \,\mathrm{d}R_s \tag{59}$$

with Z_s as the solution of the stochastic integral equation (58). In view of the objects appearing in (58), we note the following restrictions. Let A be a positive definite matrix such that $T^T A T = D$ holds where D is a diagonal matrix with its diagonal entries λ_i. Furthermore, let M_t be a locally square integrable martingale and R_t be the process described in Chapter 2. In view of applying Theorem 5.1 to a special Kiefer-Wolfowitz process, the assumption on the matrix A is not restrictive, because in such an application the Hessian plays the part of the matrix A. The following theorem states asymptotic normality of the process $\bar{Z}_t$. Figure 9 shows how the process $\bar{Z}_t$ (59) works in the Itô setting. There, a simulated path of the process $\bar{Z}_t$ as well as of the underlying process Z_t of (58) are depicted.

Theorem 5.1. *Let $x^\star$ be the unique solution of $Ax = b$. We assume $a \in \mathbb{R}_+$, $\Delta R_t \leq C(\omega) < \infty$, $a\lambda_j \Delta R_t \neq 1$ for all $j \in \{1, \ldots, d\}$, and*

$$\textbf{(N)} \qquad \frac{a\int_0^t T\phi_s^{-1}\int_s^t \phi_r \,\mathrm{d}R_r\, T^T \,\mathrm{d}M_s}{(1+R_t)^{\frac{1}{2}}} \xrightarrow{\mathcal{D}} N(0,\Sigma),$$

where $\phi_t^{ij} := \mathcal{E}_t(-a\lambda_i R)\,\delta_{ij}$. Then we have

$$(1+R_t)^{\frac{1}{2}}(\bar{Z}_t - x^\star) \xrightarrow{\mathcal{D}} N(0,\Sigma).$$

We discuss condition (N) appearing in Theorem 5.1 in Section 5.3 to get a better understanding of it.

Remark 5.2. In the theorem above, condition (N) saves us some work. Considering the condition in more detail, the question arises of whether it can be simplified. If we had the process ϕ_{r-} instead of ϕ_r in $\int_s^t \phi_r \,\mathrm{d}R_r$, we could easily simplify the condition since

$$\int_s^t \phi_{r-}^{ii} \,\mathrm{d}R_r = \frac{1}{a\lambda_i}(\phi_s - \phi_t)$$

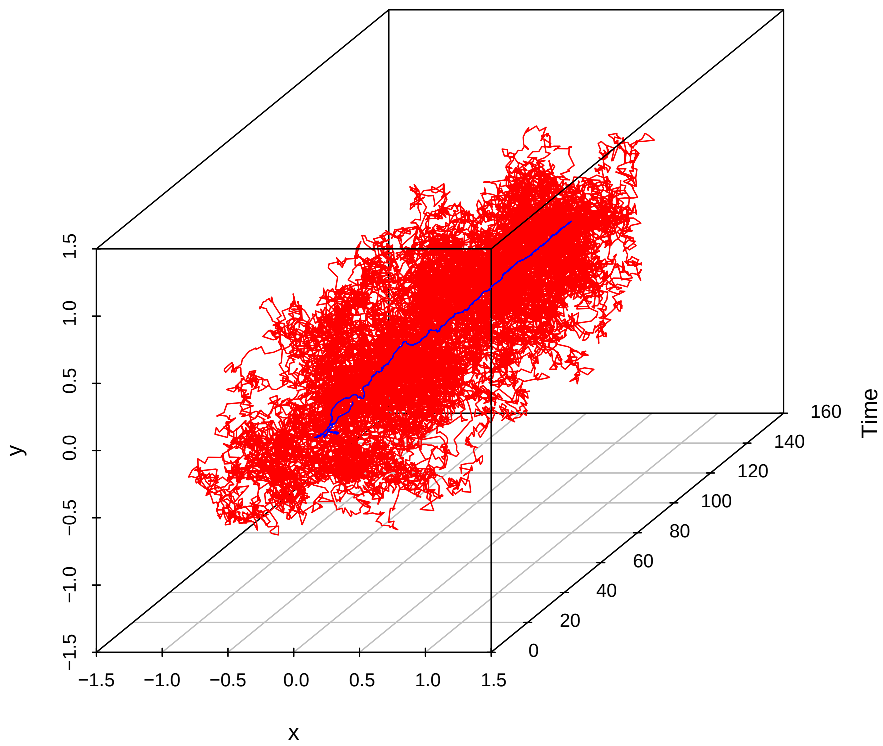

Figure 9: The plot shows the simulated path (red) of the process (58) and the corresponding averaged path (blue) of process (59).

holds. As a result, we will see later on that proving the condition is very easy in the case that R_t is a continuous process or a pure-jump process. In the case of a general process R_t, we have an extra process $\int_s^t \Delta\phi_r \,\mathrm{d}R_r$. If we have no knowledge about the process R_t, ϕ_r is a product of a continuous process and a pure-jump process, hence the calculation of such an integral is not easy. However, if we have some knowledge of the process R_t, it may be possible to calculate the integral.

Remark 5.3. The assumptions that the matrix A is symmetric and $a\lambda_j \Delta R_t \neq 1$ for all $j \in \{1, \dots, d\}$ are just for simplicity and can easily be relaxed.

Proof of Theorem 5.1

To prove asymptotic normality of the process $\bar{Z}_t$, we apply Lemma 3.3 to obtain another representation of (58). Choose

$$(AZ_{s-} - b) \quad \text{instead of} \quad (H_{x^\star} Z_{s-} + V(Z_{s-}, c_s)),$$

$\alpha = \gamma = 0$, $2c = 1$, and $M(\mathrm{d}t, Z_{t-}) = \mathrm{d}M_t$. With these choices the stochastic integral equation (5) leads to (58), which we consider here. We get the representation

$$Z_t = T\phi_t T^T Z_0 - aT\phi_t \int_0^t \phi_s^{-1} T^T \,\mathrm{d}M_s + aT\phi_t \int_0^t \phi_s^{-1} \,\mathrm{d}R_s \, T^T b \tag{60}$$

where the process ϕ_t is given by

$$\phi_t^{ii} := \mathcal{E}_t(-\int_0^\cdot a\lambda_i \,\mathrm{d}R_s) = \mathcal{E}_t(-a\lambda_i R_t).$$

Here λ_i is the i-th eigenvalue of the matrix A and $T^T A T = D$ holds. Furthermore, we have $\phi_{t-}^{ii} = \exp(-a\lambda_i R_t^c) \prod_{s<t}(1 - a\lambda_i \Delta R_s)$ and

$$\begin{aligned} \Delta\phi_t^{ii} &= \phi_t^{ii} - \phi_{t-}^{ii} = \exp(-a\lambda_i R_t^c) \left(\prod_{s<t} (1 - a\lambda_i \Delta R_s) \right) (1 - a\lambda_i \Delta R_t - 1) \\ &= -a\lambda_i \Delta R_t \phi_{t-}^{ii}. \end{aligned} \tag{61}$$

Now we consider the averaged process $\bar{Z}_t$ in more detail. We obtain

$$\begin{aligned} \bar{Z}_t &= (1 + R_t)^{-1} \int_0^t Z_s \,\mathrm{d}R_s \\ &= T\,(1 + R_t)^{-1} \int_0^t \phi_s \,\mathrm{d}R_s T^T Z_0 - aT\,(1 + R_t)^{-1} \int_0^t \phi_s \int_0^s \phi_r^{-1} T^T \,\mathrm{d}M_r \,\mathrm{d}R_s \\ &\quad + (1 + R_t)^{-1} aT \int_0^t \phi_s \int_0^s \phi_r^{-1} \,\mathrm{d}R_r \,\mathrm{d}R_s \, T^T b. \end{aligned}$$

To prove the assertion of the theorem, that is $(1 + R_t)^{\frac{1}{2}}(\bar{Z}_t - x^\star) \xrightarrow{\mathcal{D}} N(0, \Sigma)$, we will show

$$\begin{aligned} &\textbf{(I)} \quad T\,(1 + R_t)^{-\frac{1}{2}} \int_0^t \phi_s \,\mathrm{d}R_s T^T Z_0 \xrightarrow{\mathbf{P}} 0 \\ &\textbf{(II)} \quad (1 + R_t)^{\frac{1}{2}} \left((1 + R_t)^{-1} aT \int_0^t \phi_s \int_0^s \phi_r^{-1} \,\mathrm{d}R_r \,\mathrm{d}R_s \, T^T b - x^\star \right) \xrightarrow{\mathbf{P}} 0 \\ &\textbf{(III)} \quad aT\,(1 + R_t)^{-\frac{1}{2}} \int_0^t \phi_s \int_0^s \phi_r^{-1} T^T \,\mathrm{d}M_r \,\mathrm{d}R_s \xrightarrow{\mathcal{D}} N(0, \Sigma). \end{aligned}$$

Using Slutsky's theorem, we then get the desired result.

Part (I): We have

$$\begin{aligned} |(1 + R_t)^{-\frac{1}{2}} \int_0^t \phi_s^{ii} \,\mathrm{d}R_s| &= |(1 + R_t)^{-\frac{1}{2}} \int_0^t (\phi_{s-}^{ii} + \Delta\phi_s^{ii}) \,\mathrm{d}R_s| \\ &= (1 + R_t)^{-\frac{1}{2}} |\int_0^t (1 - a\lambda_i \Delta R_s) \phi_{s-}^{ii} \,\mathrm{d}R_s| \\ &\le \frac{|1 - a\lambda_i C(\omega)|}{a\lambda_i} \frac{1 - \phi_t^{ii}}{(1 + R_t)^{\frac{1}{2}}} \to 0. \end{aligned}$$

Part (II): We observe

$$\int_0^s \phi_r^{ii^{-1}} \,\mathrm{d}R_r = \frac{1}{a\lambda_i}\int_0^s a\lambda_i \phi_r^{ii^{-1}} \,\mathrm{d}R_r = \frac{1}{a\lambda_i}\int_0^s \mathrm{d}\phi_r^{ii^{-1}} = \frac{\phi_s^{ii^{-1}} - 1}{a\lambda_i}$$

and obtain

$$\begin{aligned}
&\|(1+R_t)^{-1} aT \int_0^t \phi_s \int_0^s \phi_r^{-1} \,\mathrm{d}R_r \,\mathrm{d}R_s \, T^T b - x^\star\| \\
&\quad = \|(1+R_t)^{-1} T \int_0^t \phi_s \,(\phi_s^{-1} - I) D^{-1} \,\mathrm{d}R_s \, T^T b - x^\star\| \\
&\quad \leq \|\frac{R_t}{(1+R_t)} T D^{-1} T^T b - x^\star\| + \|(1+R_t)^{-1} T \int_0^t \phi_s \,\mathrm{d}R_s \, D^{-1} T^T b\| \\
&\quad = \|\frac{R_t}{(1+R_t)} A^{-1} b - x^\star\| + \mathcal{C}_\omega (1+R_t)^{-1} = \|\left(\frac{R_t - (1+R_t)}{(1+R_t)}\right) x^\star\| + \mathcal{C}_\omega (1+R_t)^{-1} \\
&\quad = \mathcal{C}_\omega (1+R_t)^{-1}.
\end{aligned}$$

Part (III): We have to verify

$$a\,(1+R_t)^{-\frac{1}{2}} \int_0^t T\phi_s \int_0^s \phi_r^{-1} T^T \,\mathrm{d}M_r \,\mathrm{d}R_s \xrightarrow{\mathcal{D}} N(0,\Sigma).$$

Here the problem is that the integrand is a local martingale and not the integrator. Hence in this representation it is hard to prove asymptotic normality. Therefore, we will use the product rule to rewrite the above process. We rewrite the process as stated above and explain the steps. First, we consider

$$a \sum_{j,k=1}^{d} T_{ij} T_{jk}^T (1+R_t)^{-\frac{1}{2}} \underbrace{\int_0^t \phi_s^{jj} \int_0^s \phi_r^{jj^{-1}} \,\mathrm{d}M_r^k \,\mathrm{d}R_s}_{=:\, X_t} .$$

Then we deal with the process X_t more precisely using $L_t := \int_0^t \phi_r^{jj^{-1}} \,\mathrm{d}M_r^k$ and $B_t := \int_0^t \phi_s^{jj} \,\mathrm{d}R_s$. Using the representation

$$\begin{aligned}
X_t &= \int_0^t L_s \,\mathrm{d}B_s = \int_0^t L_{s-} \,\mathrm{d}B_s + \int_0^t \Delta L_s \,\mathrm{d}B_s \\
&= B_t L_t - \int_0^t B_{s-} \,\mathrm{d}L_s - [B,L]_t + \int_0^t \Delta L_s \,\mathrm{d}B_s \\
&= B_t L_t - \int_0^t B_{s-} \,\mathrm{d}L_s = \int_0^t (B_t - B_{s-}) \,\mathrm{d}L_s \\
&= \int_0^t \phi_s^{jj^{-1}} \int_s^t \phi_r^{jj} \,\mathrm{d}R_r \,\mathrm{d}M_s^k
\end{aligned}$$

and the assumption (N), we obtain

$$a(1+R_t)^{-\frac{1}{2}}\int_0^t T\phi_s\int_0^s \phi_r^{-1}T^T\,\mathrm{d}M_r\,\mathrm{d}R_s = \underbrace{a(1+R_t)^{-\frac{1}{2}}\int_0^t T\phi_s^{-1}\int_s^t \phi_r\,\mathrm{d}R_r\,T^T\,\mathrm{d}M_s}_{\xrightarrow{\mathcal{D}} N(0,\Sigma)}.$$

Hence we have proven the conclusion of the theorem. □

5.2 Asymptotic Normality of the Averaged Kiefer-Wolfowitz Process in the Case of a Quadratic Regression Function

Here we discuss the averaged Kiefer-Wolfowitz process with constant weights using the discussion of the averaged linear Robbins-Monro process in the previous section. We consider the stochastic integral equation (5) with constant weights $a_t := a$, $c_t := c$ that is $\alpha = \gamma = 0$ and the simplification $M(\mathrm{d}s, Z_{s-}) = \mathrm{d}M_s$. From (5) we get the stochastic integral equation

$$Z_t = Z_0 - a\int_0^t \frac{1}{2c}\left\{f(Z_{s-}+ce_i) - f(Z_{s-}-ce_i)\right\}_{i\in\{1,\ldots,d\}}\mathrm{d}R_s - \int_0^t \frac{a}{2c}\,\mathrm{d}M_s. \tag{62}$$

We focus our attention on quadratic regression functions f with

$$f(x) = x^T Ax + b^T x + d \tag{63}$$

and a symmetric, positive definite matrix A which allows us to apply Theorem 5.1. Then

$$\nabla f(x) = 2Ax + b, \qquad\qquad H = 2A,$$

and there exists a T such that $T^T HT = D$ holds where D is a diagonal matrix with its diagonal entries λ_i. Furthermore, we obtain

$$\frac{1}{2c}\{f(x+ce_i) - f(x-ce_i)\} = \sum_{j=1}^d A_{ji}x_j + \sum_{j=1}^d A_{ij}x_j + b_i = (2Ax+b)_i$$

and thus

$$\left\{\frac{f(Z_{s-}+ce_i) - f(Z_{s-}-ce_i)}{2c}\right\}_{i\in\{1,\ldots,d\}} = HZ_{s-} + b. \tag{64}$$

This representation is useful to apply Theorem 5.1. The following theorem gives the asymptotic behavior of the averaged Kiefer-Wolfowitz process with constant weights in the case of a quadratic regression function.

Theorem 5.2. *Let f be a function of type* (63) *with the stationary point $x^\star$ and $\bar{Z}_t$ be the averaged process* (59)*, which uses the unique solution of* (62)*. We assume $a \in \mathbb{R}_+$, $c \in \mathbb{R}_+$, $\Delta R_t \leq C(\omega) < \infty$, $a\lambda_j \Delta R_t \neq 1$ for all $j \in \{1, \dots, d\}$, and*

$$\text{(M)} \qquad \frac{\frac{a}{2c} \int_0^t T\phi_s^{-1} \int_s^t \phi_r \,\mathrm{d}R_r \,T^T \,\mathrm{d}M_s}{(1+R_t)^{\frac{1}{2}}} \xrightarrow{\mathcal{D}} N(0, \Sigma)$$

where $\phi_t^{ij} := \mathcal{E}_t(-a\lambda_i R)\,\delta_{ij}$. Then we have

$$(1+R_t)^{\frac{1}{2}}(\bar{Z}_t - x^\star) \xrightarrow{\mathcal{D}} N(0, \Sigma).$$

Remark 5.4. In view of the validity of (M) and the precise form of Σ the figure of R_t is crucial. In Corollary 5.1 and its proof is a verification of (N), that is nearly the same as (M), done in the standard setting, which means the recursion with $R_t = \lfloor t \rfloor$, in the Itô setting with $R_t = t$ and in a further setting. As we considered this condition in Corollary 5.1 we do not focus our attention on it again.

Remark 5.5. Considering (M) we recognize the influence of $\frac{1}{2c}$ on the matrix Σ.

Proof of Theorem 5.2
Without loss of generality we assume $x^\star = 0$ and therefore implicitly $b = 0$. Since we consider $\bar{Z}_t - x^\star$, this is not restrictive. Furthermore we assume in a first step $2c = 1$. Applying (64) to (62) leads to

$$Z_t = Z_0 - a \int_0^t H Z_{s-} \mathrm{d}R_s - aM_t,$$

which equals (58) using H instead of A. We apply Theorem 5.1 on Z_t. Validity of the assumptions in Theorem 5.1 follows from the assumptions in Theorem 5.2 directly. In the next step we examine the case of a general c and get

$$Z_t = Z_0 - a \int_0^t H Z_{s-} \mathrm{d}R_s - \frac{a}{2c} M_t.$$

Considering the proof of Theorem 5.1, we find the process $\frac{a}{2c} M_t$ instead of aM_t, which effects only assumption (N). Hence we have proven the theorem since the assumption (M) includes the term $\frac{a}{2c}$ and the assumption (N) includes the term a. □

5.3 Discussion of Special Settings

Here we discuss Theorem 5.1 in the case of special figures of R_t and present the results in Corollary 5.1. In these special settings we eliminate the technical condition (N), which we need in Theorem 5.1.

Corollary 5.1. *We consider Theorem 5.1 in the case of a*

(a) continuous deterministic process $(R_t)_{t\geq 0}$*, that is,* $\Delta R_s \equiv 0$.

(b) deterministic pure-jump process $(R_t)_{t\geq 0}$ *with fixed jump heights* J*, that is,* $R^c_s \equiv 0$ *and* $\Delta R_s = J\mathbb{1}_{\{\Delta R_s \neq 0\}}$*. Here we assume* $a < \frac{2}{\lambda_{\max} J}$.

(c) process $R_t := t + \lfloor t \rfloor$*, which is a sum of a continuous and a pure-jump process, a degenerated matrix* $A = \lambda$*, and a one-dimensional process* M_t*. Here we assume* $a < \frac{2}{\lambda}$.

If in each case we have

$$(L) \qquad \frac{\int_0^t \int_{G_t^\epsilon} \|x\|^2 \, \nu^M(\mathrm{d}s, \mathrm{d}x)}{1 + R_t} \xrightarrow{\mathbf{P}} 0 \quad (t \to \infty)$$

for all $\epsilon \in (0,1)$ *where* $G_t^\epsilon := \{x \in \mathbb{R}^d : \|x\| > \epsilon(1+R_t)^{\frac{1}{2}}\}$*,* h_s *is bounded and* $\lim_{s\to\infty} h_s = h$ *with* $\lceil M \rceil_t = \int_0^t h_s \, \mathrm{d}R_s$*, then condition (N) is fulfilled. We further have*

$$(1+R_t)^{\frac{1}{2}}(\bar{Z}_t - x^\star) \xrightarrow{\mathcal{D}} N(0, \Sigma)$$

where $\Sigma := A^{-1} h A^{-1}$ *in part (a), part (b), and*

$$\Sigma := \left(\lambda^{-2} - \frac{a}{\lambda(1-a\lambda)} + a^3\lambda \frac{e^{2a\lambda}(1+a^2\lambda^2) - (1-a\lambda)^2}{4(1-a\lambda)^2(1-a\lambda-e^{a\lambda})^2}\right) h$$

in part (c).

Figure 10 shows the density of the limit distribution of the processes $(1+t)^{1/2}(\bar{Z}_t - x^\star)$ considered in Corollary 5.1. The histogram is constructed by evaluation of paths of the corresponding process at the time $t = 250$. The density of the theoretical limit distribution as given in Corollary 5.1 are presented as well. We choose $A = 0.5I$, $a = 2$ and use the Milstein scheme (see, e.g., [14]) to simulate paths of the process Z_t. The starting points are taken from a uniform distribution on $[-4,4] \times [-4,4]$.

Remark 5.6. We see that the Itô version of (58) is included, since $R_t = t$ is continuous and the Brownian motion that appears as a martingale is also continuous. Furthermore, the standard recursion of (58) is included, since $R_t = \lfloor t \rfloor$ is a pure-jump process with the unique jump height 1. To get an indication of how to verify the Lindeberg-type condition (L), it is helpful to consider Corollaries 3.2, 4.2, and their proofs.

Remark 5.7. In view of the proof of Theorem 5.1 and Corollary 5.1, it is easy to see that the asymptotic results also hold for processes with jumps that “die” out.

Remark 5.8. In part (c) of Corollary 5.1 we considered the 1-dimensional case. With exactly the same ideas and methods it is possible to prove a d-dimensional version of the statement. In the case of a continuous process (see (a)) or a jump process (see (b)) we see that the constant weight a does not influence the variance matrix. This is a well-known phenomenon in the discrete-time setting (see [11] and [32]). But considering a mixture as in part (c), the special feature is that the constant weight a influences the matrix.

Remark 5.9. The boundedness condition on h_s is just for simplicity and can easily be relaxed.

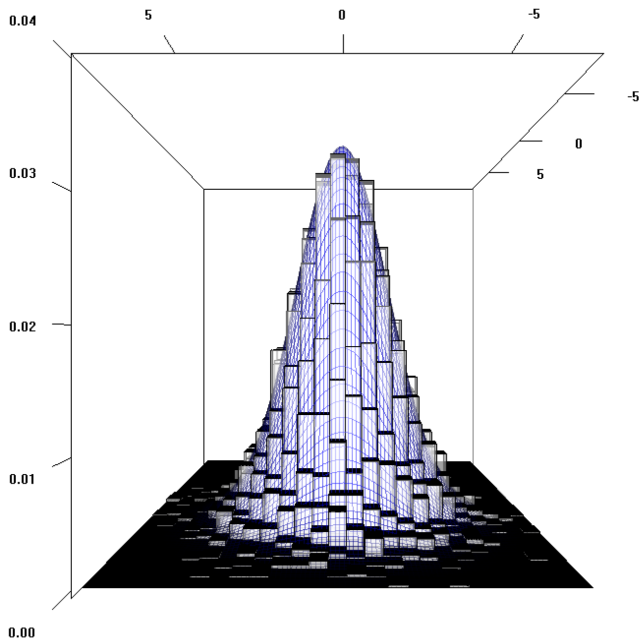

Figure 10: This plot shows the histogram as an empirical approximation of the density of the limit distribution of the process considered in Corollary 5.1(a) and the density of the theoretical asymptotic distribution overlaid in blue.

Proof of Corollary 5.1

We will verify the corollary using Theorem 5.1. Essentially we have to prove condition (N). First we consider ϕ_t, which appears in (N), in more detail. Let J be the jump height of the process R_s. Using $\Delta R_s = J\mathbb{1}_{\{\Delta R_s \neq 0\}}$ and $\#_s := \frac{R_s^d}{J}$, we get

$$\begin{aligned}\phi_t^{ii} &= \mathcal{E}_t(-a\lambda_i R_t) = \exp(-a\lambda_i R_t) \prod_{s\leq t} (1 + \Delta(-a\lambda_i R_s)) \exp(-\Delta(-a\lambda_i R_s)) \\ &= \exp(-a\lambda_i R_t) \prod_{\substack{s\leq t \\ \Delta R_s \neq 0}} (1 - a\lambda_i J) \prod_{r\leq t} \exp(a\lambda_i \Delta R_r)\end{aligned}$$

$$
\begin{aligned}
&= \exp(-a\lambda_i R_t)(1 - a\lambda_i J)^{\#_t} \exp(a\lambda_i \sum_{r\le t} \Delta R_r) \\
&= \exp(-a\lambda_i R_t)(1 - a\lambda_i J)^{\#_t} \exp(a\lambda_i R_t^d) = \exp(-a\lambda_i R_t^c)\,(1 - a\lambda_i J)^{\#_t}, \qquad (65)
\end{aligned}
$$

and hence

$$
\phi_t^{ii} = \exp(-a\lambda_i R_t^c)\,(1 - a\lambda_i J)^{\#_t} \quad \text{where} \quad \begin{cases} (1 - a\lambda_i J)^{\#_t} \to 0 & \text{if } \#_t \to \infty \\ (1 - a\lambda_i J)^{\#_t} \to G < \infty & \text{if } \#_t < K < \infty. \end{cases}
$$

It still remains to prove the indicated convergences. Here we distinguish between $\#_t \to \infty$ and $\#_t < K$. In the case $\#_t < K$ there is nothing to verify. In the case $\#_t \to \infty$ we have

$$
(1 - a\lambda_i J)^{\#_t} \to 0 \qquad \iff \qquad |1 - a\lambda_i J| < 1.
$$

We use

$$
0 < a < \frac{2}{\lambda_{\max} J} \qquad \Longrightarrow \qquad |1 - a\lambda_i J| < 1
$$

and obtain convergence. Furthermore, we have $\phi_{t-}^{ii} = \exp(-a\lambda_i R_t^c)\prod_{s<t}(1 - a\lambda_i \Delta R_s)$,

$$
\begin{aligned}
\Delta\phi_t^{ii} = \phi_t^{ii} - \phi_{t-}^{ii} &= \exp(-a\lambda_i R_t^c)\left(\prod_{s<t}(1 - a\lambda_i \Delta R_s)\right)(1 - a\lambda_i \Delta R_t - 1) \\
&= -a\lambda_i \Delta R_t \phi_{t-}^{ii}, \qquad (66)
\end{aligned}
$$

and ${\phi_t^{ii}}^{-1} = \exp(a\lambda_i R_t^c)\,(1 - a\lambda_i J)^{-\#_t}$.

(a) Case: R_t is a continuous deterministic process. This is the easiest case as we already mentioned in the remark. Since the process R_t is continuous, we get for condition (N)

$$
\begin{aligned}
a(1 + R_t)^{-\frac{1}{2}} \int_0^t T\phi_s^{-1} \underbrace{\int_s^t \phi_r \,\mathrm{d}R_r}\, T^T \,\mathrm{d}M_s& \\
= \int_s^t \phi_{r-} \,\mathrm{d}R_r = -\frac{1}{a} D^{-1} \int_s^t \mathrm{d}\phi_r = \frac{1}{a}(D^{-1}\phi_s - D^{-1}\phi_t)& \\
= (1 + R_t)^{-\frac{1}{2}} \int_0^t T\phi_s^{-1} D^{-1}\phi_s\, T^T \,\mathrm{d}M_s - (1 + R_t)^{-\frac{1}{2}} \int_0^t T\phi_s^{-1} D^{-1}\phi_t\, T^T \,\mathrm{d}M_s& \\
\overset{(\star)}{=} \underbrace{TD^{-1}T^T}_{A^{-1} =}\ \underbrace{\frac{\int_0^t \mathrm{d}M_s}{(1 + R_t)^{\frac{1}{2}}}}_{\xrightarrow{\mathcal{D}} N(0,h)} - TD^{-1} \underbrace{\frac{\phi_t \int_0^t \phi_s^{-1}\, T^T \,\mathrm{d}M_s}{(1 + R_t)^{\frac{1}{2}}}}_{\xrightarrow{\mathbf{P}} 0}.&
\end{aligned}
$$

Here $(\star)$ holds, since ϕ, D, and D^{-1} are diagonal matrices. Using Slutsky's theorem, we get the desired result $\xrightarrow{\mathcal{D}} N(0, \Sigma)$, where $\Sigma = A^{-1}hA^{-1}$. First, we verify convergence in probability to 0 of the third indicated term. We use an L^2-view, because we want to apply the Itô isometry. It is sufficient to verify

$$\frac{\mathbf{E}\left(\int_0^t \phi_s^{jj^{-1}}\,\mathrm{d}M_s^k\right)^2}{(1+R_t)\phi_t^{jj^{-2}}} = \frac{\mathbf{E}\lceil \int_0^t \phi_s^{jj^{-1}}\,\mathrm{d}M_s^k\rceil}{(1+R_t)\phi_t^{jj^{-2}}} = \frac{\mathbf{E}\int_0^t \phi_s^{jj^{-2}} h_s^{kk}\,\mathrm{d}R_s}{(1+R_t)\phi_t^{jj^{-2}}} \le \mathcal{C}\frac{\int_0^t \phi_s^{jj^{-1}}\,\mathrm{d}R_s}{(1+R_t)\phi_t^{jj^{-1}}}$$
$$= \mathcal{C}\frac{\phi_t^{jj^{-1}}}{(1+R_t)\phi_t^{jj^{-1}}} = \mathcal{C}(1+R_t)^{-1} \to 0.$$

Now we verify convergence in distribution of the second indicated term. To this end we restrict ourselves to sequences and use a variant of a central limit theorem (see part (III) of the proof of Theorem 3.2 or 4.2). Mainly we have to prove the convergence of the predictable quadratic variation

$$\left\lceil \frac{\int_0^{\cdot t_n}\mathrm{d}M_s}{(1+R_{t_n})^{\frac{1}{2}}}\right\rceil_1^{ij} = \frac{\int_0^{t_n}\mathrm{d}\lceil M\rceil_s}{(1+R_{t_n})} = \frac{\int_0^{t_n} h_s^{ij}\,\mathrm{d}R_s}{(1+R_{t_n})} = h^{ij}\frac{\int_0^{t_n}(1+\mathrm{o_b}(1))\,\mathrm{d}R_s}{(1+R_{t_n})} \xrightarrow{\mathbf{P}} h^{ij}$$

and the Lindeberg-type condition. With $M_s^n := (1+R_{t_n})^{-\frac{1}{2}} M_{st_n}$ we get for $\epsilon \in (0,1)$

$$\int_0^1 \int_{\mathbb{R}^d} \|x\|^2 \mathbb{1}_{\{\|x\|>\epsilon\}}\,\nu^{M^n}(\mathrm{d}s,\mathrm{d}x) \le \frac{\int_0^{t_n}\int_{G_{t_n}^\epsilon}\|x\|^2\,\nu^M(\mathrm{d}s,\mathrm{d}x)}{1+R_{t_n}} \xrightarrow{\mathbf{P}} 0$$

where $G_{t_n}^\epsilon := \{x \in \mathbb{R}^d \,:\, \|x\| > \epsilon(1+R_{t_n})^{\frac{1}{2}}\}$.

(b) Case: R_t is a deterministic pure-jump process with fixed jump heights J. The process R_t is a pure-jump process with constant jump height J, that is, $\Delta R_s = J\mathbb{1}_{\{\Delta R_s \neq 0\}}$. As we have already mentioned, we have to use a different strategy to in part (a) because here we have jumps and hence we get a further process $\int_s^t \Delta\phi_r\,\mathrm{d}R_r$, which creates some problems. We have

$$a(1+R_t)^{-\frac{1}{2}}\int_0^t T\phi_s^{-1}\int_s^t \phi_r\,\mathrm{d}R_r\,T^T\,\mathrm{d}M_s$$
$$= \frac{\int_0^t T\left(a\phi_s^{-1}\int_s^t\phi_r\,\mathrm{d}R_r - D^{-1}\right)T^T\,\mathrm{d}M_s}{(1+R_t)^{\frac{1}{2}}} + \frac{\int_0^t TD^{-1}T^T\,\mathrm{d}M_s}{(1+R_t)^{\frac{1}{2}}}$$
$$= \underbrace{\frac{\int_0^t T\left(a\phi_s^{-1}\int_s^t\phi_r\,\mathrm{d}R_r - D^{-1}\right)T^T\,\mathrm{d}M_s}{(1+R_t)^{\frac{1}{2}}}}_{\xrightarrow{\mathbf{P}} 0} + \underbrace{A^{-1}\frac{M_t - M_0}{(1+R_t)^{\frac{1}{2}}}}_{\xrightarrow{\mathcal{D}} N(0,h)}.$$

Thus we get the desired conclusion $\xrightarrow{\mathcal{D}} N(0,h)$, where $\Sigma = A^{-1}hA^{-1}$ using Slutsky's theorem. The proof of convergence in distribution can essentially be done as in (a). Now we verify convergence in probability to 0. We use a L^2-argument and obtain

$$\lim_t \frac{\mathbf{E}\left(\int_0^t \left(\, a\phi_s^{ii^{-1}}\int_s^t \phi_r^{ii}\,\mathrm{d}R_r - \frac{1}{\lambda_i}\,\right)\mathrm{d}M_s^k\right)^2}{(1+R_t)}$$

$$= \lim_t \frac{\mathbf{E} \int_0^t \left(a\phi_s^{ii^{-1}} \int_s^t \phi_r^{ii} \, \mathrm{d}R_r - \frac{1}{\lambda_i} \right)^2 h_s^{kk} \, \mathrm{d}R_s}{(1+R_t)}$$

$$= \mathcal{C} \lim_t \frac{\int_0^t \left(a\phi_s^{ii^{-1}} \int_s^t \phi_r^{ii} \, \mathrm{d}R_r - \frac{1}{\lambda_i} \right)^2 \mathrm{d}R_s}{(1+R_t)} \stackrel{(\star)}{=} 0.$$

The verification of ($\star$) remains. Usage of the fact that R_t is a jump process, (65), and (66) gives

$$\phi_t^{ii} = (1 - Ja\lambda_i)^{\#_t} \qquad \text{and} \qquad \Delta\phi_t^{ii} = -a\lambda_i \Delta R_t \phi_{t-}^{ii}.$$

By summing up a finite geometric series, we obtain

$$\begin{aligned}
\phi_s^{ii^{-1}} \int_s^t \phi_r^{ii} \, \mathrm{d}R_r &= J(1 - Ja\lambda_i)^{-\#_s} \sum_{\substack{s \le r \le t \\ \Delta R_r \neq 0}} (1 - Ja\lambda_i)^{\#_r} = J(1 - Ja\lambda_i)^{-\#_s} \sum_{k=\#_s}^{\#_t} (1 - Ja\lambda_i)^k \\
&= J(1 - Ja\lambda_i)^{-\#_s} \left(\sum_{k=0}^{\#_t} (1 - Ja\lambda_i)^k - \sum_{k=0}^{\#_s - 1} (1 - Ja\lambda_i)^k \right) \\
&= \frac{J}{Ja\lambda_i} (1 - Ja\lambda_i)^{-\#_s} \left((1 - Ja\lambda_i)^{\#_s} - (1 - Ja\lambda_i)^{\#_t + 1} \right) \\
&= \frac{1}{a\lambda_i} \left(1 - (1 - Ja\lambda_i)^{\#_t - \#_s + 1} \right),
\end{aligned}$$

and hence

$$\phi_s^{ii^{-1}} \int_s^t \phi_r^{ii} \, \mathrm{d}R_r = \frac{1}{a\lambda_i} - \frac{1}{a\lambda_i} (1 - Ja\lambda_i)^{\#_t - \#_s + 1},$$

if $\#_{s-} \neq \#_t$ holds, and 0 otherwise. Now convergence follows

$$\begin{aligned}
\frac{\int_0^t \left(a\phi_s^{ii^{-1}} \int_s^t \phi_r^{ii} \, \mathrm{d}R_r - \frac{1}{\lambda_i} \right)^2 \mathrm{d}R_s}{(1+R_t)} &= \frac{J}{a^2\lambda_i^2} \frac{\sum_{0 \le s \le t} \mathbb{1}_{\{\Delta R_s \neq 0\}} (1 - Ja\lambda_i)^{2\#_t - 2\#_s + 2}}{(1+R_t)} \\
&= \mathcal{C} \frac{\sum_{k=0}^{\#_t} (1 - Ja\lambda_i)^{2k}}{(1+R_t)} = \mathcal{C}(1+R_t)^{-1} \stackrel{t\to\infty}{\longrightarrow} 0,
\end{aligned}$$

because the implication

$$\#_t = \#_{s-} \quad \Rightarrow \quad R_r \equiv R_r^d \text{ is constant on } [s,t]$$

is true.

(c) Case: $R_t := t + \lfloor t \rfloor$. We consider the process $R_t = t + \lfloor t \rfloor$, which is a sum of a continuous process $R_t^c = t$ and a pure-jump process $R_t^d = \lfloor t \rfloor$. We have

$$\int_s^t \phi_r \, \mathrm{d}R_r = \int_s^t \phi_{r-} \, \mathrm{d}R_r + \int_s^t \Delta\phi_r \, \mathrm{d}R_r = \frac{\phi_s}{a\lambda} - \frac{\phi_t}{a\lambda} - a\lambda \int_s^t \Delta R_r \phi_{r-} \, \mathrm{d}R_r$$

$$= \frac{\phi_s}{a\lambda} - \frac{\phi_t}{a\lambda} - a\lambda \sum_{j=\lceil s\rceil}^{\lfloor t\rfloor} e^{-a\lambda j}(1-a\lambda)^{j-1}$$
$$= \frac{\phi_s}{a\lambda} - \frac{\phi_t}{a\lambda} + \tilde{c}\left(\frac{1-a\lambda}{e^{a\lambda}}\right)^{\lceil s\rceil} - \tilde{c}\left(\frac{1-a\lambda}{e^{a\lambda}}\right)^{\lfloor t\rfloor+1}$$

where $\tilde{c} := \frac{a\lambda e^{a\lambda}}{(1-a\lambda)(1-a\lambda-e^{a\lambda})}$, and thus

$$\phi_s^{-1}\int_s^t \phi_r \,\mathrm{d}R_r = \frac{1}{a\lambda} + \tilde{c}e^{-a\lambda(\lceil s\rceil - s)}(1-a\lambda)^{\lceil s\rceil-\lfloor s\rfloor} - \frac{\phi_s^{-1}\phi_t}{a\lambda} - \tilde{c}\phi_s^{-1}\left(\frac{1-a\lambda}{e^{a\lambda}}\right)^{\lfloor t\rfloor+1}.$$

Here we see clearly why the situation changes, if we consider a process which is a mixture of a continuous and a pure-jump process. Now we separate the first two and the last two processes, since the first two processes influence the matrix and the last two do not. We write

$$a(1+R_t)^{-\frac{1}{2}}\int_0^t \phi_s^{-1}\int_s^t \phi_r\,\mathrm{d}R_r\,\mathrm{d}M_s = \tilde{M}_t - \hat{M}_t^1 - \hat{M}_t^2,$$

where

$$\tilde{M}_t := a(1+R_t)^{-\frac{1}{2}}\int_0^t \left(\frac{1}{a\lambda} + \tilde{c}\,e^{-a\lambda(\lceil s\rceil - s)}(1-a\lambda)^{\lceil s\rceil-\lfloor s\rfloor}\right)\mathrm{d}M_s$$
$$\hat{M}_t^1 := \lambda^{-1}(1+R_t)^{-\frac{1}{2}}\int_0^t \phi_s^{-1}\phi_t\,\mathrm{d}M_s$$
$$\hat{M}_t^2 := a\tilde{c}(1+R_t)^{-\frac{1}{2}}\int_0^t \phi_s^{-1}\left(\frac{1-a\lambda}{e^{a\lambda}}\right)^{\lfloor t\rfloor+1}\mathrm{d}M_s.$$

We will show

$$(i)\;\; \hat{M}_t^1 \xrightarrow{\mathbf{P}} 0 \qquad (ii)\;\; \hat{M}_t^2 \xrightarrow{\mathbf{P}} 0 \qquad (iii)\;\; \tilde{M}_t \xrightarrow{\mathcal{D}} N(0,\Sigma).$$

Analyzing the proof of part (a) and part (b), (i) follows easily. We consider

$$\phi_s^{-1}\left(\frac{1-a\lambda}{e^{a\lambda}}\right)^{\lfloor t\rfloor+1} = \left(\frac{1-a\lambda}{e^{a\lambda}}\right)\phi_s^{-1}e^{-a\lambda\lfloor t\rfloor}(1-a\lambda)^{\lfloor t\rfloor}$$
$$= \left(\frac{1-a\lambda}{e^{a\lambda}}\right)\phi_s^{-1}\phi_t e^{-a\lambda(t-\lfloor t\rfloor)} \le \mathcal{C}\,\phi_s^{-1}\phi_t$$

to verify (ii) and it turns out that (i) implies (ii). The proof of (iii) remains. Here we take the scheme we used several times before, that is, we restrict ourselves to sequences and apply Theorem 4 on page 435 in [19]. We consider the sequence of locally square integrable martingales

$$\tilde{M}_s^n := a(1+R_{t_n})^{-\frac{1}{2}}\int_0^{st_n}\left(\frac{1}{a\lambda} + \tilde{c}\,e^{-a\lambda(\lceil r\rceil - r)}(1-a\lambda)^{\lceil r\rceil-\lfloor r\rfloor}\right)\mathrm{d}M_r.$$

The predictable quadratic variation satisfies

$$\lceil \tilde{M}^n \rceil_1 = a^2(1+R_{t_n})^{-1} \int_0^{t_n} \left(\frac{1}{a\lambda} + \tilde{c} e^{-a\lambda(\lceil s \rceil - s)} (1-a\lambda)^{\lceil s \rceil - \lfloor s \rfloor} \right)^2 h_s \, \mathrm{d}R_s.$$

Since $(1+R_{t_n})^{-1}$ appears, $h_s \to h$, and h is bounded, we substitute h_s by h. We expand the integrand and obtain afterwards three processes. In view of the first process, we get

$$\frac{1}{\lambda^2} \frac{\int_0^{t_n} h_s \, \mathrm{d}R_s}{1+R_{t_n}} \to \frac{h}{\lambda^2}.$$

In view of the second process, we have

$$\begin{aligned}
&2\frac{ah\tilde{c}}{\lambda} \frac{\int_0^{t_n} e^{-a\lambda(\lceil s \rceil - s)} (1-a\lambda)^{\lceil s \rceil - \lfloor s \rfloor} \, \mathrm{d}R_s}{1+R_{t_n}} \\
&\quad = 2\frac{ah\tilde{c}}{\lambda} \left(\frac{\int_0^{t_n} e^{-a\lambda(\lceil s \rceil - s)} (1-a\lambda)^{\lceil s \rceil - \lfloor s \rfloor} \, \mathrm{d}s}{1+R_{t_n}} + \frac{\int_0^{t_n} e^{-a\lambda(\lceil s \rceil - s)} (1-a\lambda)^{\lceil s \rceil - \lfloor s \rfloor} \, \mathrm{d}\lfloor s \rfloor}{1+R_{t_n}} \right) \\
&\quad \overset{(\star)}{=} 2\frac{ah\tilde{c}}{\lambda} \left((1-a\lambda) \frac{\sum_{i=1}^{\lfloor t_n \rfloor} \int_{i-1}^{i} e^{-a\lambda(\lceil s \rceil - s)} \, \mathrm{d}s + \int_{\lfloor t_n \rfloor}^{t_n} e^{-a\lambda(\lceil s \rceil - s)} \, \mathrm{d}s}{1+R_{t_n}} + \frac{\int_0^{t_n} \mathrm{d}\lfloor s \rfloor}{1+R_{t_n}} \right) \\
&\quad = 2\frac{ah\tilde{c}}{\lambda} \left(\frac{1-a\lambda}{e^{a\lambda}} \frac{\lfloor t_n \rfloor \int_0^1 e^{a\lambda s} \, \mathrm{d}s}{1+R_{t_n}} + (1-a\lambda) \frac{\int_{\lfloor t_n \rfloor}^{t_n} e^{-a\lambda(\lceil s \rceil - s)} \, \mathrm{d}s}{1+R_{t_n}} + \frac{\lfloor t_n \rfloor}{1+R_{t_n}} \right) \\
&\quad = 2\frac{ah\tilde{c}}{\lambda} \underbrace{\frac{\lfloor t_n \rfloor}{1+t_n+\lfloor t_n \rfloor}}_{\to \frac{1}{2}} \left(\frac{(1-a\lambda)(e^{a\lambda}-1)}{a\lambda e^{a\lambda}} + \underbrace{\frac{\int_{\lfloor t_n \rfloor}^{t_n} e^{-a\lambda(\lceil s \rceil - s)} \, \mathrm{d}s}{\lfloor t_n \rfloor}}_{|\cdot| \le \mathcal{C} \lfloor t_n \rfloor^{-1} \to 0} (1-a\lambda) + 1 \right) \\
&\quad \longrightarrow \frac{ah\tilde{c}}{\lambda} \left(\frac{(1-a\lambda)(e^{a\lambda}-1)}{a\lambda e^{a\lambda}} + 1 \right) = -\frac{a}{\lambda(1-a\lambda)} h
\end{aligned}$$

with $(\star)$ as $\lceil s \rceil - \lfloor s \rfloor = 1$ for $s \notin \mathbb{N}$ and $s \in \mathbb{N}$ does not matter, because we integrate with respect to the Lebesgue measure. Furthermore $\lceil s \rceil - \lfloor s \rfloor = \lceil s \rceil - s = 0$ holds for $s \in \mathbb{N}$. In view of the last process, we have

$$a^2 \tilde{c}^2 h \frac{\int_0^t e^{-2a\lambda(\lceil s \rceil - s)} (1-a\lambda)^{2(\lceil s \rceil - \lfloor s \rfloor)} \, \mathrm{d}R_s}{1+R_t} \to a^3 \lambda \frac{e^{2a\lambda}(1+a^2\lambda^2) - (1-a\lambda)^2}{4(1-a\lambda)^2(1-a\lambda-e^{a\lambda})^2} h.$$

We add the results above and obtain the variance term, which appears in the limit distribution. It remains to verify the Lindeberg condition. Considering $|\Delta \tilde{M}_s^n|^2$ and $\mathbb{1}_{\{|\Delta \tilde{M}_s^n| > \delta\}}$, we obtain

$$\Delta \tilde{M}_s^n = a(1+R_{t_n})^{-\frac{1}{2}} \left(\frac{1}{a\lambda} + \tilde{c} \, e^{-a\lambda(\lceil st_n \rceil - st_n)} (1-a\lambda)^{\lceil st_n \rceil - \lfloor st_n \rfloor} \right) \Delta M_{st_n}$$

$$\leq \mathcal{C}\,(1 + R_{t_n})^{-\frac{1}{2}}\,\Delta M_{st_n}$$

with $\mathcal{C}$ independent of t. Thus it is sufficient, if M_t in connection with $(1+R_t)^{-\frac{1}{2}}$ fulfills the Lindeberg condition, which is ensured by assumption. □

Notations

$\{X + e_i\}_{i\in\{1,\dots,d\}}$	another notation for $X + \vec{1}$
$e_1, \dots, e_d$	the unit vectors of the Euclidean space $\mathbb{R}^d$
W_t	Brownian motion
$\mathcal{M}_{\text{loc}}$	the collection of local martingales
$\mathcal{M}^2_{\text{loc}}$	the collection of locally square integrable martingales
$\langle \cdot, \cdot \rangle$	the usual inner product of the Euclidean space $\mathbb{R}^d$
$[X]_t$	quadratic variation of the process X
$\\| \cdot \\|$	the usual norm of the Euclidean space $\mathbb{R}^d$
$[X, Y]_t$	covariation of the processes X and Y
$\lceil X \rceil_t$	predictable quadratic variation of the process X
$\lceil X, Y \rceil_t$	predictable covariation of the processes X and Y
M_t^c	purely continuous part of the local martingale M_t ($M_t^c \perp M_t^d$)
M_t^d	purely discontinuous part of the local martingale M_t ($M_t^c \perp M_t^d$)
R_t^c	continuous part of the process R_t
R_t^d	sum of the jumps of the process R_s up to the time t
ΔX_t	jump height of the process X at time t
$\{X \to\}$	set where X_∞ exists and is a finite random variable
$\{\cdot\}_i$	short notation for $\{f(Z_{s-} + c_s e_i) - f(Z_{s-} - c_s e_i)\}$
$\{\cdot\}_\cdot$	another notation for $\{\{\cdot\}_i\}_{i\in\{1,\dots,d\}}$
X_{t-}	left continuous version of the process X_t
$\nabla f(x)$	gradient of the function f at the point x
$\nabla_i f(x)$	i-th coordinate of $\nabla f(x)$
$\mathcal{E}(M)_t$	stochastic exponential of the process M
$\mathbb{1}$	indicator function
H_x	Hessian at the point x of a considered function f
λ_i	i-th eigenvalue of the Hessian

$\lambda_{\min}$	lowest eigenvalue of the Hessian
$\lambda_{\max}$	highest eigenvalue of the Hessian
$\mathrm{o}(\cdot)$	Landau symbol
$\mathcal{O}(\cdot)$	Landau symbol
$\mathbf{E}$	expectation
$\mathbf{P}$	probability measure
Z_t^D	process Z_t stopped at time D
I	identity matrix
T^T	transpose of the matrix T
δ_{ij}	Kronecker's delta (equals one if $i = j$ and zero otherwise)
$\lfloor \cdot \rfloor$	Gauss bracket
max	maximum
sup	supremum
inf	infimum
$\xrightarrow{\mathbf{P}}$	convergence in probability
$\xrightarrow{\mathcal{D}}$	convergence in distribution
$\simeq$	asymptotically equal
$b\mathcal{P}$	set of bounded predictable processes
$\mathcal{P}$	set of predictable processes
$\mathcal{V}$	set of all real-valued processes that are càdlàg, adapted, starting at zero and whose each path has a finite-variation on a finite interval
$\mathcal{V}^+$	set of all real-valued processes that are càdlàg, adapted, starting at zero and whose each path is non-decreasing

List of Figures

References

[1] Blum, J. R. (1954). Multidimensional stochastic approximation methods. *Ann. Math. Statist.* **25** 4 737-744.

[2] Crimaldi, I., Pratelli, L. (2005). Convergence results for multivariate martingales. *Stochastic Process. Appl.* **115** 4 571-577.

[3] Dippon, J. (2003). Accelerated randomized stochastic optimization. *Ann. Statist.* **31** 4 1260-1281.

[4] Dippon, J., Renz, J. (1997). Weighted means in stochastic approximation of minima. *SIAM J. Control Optim.* **35** 5 1811-1827.

[5] Dippon, J., Renz, J. (1996). Weighted means of processes in stochastic approximation. *Math. Meth. Statist.* **5** 1 32-60.

[6] Doléans-Dade, C. (1976). On the existence and unicity of solutions of stochastic differential equations. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **36** 93-101.

[7] Emery, M. (1978). Stabilité des solutions des équations différentielles stochastiques application aux intégrales multiplicatives stochastiques. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **41** 241-262.

[8] Fabian, V. (1968). On asymptotic normality in stochastic approximation. *Ann. Math. Statist.* **39** 4 1327-1332.

[9] Fabian, V. (1967). Stochastic approximation of minima with improved asymptotic speed. *Ann. Math. Statist.* **38** 1 191-200.

[10] Gal'chuk, L. I. (1978). Existence and uniqueness of a solution for stochastic equations with respect to semimartingales. *Theory Probab. Appl.* **23** 4 751-763.

[11] Györfi, L., Walk, H. (1996). On the averaged stochastic approximation for linear regression. *SIAM J. Control Optim.* **34** 1 31-61.

[12] Jacod, J., Shiryaev, A. N. (2002). Limit Theorems for Stochastic Processes. Springer. Berlin, Heidelberg, New York.

[13] Kiefer, J., Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* **23** 462-466.

[14] Kloeden, P. E., Platen E. (1992). Numerical Solution of Stochastic Differential Equations. Springer. Berlin, Heidelberg, New York.

[15] Kushner, H. J., Yin, G. G. (2003). Stochastic Approximation and recursive Algorithms and Applications. Springer. New York, Berlin, Heidelberg.

[16] Lazrieva, N. et al. (1997). The Robbins-Monro type stochastic differential equations. I. Convergence of solutions. *Stochastics Stochastics Rep.* **61** 1-2 67-87.

[17] Lazrieva, N. et al. (2003). The Robbins-Monro type stochastic differential equations. II. Asymptotic behaviour of solutions. *Stochastics Stochastics Rep.* **75** 3 153-180.

[18] Lazrieva, N. et al. (2008). Semimartingale stochastic approximation procedure and recursive estimation. *J. of Math. Sciences* **153** 3 211-261.

[19] Lipster, R. Sh., Shiryayev, A. N. (1986). Theory of Martingales. Kluwer Academic Publisher. Dordrecht, Boston, London.

[20] Ljung, L. et al. (1992). Stochastic Approximation and Optimization of Random Systems. Birkhäuser. Basel, Boston, Berlin.

[21] Mao, X. (1989). Exponential stability for nonlinear stochastic differential equations with respect to semimartingales. *Stochastics Stochastics Rep.* **28** 343-355.

[22] Melnikov, A. V., Valkeila E. (1997). Martingale models of stochastic approximation and their convergence. *Theory Probab. Appl.* **44** 2 333-360.

[23] Melnikov, A. V. et al. (1993). On a general class of stochastic approximation algorithms. *Front. Pure Appl. Probab.* **1** 183-196.

[24] Melnikov, A .V., Rodkina, A. E. (1993). Martingale approach to the procedures of stochastic approximation. *Front. Pure Appl. Probab.* **1** 165-182.

[25] Melnikov, A. V., Rodkina, A. E. (1992). Consistent statistical estimation in semimartingale models of stochastic approximation. *Ann. AcademiæScientiarum Fennicæ* **17** 85-91.

[26] Melnikov, A. V. (1989). Stochastic approximation procedures for semimartingales. *Statist. and control of random processes* 147-156.

[27] Melnikov, A. V. (1982). On strong solutions of stochastic equations with respect to semimartingales. *Lecture Notes in Control and Information Sciences* **43**. Springer. Berlin, Heidelberg.

[28] Melnikov, A. V. (1982). On properties of strong solutions of stochastic equations with respect to semimartingales. *Stochastics* **8** 2 103-119.

[29] Melnikov, A. V. (1979). On the theory of stochastic equations with respect to semimartingales. *Math. sbornik* **110** 3 414-427.

[30] Nevel'son, M. B., Has'minskiĭ, R. Z. (1973). Stochastic Approximation and Recursive Estimation. *Translations of Mathematical Monographs.*

[31] Polyak, B. T. (1990). New stochastic approximation type procedures. *Automat. Remote Control* **7** 98-107.

[32] Polyak, B. T., Juditsky, A. B. (1992). Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.* **30** 4 838-855.

[33] Protter, P. E. (2005). Stochastic Integration and Differential Equations. Springer. Berlin, Heidelberg.

[34] Protter, P. E. (1977). On the existence, uniqueness, convergence, and explosions of solutions of systems of stochastic integral equations. *Ann. Probab.* **5** 2 243-261.

[35] Robbins, H., Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22** 400-407.

[36] Sacks, J. (1957). Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.* **29** 2 373-405.

[37] Wang, J.-G. (1995). The asymptotic behaviour of locally square integrable martingales. *Ann. Probab.* **23** 2 552-585.

[38] Wang, J.-G. (1993). A law of the iterated logarithm for stochastic integrals. *Stochastic Process. Appl.* **47** 215-228.