Sharp nonlinear generalizations of stochastic Gronwall inequalities

Sarah Geiss∗†

December 8, 2021

Abstract

We provide nonlinear generalizations of a class of stochastic Gronwall inequalities that have been studied by von Renesse and Scheutzow (2010), Scheutzow (2013), Xie and Zhang (2020), Mehri and Scheutzow (2021) and Le and Ling (2021). This class of stochastic Gronwall inequalities is a useful tool for SDEs.

Our main focus are nonlinear generalizations of the Bihari-LaSalle type. Most of our estimates are sharp, in particular, we provide sharp constants for the stochastic Gronwall inequalities. The proofs are connected to the proof of a domination inequality by Lenglart (1977) and a proof by Pratelli (1976).

We apply our results to prove existence and uniqueness of global solutions of path-dependent SDEs driven by Lévy processes under a one-sided non-Lipschitz condition. Furthermore, we provide conditions for the existence of exponential moments of solutions of path-dependent SDEs driven by Brownian motion, which are similar to the conditions known for non-path-dependent SDEs.

Keywords: stochastic Gronwall inequality, stochastic Bihari-LaSalle inequality, Lenglart’s domination inequality, sharp constants, exponential moments of path-dependent SDEs, existence and uniqueness of path-dependent SDEs

MSC2020 subject classifications: 34K50, 60H10, 60G44, 60G51, 60J65

Contents

1 Introduction

2 Notation, assumptions and summary

2.1 Notation and constants

2.2 Assumptions

2.3 Summary of the results

∗Technische Universität Berlin, Germany. E-mail: geiss@math.tu-berlin.de
†The author was supported by an Elsa-Neumann-Stipendium des Landes Berlin.
1 Introduction

In this article we study nonlinear generalizations of two stochastic Gronwall inequalities. Furthermore, we sharpen some of the constants. We apply these results to study existence and uniqueness of global solutions to path-dependent SDEs. Furthermore, we provide a criterion for the finiteness of exponential moments of solutions to path-dependent SDEs, which is similar to that known for non-path-dependent SDEs.

The following stochastic Gronwall inequality with supremum is due to von Renesse and Scheutzow [33, Lemma 5.4] and was generalized by Mehri and Scheutzow [24, Theorem 2.1]: Let \( \{X_t, t \geq 0\} \) be a non-negative stochastic process, that satisfies

\[
X_t \leq \int_{(0,t]} X^*_s \, dA_s + M_t + H_t \quad \text{for all } t \geq 0,
\]

where \( X^*_s = \sup_{u \in [0,s]} X_u \) denotes the running supremum. Here, \( M \) is a càdlàg local martingale that starts in 0, and \( H \) and \( A \) are suitable non-decreasing stochastic processes.
Then, for all $T > 0$ and $p \in (0, 1)$ there exists an explicit upper bound for $\mathbb{E}[\sup_{t \leq T} X_t^p]$ which does not depend on the martingale $M$ (i.e. it only depends on $p$, $T$, $A$ and $H$).

There is also a stochastic Gronwall inequality without supremum which is closely connected to the previous inequality with supremum. This result is due to Scheutzow [30, Theorem 4] and was generalized by Xie and Zhang [36, Lemma 3.8] to càdlàg martingales: If we assume instead of (1) the slightly stronger assumption that

$$X_t \leq \int_{[0,t]} X_s - dA_s + M_t + H_t \quad \text{for all } t \geq 0,$$

(2)

sharper bounds can be obtained.

Both previously mentioned inequalities are useful tools for SDEs. The stochastic Gronwall inequality with supremum is applied to study SDEs with memory, see for example [2], [3], [4], [5], [24], [31] and [33]. The stochastic Gronwall inequality without supremum is applied to study various SDEs without memory, see e.g. [8], [11], [12], [15], [18], [23], [28], [29], [32], [35] and [37].

Also nonlinear extensions of the stochastic Gronwall inequalities have been studied. We refer to the following extensions as stochastic $L^\theta$ Gronwall inequalities. Assume

$$X_t \leq \left( \int_{[0,t]} (X_s)^\theta dA_s \right)^{1/\theta} + M_t + H_t \quad \text{for all } t \geq 0,$$

(3)

Then, for $\theta \in (0, 1)$, a continuous local martingale $M$ and otherwise similar assumptions as before, upper bounds on $\mathbb{E}[\sup_{t \leq T} X_t^p]$, $p \in (0, \theta)$ which do not depend on the martingale $M$ can be derived, see Makasu [22, Theorem 2.2]. If in addition $X$ is assumed to be non-decreasing, estimates of $\mathbb{E}[\sup_{t \leq T} X_t^p]$ can by obtained all $\theta \in (0, \infty)$ by Le and Ling [20, Lemma 3.8], where the range of $p$ depends on whether $A$ is deterministic or random.

A further nonlinear extension of the stochastic Gronwall inequalities has been studied by Mekki, Nieto and Ouahab [25, Theorem 2.4]: For continuous local martingales a stochastic Henry Gronwall’s inequality with upper bounds that do not depend on the local martingale $M$ can be proven.

Further extensions of the stochastic Gronwall inequalities have been studied, where the upper bounds depend on the quadratic variation of the martingale $M$, see e.g. Makasu [22], Makasu [21] and Mekki, Nieto and Ouahab [25]. In the present paper, we focus on bounds which do not depend on the local martingale $M$. Furthermore, Hudde, Hutzenthaler and Mazzonetto [13] have extended the stochastic Gronwall inequality without supremum to the setting of Itô processes which satisfy a suitable one-sided affine-linear growth condition.

In this paper, we study the following nonlinear generalization of the above mentioned stochastic Gronwall inequalities: We replace the assumptions (1) and (2) by

$$X_t \leq \int_{[0,t]} \eta(X^*_s) dA_s + M_t + H_t \quad \text{for all } t \geq 0,$$

(4)

and

$$X_t \leq \int_{[0,t]} \eta(X^-_s) dA_s + M_t + H_t \quad \text{for all } t \geq 0,$$

(5)
respectively, where \(\eta : [0, \infty) \to [0, \infty)\) is a suitable non-decreasing function.

For continuous martingales, (4) has been studied by von Renesse and Scheutzow [33] in the context of global solutions of stochastic functional differential equations, but no explicit upper bounds for \(E\) were derived.

The Bihari-LaSalle inequality provides an upper bound for \(X_t\) in the deterministic case (i.e. \(M \equiv 0\)), see [6]. In [6] and [16] the integrator is of the form \(A(t) = \int_0^t \varphi(s)ds\), but we need a version where the integrator \(A\) is only càdlàg. As this version is difficult to find in the literature, we provide a short proof in the appendix, which goes along the lines of the standard proof.

**Lemma 1.1** ((Deterministic) Bihari-LaSalle inequality). Let \(H > 0\) be a constant, \(x : [0, \infty) \to [0, \infty)\) a càdlàg function and \(A : [0, \infty) \to [0, \infty)\) a non-decreasing càdlàg function with \(A(0) = 0\). Let \(\eta : [0, \infty) \to [0, \infty)\) be a continuous non-decreasing function satisfying \(\eta(u) > 0\) for \(u > 0\). Define \(G(v) := \int_C^v \frac{du}{\eta(u)}\) for some constant \(C > 0\). Assume that the function \(x\) satisfies

\[
x(t) \leq \int_{[0,t]} \eta(x(s^-))dA(s) + H \quad \forall t \in [0,T]
\]

for some \(T > 0\). If \(G(H) + A(T) \in \text{domain}(G^{-1})\), then the following inequality holds true:

\[
x(t) \leq G^{-1}(G(H) + A(t)) \quad \forall t \in [0,T].
\]

Note that the upper bound on \(x\) does not depend on the choice of the constant \(C > 0\) in the definition of \(G\). We obtain the well-known Gronwall inequality by choosing in Lemma 1.1 \(\eta(u) \equiv u\) i.e. \(G(u) \equiv \log(u)\), which implies the upper bound

\[
x(t) \leq H \exp(A(t)).
\]

For (4) we prove upper bounds on \(\text{E}[\sup_{t\in[0,T]} X_t^p]\), \(p \in (0, 1)\), which depend on whether \(\eta\) is concave or convex. For concave \(\eta\) we get bounds of the type:

\[
\|X_T^p\|_p \leq G^{-1}(G(\hat{c}_p\|H\|_p) + \hat{c}_pA_T)
\]

where \(G\) is defined as in Lemma 1.1 and \(\hat{c}_p\) and \(\hat{c}_p\) are constants which only depend on \(p\). For convex \(\eta\) the above estimate does not hold true in general. Instead, the following type of estimate can be shown:

\[
\|G^{-1}(G(X_T^p) - \hat{c}_pA_T)\|_p \leq \hat{c}_p\|H_T\|_p.
\]

For (5) we obtain similar bounds with improved constants. A summary of the stochastic Bihari-LaSalle inequalities obtained can be found in Table 2, see Section 2.3.

We also study extensions of the previously mentioned stochastic \(L^\theta\) Gronwall inequalities. We derive upper bounds for \(\text{E}[\sup_{t\in[0,T]} X_t^p]\) when (1) is replaced by

\[
X_t \leq \left(\int_{[0,t]} (X_{s^-})^\theta dA_s\right)^{1/\theta} + M_t + H_t \quad \mathbb{P}\text{-a.s } \forall t \geq 0,
\]

for some \(\theta \in (0, \infty)\). This inequality is a weaker assumption than (3), but coincides with (3) in the special case when \(X\) is non-decreasing. A summary of the bounds obtained can be found in Table 3, see Section 2.3.
We prove most nonlinear generalizations with a unified approach: Our procedure can be seen as a generalization of the proof of Lenglart’s inequality \[17\] Théorème I, Corollaire II. Our proofs are also linked to the proof of an inequality by Pratelli \[27\] Proposition 1.2 and to the proofs of the stochastic Gronwall inequalities by Mehri and Scheutzow \[24\] Theorem 2.1 and Xie and Zhang \[36\] Lemma 3.8.

## 2 Notation, assumptions and summary

We assume that all processes are defined on an underlying filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})\) satisfying the usual conditions.

### 2.1 Notation and constants

For \(p \in (0, 1)\) define the following constants:

\[
\beta = (1 - p)^{-1}, \quad \alpha_1 = (1 - p)^{-1/p}, \quad \alpha_2 = p^{-1}.
\]

Let \(Y\) be a random variable. We use for \(p \in (0, 1]\) the notation (if well-defined)

\[
\mathbb{E}_{\mathcal{F}_0}[Y] := \mathbb{E}[Y \mid \mathcal{F}_0], \quad \|Y\|_p := \mathbb{E}[|Y|^p]^{1/p}, \quad \|Y\|_{p, \mathcal{F}_0} := \mathbb{E}[|Y|^p \mid \mathcal{F}_0]^{1/p}.
\]

Let \(X\) be a non-negative stochastic process with right-continuous paths. We use the following notation for the running supremum and its left limits:

\[
X^*_t := \sup_{0 \leq s \leq t} X_s \quad \forall t \geq 0, \quad X^*_{t-} := \lim_{s \uparrow t} X^*_s = \sup_{s < t} X_s \quad \forall t > 0.
\]

As usual, we set \(X^*_{0-} := X_0\). If \(X\) is càdlàg, then also \(X^*_{t-} = \sup_{s \leq t} X_s^*\) holds true. If \(X\) is only right-continuous, \(X^*_t\) can take values in \([0, +\infty]\).

Let \(\eta : [0, \infty) \to [0, \infty)\) be a non-decreasing function with \(\eta(u) > 0\) for \(u > 0\). As mentioned in the introduction, we define for some \(c > 0\)

\[
G(x) := \int_c^x \frac{du}{\eta(u)} \quad \forall x \in [0, \infty),
\]

noting that \(G(0) \in [-\infty, 0)\). Due to \(G\) being an increasing function, the inverse is well-defined on the range of \(G\). If \(G(0) = -\infty\), then we define the term \(G^{-1}(G(0) + a)\) for \(a \in (-\infty, \infty)\) by

\[
G^{-1}(G(0) + a) := \lim_{\varepsilon \to 0} G^{-1}(G(\varepsilon) + a) = 0.
\]

We also define for all \(x > 0\) and \(p \in (0, 1)\):

\[
\eta_p(x) := \frac{p}{1 - p} \eta(x^{1/p}) x^{1-1/p}, \quad \tilde{G}_p(x) := \int_c^x \frac{du}{\eta_p(u)}. \tag{9}
\]

The functions \(G\) and \(\tilde{G}_p\) are related as follows:

\[
\tilde{G}_p(x) = \int_c^x \frac{du}{\eta_p(u)} = \frac{1 - p}{p} \int_c^x \frac{du}{\eta(u^{1/p}) u^{1-1/p}} = (1 - p) \int_c^{x^{1/p}} \frac{du}{\eta(u)} \tag{10}
\]

In addition, a simple calculation gives \(\tilde{G}_p^{-1}(x) = (G^{-1}(\frac{1}{1-p})^p)\).
2.2 Assumptions

We study the following three cases:

**Definition 2.1 (Assumption $A_{\text{sup}}$).** Let

- $(X_t)_{t \geq 0}$ be an adapted non-negative right-continuous process,
- $(A_t)_{t \geq 0}$ be a predictable non-decreasing càdlàg process with $A_0 = 0$,
- $(H_t)_{t \geq 0}$ be an adapted non-negative non-decreasing càdlàg process,
- $(M_t)_{t \geq 0}$ be a càdlàg local martingale with $M_0 = 0$,
- $\eta : [0, \infty) \to [0, \infty)$ be a non-decreasing continuous function with $\eta(x) > 0$ for all $x > 0$.

We say the processes $X$, $A$, $H$, $M$ satisfy $A_{\text{sup}}$ if they satisfy the inequality below for all $t \geq 0$:

$$X_t \leq \int_{(0,t]} \eta(X_s^*)dA_s + M_t + H_t \quad \mathbb{P}\text{-a.s.} \quad (11)$$

The following assumption is slightly stronger:

**Definition 2.2 (Assumption $A_{\text{no sup}}$).** Under the same assumptions on the processes as in the previous definition, we say that the processes satisfy $A_{\text{no sup}}$ if in addition $X$ has left limits and the processes satisfy the following inequality for all $t \geq 0$:

$$X_t \leq \int_{(0,t]} \eta(X_s^-)dA_s + M_t + H_t \quad \mathbb{P}\text{-a.s.} \quad (12)$$

**Definition 2.3 (Assumption $A_{\text{sup, }\theta}$).** If in Definition 2.1 inequality (11) is replaced by

$$X_t \leq \left( \int_{(0,t]} (X_s^*)^{\theta}dA_s \right)^{1/\theta} + M_t + H_t \quad \mathbb{P}\text{-a.s.} \forall t \geq 0, \quad (13)$$

for some $\theta \in (0, \infty)$ we say the processes $X$, $A$, $H$, $M$ satisfy $A_{\text{sup, }\theta}$.

2.3 Summary of the results

The following three tables summarize some of the results in the literature and the results of this paper. The first table summarizes the stochastic Gronwall inequalities, the other two tables summarize the nonlinear generalizations. The constant $\alpha_1\alpha_2$ is the sharp constant from Lenglart’s inequality, $\alpha_2$ is the sharp constant from a monotone version of Lenglart’s inequality.
Assumption $A_{\sup}$, $\eta(x) \equiv x$ (Special case of $A_{\sup}$)

$A$ deterministic, $p \in (0, 1)$
- $H$ predictable or $\Delta M \geq 0$: $\|X_T^*\|_p \leq \alpha_1 \alpha_2 \|H_T\|_p e^{A_T}$
  
  $\|X_T^\ast\|_p \leq \alpha_1 \|H_T\|_p e^{A_T}$
  See Corollary 4.4

$E[H_T] < \infty$:
$\|X_T\|_p \leq \alpha_1 \|H_T\|_1 e^{A_T}$
  See Corollary 4.4

von Renesse and Scheutzow [33, Lemma 5.4], Mehri and Scheutzow [21, Theorem 2.1]:
- $H$ predictable: $\|X_T^\ast\|_p \leq c_1 \|H_T\|_p e^{c_2 A_T}$
  $c_1 = p^{-1} \alpha_1 \alpha_2$, $c_2 = p^{-1} c_p^2$

$\Delta M \geq 0$: $\|X_T^\ast\|_p \leq c_1 \|H_T\|_p e^{c_2 A_T}$
  $c_1 = (c_p + 1)/p$, $c_2 = p^{-1} (c_p + 1)/p$

$E[H_T] < \infty$: $\|X_T\|_p \leq c_1 \|H_T\|_1 e^{c_2 A_T}$
  $c_1 = \alpha_1 \alpha_2 p^{-1} / p$, $c_2 = p^{-1} c_p^2$

Corollary 4.2: Sharp constants for the three cases above
- $H$ predictable or $\Delta M \geq 0$:
  $c_1 = \alpha_1 \alpha_2$, $c_2 = \beta$

$E[H_T] < \infty$:
$\|X_T^\ast\|_p \leq c_1 \|H_T\|_1 e^{c_2 A_T}$
  $c_1 = \alpha_1 \alpha_2 p^{-1} / p$, $c_2 = p^{-1} c_p^2$

$A$ random, $0 < q < p < 1$
- $H$ predictable or $\Delta M \geq 0$:
  $\|X_T^\ast\|_q \leq \alpha_1 \alpha_2 \|H_T\|_p e^{A_T} \|e_{\beta A_T}\|_{qp/(p-q)}$
  See Corollary 4.4

$E[H_T] < \infty$:
$\|X_T^\ast\|_q \leq C \|H_T\|_1 e^{A_T} \|e_{\beta A_T}\|_{qp/(p-q)}$
  See Corollary 4.4

where $C$ is given by:

[Scheutzow [30, Theorem 4],
Xie and Zhang [36, Lemma 3.8]:
$C = (pq^{-1} (1 - p)^{-1} - 1)^{1/q}$

Slightly improved constant (Corollary 4.4):
$C = \alpha_1$

$E[H_T] < \infty$:
$\|X_T^\ast\|_q \leq C \|H_T\|_1 e^{\beta A_T} \|e_{\beta A_T}\|_{qp/(p-q)}$
  See Corollary 4.4

where $C$ is given by:

[Scheutzow [30, Theorem 4],
Xie and Zhang [36, Lemma 3.8]:
$C = (pq^{-1} (1 - p)^{-1} - 1)^{1/q}$

Slightly improved constant (Corollary 4.4):
$C = \alpha_1$

$\beta = (1 - p)^{-1}$, $\alpha_1 = (1 - p)^{-1/p}$, $\alpha_2 = p^{-1}$, $c_p = (\alpha_1 \alpha_2)^p$

Table 1: Summary of stochastic Gronwall inequalities (i.e. $\eta(x) \equiv x$)
| Assumption $A_{\alpha,sup}$ |
|----------------------------|
| (Special case of $A_{sup}$) |

$\sup_{x \in (0,1)} \frac{f(x)}{x} < \infty$, $A$ random

$\eta$ concave, $A$ deterministic, $p \in (0,1)$

- $H$ predictable or $\Delta M \geq 0$:
  $$||X_t^\eta||_p \leq \alpha_1 G^{-1}(G(\alpha_2||H_T||_p) + A_T)$$
  See Theorem 3.1

- $\mathbb{E}[H_T] < \infty$:
  $$||X_t^\eta||_p \leq \alpha_1 G^{-1}(||H_T||_1 + A_T)$$
  See Theorem 3.1

$\eta$ convex, $A$ continuous, $p \in (0,1)$

- $H$ predictable or $\Delta M \geq 0$:
  $$||\sup_{t \in (0,T]} G^{-1}(G(X_t)-A_t)||_p \leq \alpha_1 \alpha_2 ||H_T||_p$$
  See Theorem 3.6

- $\mathbb{E}[H_T] < \infty$:
  $$||\sup_{t \in (0,T]} G^{-1}(G(X_t)-A_t)||_p \leq \alpha_1 ||H_T||_1$$
  See Theorem 3.6

Constants:

$$\beta = (1-p)^{-1}$$
$$\alpha_1 = (1-p)^{-1/p}$$
$$\alpha_2 = p^{-1}$$

Table 2: Summary of stochastic Bihari-LaSalle inequalities

| Assumption $A_{\alpha,sup,\theta}$ and $X$ non-decreasing |
|-----------------------------------------------------------|

$\theta > 1$, $A$ random, $p \in (0,1)$

- $H$ predictable or $\Delta M \geq 0$:
  $$||e^{-\lambda T} X_t^\eta||_p \leq \beta^{1/p} \alpha_1 \alpha_2 ||H_T||_p$$
  See Theorem 3.13

- $\mathbb{E}[H_T] < \infty$ and $A$ continuous:
  Le and Ling [20] Lemma 3.8:
  $$||e^{-\lambda T} X_t^\eta||_1 \leq 2||H_T||_1$$
  See Theorem 3.13

$\theta < 1$, $A$ random, $p \in (0,1)$

- $H$ predictable or $\Delta M \geq 0$:
  $$||e^{-\lambda T} X_t^\eta||_p \leq \alpha_2 ||H_T||_p$$
  See Theorem 3.13

- $\mathbb{E}[H_T] < \infty$ and $A$ continuous:
  Le and Ling [20] Lemma 3.8:
  $$||e^{-\lambda T} X_t^\eta||_1 \leq ||H_T||_1$$
  See Theorem 3.13

Constants, notation:

$$\beta = (1-p)^{-1}$$
$$\alpha_1 = (1-p)^{-1/p}$$
$$\alpha_2 = p^{-1}$$

Table 3: Summary of stochastic $L^0$ Gronwall inequalities
3 Main results

3.1 Stochastic Bihari-LaSalle inequalities

3.1.1 Estimates for concave $\eta$

In the literature stochastic Gronwall inequalities have been studied, Theorem 3.1 provides a nonlinear generalization. The related results are mentioned in detail in Section 4. For Theorem 3.1 discussed in detail in Section 3.3.

Theorem 3.1 (A sharp stochastic Bihari-LaSalle inequality for concave $\eta$). Assume that $A$ is deterministic and $p \in (0, 1)$. For $\eta$ from $A_{\sup}$ or $A_{\nosup}$ we use the notation $\eta_p : (0, \infty) \mapsto [0, \infty)$, $x \mapsto \eta(x^{1/p})x^{1-1/p}$.

a) Assume Assumption $A_{\sup}$ (see Definition 2.1) and that $\eta_p$ is concave and non-decreasing. Then, the following assertions hold for all $t \geq 0$:

$$
\|X^*_t\|_{p,F_0} \leq \begin{cases} 
G^{-1}\left(G(\alpha_1\alpha_2\|H_t\|_{p,F_0}) + \beta A_t\right) & \text{if } \mathbb{E}[H_t^p] < \infty, H \text{ is predictable}, \\
G^{-1}\left(G(\alpha_1\alpha_2\|H_t\|_{p,F_0}) + \beta A_t\right) & \text{if } \mathbb{E}[H_t^p] < \infty, \Delta M \geq 0,
\end{cases}
$$

b) Assume Assumption $A_{\nosup}$ (see Definition 2.2). Then, the following assertions hold for all $t \geq 0$:

$$
\|X^*_t\|_{p,F_0} \leq \begin{cases} 
\alpha_1 G^{-1}\left(G(\alpha_2\|H_t\|_{p,F_0}) + A_t\right) & \text{if } \mathbb{E}[H_t^p] < \infty, H \text{ is predictable}, \\
\alpha_1 G^{-1}\left(G(\alpha_2\|H_t\|_{p,F_0}) + A_t\right) & \text{if } \mathbb{E}[H_t^p] < \infty, \Delta M \geq 0,
\end{cases}
$$

Remark 3.2 (Well-definedness of the upper bounds). a) In the deterministic Bihari-LaSalle inequality Lemma 1.1 the assumption $G(H) + A(T) \in \text{domain}(G^{-1})$ is needed. In the Theorem 3.1 we do not need a corresponding assumption, because this is automatically satisfied if $\eta$ or $\eta_p$ is concave: We have domain$(G^{-1}) = \text{range}(G) = (\lim_{\varepsilon \to 0} G(\varepsilon), \lim_{x \to \infty} G(x))$. Hence, it suffices to show $\lim_{x \to \infty} G(x) = \infty$. Concavity of $\eta$ implies that there exists some $K > 0$ such that $\eta(u) \leq Ku$ for all $u \in [c, \infty)$, and hence

$$
\lim_{x \to \infty} G(x) = \int_{\varepsilon}^{\infty} \frac{du}{\eta(u)} \geq \int_{\varepsilon}^{\infty} \frac{du}{Ku} = +\infty.
$$

Due to (10) a similar argument implies $\lim_{x \to \infty} G(x) = \infty$ if $\eta_p$ is concave.
b) Recall that if \( G(0) = -\infty \), we defined the term \( G^{-1}(G(0) + a) \) for \( a \in (-\infty, \infty) \) by
\[
G^{-1}(G(0) + a) := \lim_{\varepsilon \to 0} G^{-1}(G(\varepsilon) + a) = 0.
\]

**Remark 3.3** (Random \( A \)). Assume Assumption \( \mathcal{A}_{\text{sup}} \) and that \( A \) is increasing, continuous and not deterministic. Then, by a time shift, we can transform \( \mathcal{A}_{\text{sup}} \) to the case with a deterministic integrator. Using \( \int_{(0, A_t^{-1})} \eta(X^*_{s^{-}})dA_s = \int_{(0, t)} \eta(X^*_{A^{-1}})ds \), we have:
\[
X_{A_t^{-1}} \leq \int_{[0, t]} \eta(X^*_{A_s^{-}})ds + M_{A_t^{-1}} + H_{A_t^{-1}}. \quad \mathbb{P}\text{-a.s.}
\]
Setting for all \( t \geq 0 \)
\[
\tilde{X}_t \equiv X_{A_t^{-1}}, \quad \tilde{A}_t \equiv t, \quad \tilde{M}_t \equiv M_{A_t^{-1}}, \quad \tilde{H}_t \equiv H_{A_t^{-1}},
\]
we have that \( \tilde{X}, \tilde{A}, \tilde{M}, \tilde{H} \) with filtration \( (\mathcal{F}_{A_t^{-1}})_{t \geq 0} \) satisfy the assumption \( \mathcal{A}_{\text{sup}} \). In particular, **Theorem 3.1** implies in this case upper bounds on \( \mathbb{E}[\sup_{s \in [0, A_t^{-1}]} X^*_s] \).

**Remark 3.4** (On the relation between the concavity of \( \eta_p \) and \( \eta \)). Assume as in **Theorem 3.1** that \( \eta_p \) is concave and non-decreasing. Assume further, that \( \eta \) is twice differentiable. Then, concavity of \( \eta_p \) implies concavity of \( \eta \) which can be seen as follows. Fix some \( p \in (0, 1) \). We have for all \( y > 0 \):
\[
\eta(y) = \eta_p(y^p)y^{1-p} \quad \text{and} \quad \eta''(y) = p^2\eta''_p(y^p)y^{p-1} - p(1-p)y^{-p}(\eta_p(y^p) - \eta'_p(y^p)y^p).
\]
Due to concavity of \( \eta_p \) we have \( 0 \leq \eta_p(0) \leq \eta_p(z) + \eta'_p(z)(0 - z) \) and \( \eta''_p(z) \leq 0 \) for all \( z > 0 \), therefore \( \eta''(x) \leq 0 \) and hence \( \eta \) is concave.

**Remark 3.5** (Comparison of bounds in the cases \( \mathcal{A}_{\text{nosup}} \) and \( \mathcal{A}_{\text{sup}} \)). Recall that Assumption \( \mathcal{A}_{\text{nosup}} \) is a special case of Assumption \( \mathcal{A}_{\text{sup}} \). The bound given by **Theorem 3.1** b) is indeed better than the bound of **Theorem 3.1** a). This can be seen as follows: We show for all \( h > 0, x \geq 0 \)
\[
\alpha_1 G^{-1}(G(h) + x) \leq G^{-1}(G(\alpha_1 h) + \beta x)
\]
holds true. Fix some \( h > 0 \). Define for all \( x \geq 0 \):
\[
f(x) := \alpha_1^p G^{-1}(G(h) + x)^p = \alpha_1^p (\tilde{G}_p^{-1}(\tilde{G}_p(h^p) + (1-p)x)), \quad g(x) := G^{-1}(G(\alpha_1 h) + \beta x)^p = (\tilde{G}_p^{-1}(\tilde{G}_p(\alpha_1^p h^p) + (1-p)\beta x)).
\]
Here, we used \( (10) \). Note that \( f(0) = \alpha_1^p h^p = g(0) \). Furthermore, we have \( \alpha_1^p = \beta, \alpha_1^{-p} \in (0, 1) \) and
\[
f'(x) = (1-p)\alpha_1^p \frac{1}{\tilde{G}_p'(\alpha_1^{-p} f(x))} = (1-p)\alpha_1^p \eta_p(\alpha_1^{-p} f(x)) \quad \text{and} \quad g'(x) = (1-p)\beta \frac{1}{\tilde{G}_p'(g(x))} = (1-p)\alpha_1^p \eta_p(g(x)).
\]
The assumption that \( \eta_p \) is non-decreasing implies that \( f(x) \leq g(x) \) for all \( x \geq 0 \). Taking the \( p \)-th root implies the claim.
3.1.2 Estimates for convex $\eta$

Assume Assumption $\mathcal{A}_{\text{nosup}}$ and that $H$ is constant. If $\eta$ is convex, estimates of the type
\[
\|X_t^*\|_p \leq c_1 G^{-1}(c_2 G(c_3 H) + c_4 A_t) \tag{15}
\]
(where $c_1, c_2, c_3$ and $c_4$ are constants which depend only on $p \in (0, 1)$) are in general false: For processes $X$ which satisfy Assumption $\mathcal{A}_{\text{nosup}}$ for convex $\eta$, the quantity $\|X_t^*\|_p$ may explode at finite time and this type of behaviour cannot be captured by a bound of the type (15). This is discussed in more detail in the appendix, see [Counterexample 10.1]. In particular, estimates of the type (15) do not hold under the weaker Assumption $\mathcal{A}_{\text{sup}}$.

However, the estimate of the deterministic Bihari-LaSalle inequality (see e.g. [Lemma 1.1]) can be rearranged to
\[
G^{-1}(G(x(t)) - A(t)) \leq H.
\]
This rearranged inequality can be generalized to the stochastic case for convex $\eta$. The following theorem can be used to study the finiteness of exponential moments of path-dependent SDEs, see [Section 5]

**Theorem 3.6** (A sharp stochastic Bihari-LaSalle inequality for convex $\eta$). a) Assume Assumption $\mathcal{A}_{\text{nosup}}$. Furthermore, let $A$ be continuous, $\eta$ convex and $\eta(0) = 0$. Then, for $p \in (0, 1)$ the following estimates hold for all $T \geq 0$:

\[
\left\|G^{-1}(G(X_T^*) - A_T)\right\|_p \leq \left\|\sup_{t \in [0,T]} G^{-1}(G(X_t) - A_t)\right\|_p \leq \begin{cases} 
\alpha_1 \|H_T\|_{1,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1 \alpha_2 \|H_T\|_{p,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1 \alpha_2 \|H_T\|_{p,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } H \text{ is predictable.}
\end{cases}
\]

b) Assume Assumption $\mathcal{A}_{\text{sup}}$. Furthermore, let $A$ continuous, $\eta_p$ convex and $\eta_p(0) = 0$. Then, for $p \in (0, 1)$ the following estimates hold for all $T \geq 0$:

\[
\left\|G^{-1}(G(X_T^*) - \beta A_T)\right\|_{p,\mathcal{F}_0} \leq \begin{cases} 
\alpha_1 \|H_T\|_{1,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1 \alpha_2 \|H_T\|_{p,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1 \alpha_2 \|H_T\|_{p,\mathcal{F}_0} & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } H \text{ is predictable.}
\end{cases}
\]

The assumption that $A$ is continuous is not necessary when $\eta(x) \equiv x$. The assumption $\eta(0) = 0$ and $\eta(x) > 0$ for $x > 0$ can be replaced by $\eta(c_0) = 0$ for some $c_0 \geq 0$, if $X \geq c_0$ and $H \geq c_0$. In this case, we define $G$ using some $c > c_0$.

The constants $\alpha_1$ and $\alpha_1 \alpha_2$ are sharp, they are already sharp when $A \equiv 0$. The constant $\beta$ is sharp, it is already sharp when $A$ is deterministic, $H$ is deterministic and constant and $\eta(x) \equiv x$. Sharpness is discussed in [Section 3.3]

**Remark 3.7** (On the relation between the convexity of $\eta_p$ and $\eta$). By the same calculations as in the concave case, we have: Assume that $\eta$ is twice differentiable and that $\eta_p(0) = 0$. Then, the convexity of $\eta_p$ implies that $\eta$ is convex.

**Corollary 3.8.** We get in the case $\mathcal{A}_{\text{sup}}$ and $\mathbb{E}[H_T] < \infty$ of the previous theorem the following estimates.
a) For $\eta(x) = x$ for all $x \geq 0$, we have

$$\|e^{-\beta A T} X_T^*\|_{p, F_0} \leq \alpha_1 \|H_T\|_{1, F_0}.$$ 

b) For $\eta(x) = (x + 1) \log(x + 1)$ for all $x \geq 0$, we have

$$\|(X_T^* + 1)e^{-\beta A T} - 1\|_{p, F_0} \leq \alpha_1 \|H_T\|_{1, F_0}.$$ 

c) If $X \geq 1$, $H \geq 1$ and $\eta(x) = x \log(x)$, we have

$$\|e^{-\beta A T} (X_T^* + 1)\|_{p, F_0} \leq \alpha_1 \|H_T\|_{1, F_0}.$$ 

d) For $\eta(x) = (x + e) \log(x + e) \log(\log(x + e))$ for all $x \geq 0$, we have

$$\|e^{\log^{-\beta A T} (X_T^* + e)} - e\|_{p, F_0} \leq \alpha_1 \|H_T\|_{1, F_0}.$$ 

Proof of Corollary 3.8

a) For $\eta(x) = x$ and $c = 1$ we have $G(x) = \log(x)$ and $G^{-1}(x) = e^x$.

b) For $\eta(x) = (x + 1) \log(x + 1)$ and $c = e - 1$ we have

$$G(x) = \log(\log(x + 1)) and G^{-1}(x) = \exp(\exp(x)) - 1.$$ 

c) For $\eta(x) = x \log(x)$ and $c = e$ we have

$$G(x) = \log(\log(x)) and G^{-1}(x) = \exp(\exp(x)).$$ 

d) For $\eta(x) = (x + e) \log(x + e) \log(\log(x + e))$ and $c = e^e - e$ we have

$$G(x) = \log(\log(\log(x + e))) and G^{-1}(x) = \exp(\exp(\exp(x))) - e.$$ 

3.1.3 Estimates for general $\eta$

Now we study upper bounds of $X$ without imposing concavity or convexity assumptions on $\eta$. The estimate of the deterministic Bihari-LaSalle inequality (see e.g. Lemma 1.1) can be arranged as $G(x(t)) \leq G(H) + A(t)$. Theorem 3.9 is a stochastic version of this, which can be applied to study existence and uniqueness of solutions to path-dependent SDEs, see Section 5.

Theorem 3.9 (A stochastic Bihari-LaSalle inequality for $A_{sup}$). Assume Assumption $A_{sup}$. Furthermore, assume that $k_\eta := \sup_{x>0} \eta(x) < +\infty$. Then, we have the following estimate (where the terms take values in $[0, +\infty]$):

$$E_{F_0}[G(X_T^*)] \leq E_{F_0}[A_t] + G(E_{F_0}[H_t]) + k_\eta.$$ 

Assume that $\eta(x) \equiv x$ and assumption $A_{sup}$. Note that Theorem 3.1 implies upper bounds for $\|X_t\|_{p, F_0}$, Theorem 3.6 implies upper bounds for $\|X_t e^{-\beta A t}\|_{p, F_0}$ and Theorem 3.9 implies upper bounds for $E_{F_0}[\log(X_t^*)]$. In this sense Theorem 3.9 gives the weakest result.
3.2 Stochastic $L^\theta$ Gronwall inequalities

In this section we assume the assumption $A_{\sup, \theta}$. Hence, in particular we assume for all $t \geq 0$

$$X_t \leq \left( \int_{[0,t]} (X^*_s)^\theta dA_s \right)^{1/\theta} + M_t + H_t \quad \mathbb{P}\text{-a.s.}$$

**Remark 3.10** (Deterministic case). In the deterministic case, i.e. $M = 0$ this inequality has e.g. been studied by Willett and Wong [34, Theorem 1]. For a stochastic version with upper bounds of a similar type (depending on the quadratic variation of the martingale) has been studied by Makasu [22, Theorem 2.1]. We obtain upper bounds of a different structure.

**Remark 3.11** (Upper bounds on $\|X^*_t\|_{q,F_0}$). Estimates of the form

$$\|e^{-c_1 \cdot A_t} X^*_t\|_{p,F_0} \leq c_2 \tilde{H}, \quad \text{for some } p \in (0, 1),$$

for some constants $\tilde{H} \geq 0, c_1 > 0, c_2 > 0$, imply upper estimates on $\|X^*_t\|_{q,F_0}$ via Hölder’s inequality (using the exponents $p/q$ and $p/(p-q)$) for $q \in (0, p)$:

$$\|X^*_t\|_{q,F_0} \leq c_2 \tilde{H}\|e^{-c_1 \cdot A_t}\|_{qp/(p-q),F_0}.$$

Therefore, we only study upper bounds of $\|e^{-c_1 \cdot A_t} X^*_t\|_{p,F_0}$ in this section. In particular, these types of estimates to not require $\|e^{-c_1 \cdot A_t}\|_{qp/(p-q),F_0}$ to be finite.

We extend the results of Le and Ling [20, Lemma 3.8] in this section. Under similar assumptions this inequality has also been studied by Makasu [22, Theorem 2.2].

**Theorem 3.12** (Le and Ling: A monotone stochastic $L^\theta$ Gronwall inequality). Let $(X_t)_{t \geq 0}, (H_t)_{t \geq 0}$ be non-negative non-decreasing processes, let $(A_t)_{t \geq 0}$ be a continuous non-decreasing adapted process with $A_0 = 0$ and let $(M_t)_{t \geq 0}$ be a local martingale with $M_0 = 0$. Suppose that there exists a constant $\theta \in (0, \infty)$ such that with probability one,

$$X_t \leq \left( \int_{[0,t]} X^*_s dA_s \right)^{1/\theta} + M_t + H_t, \quad \forall t \geq 0.$$

Then, for any bounded stopping time $\tau$, we have:

$$\mathbb{E}[2^{-2^\theta A_{\tau}} X_{\tau}] \leq 2\mathbb{E}[H_{\tau}] \quad \text{when } \theta \geq 1,$$

$$\mathbb{E}[2^{-4A_{\tau}^{1/\theta}} X_{\tau}] \leq \mathbb{E}[H_{\tau}] \quad \text{when } \theta < 1.$$
a) If $A$ is continuous and $\theta > 1$, then for all $t \geq 0$:
\[
\| \exp\{-\beta^\theta A_t\} X_t^* \|_{p,F_0} \leq \begin{cases} 
\theta^{1/p} \alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } H \text{ is predictable.} \\
\theta^{1/p} \alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } \Delta M \geq 0, \\
\| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\end{cases}
\]

b) If $\theta \leq 1$, then for all $t \geq 0$:
\[
\| \exp\{-\beta A_t\} X_t^* \|_{p,F_0} \leq \begin{cases} 
\alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } H \text{ is predictable.} \\
\alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } \Delta M \geq 0, \\
\| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\end{cases}
\]

and
\[
\| \exp\{-A_t\} X_t^* \|_{1,F_0} \leq \theta \| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\]

c) If $X$ is non-decreasing and $\theta > 1$, then for all $t \geq 0$:
\[
\| \exp\{-A_t\} X_t^* \|_{p,F_0} \leq \begin{cases} 
\theta^{1/p} \alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } H \text{ is predictable.} \\
\theta^{1/p} \alpha_1 \alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } \Delta M \geq 0, \\
\| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\end{cases}
\]

d) If $X$ is non-decreasing and $\theta \leq 1$, then for all $t \geq 0$:
\[
\| \exp\{-A_t^{(\theta)}\} X_t^* \|_{p,F_0} \leq \begin{cases} 
\alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } H \text{ is predictable.} \\
\alpha_2 \| H_t \|_{p,F_0} & \text{if } \mathbb{E}[H_t^p] < \infty \text{ and } \Delta M \geq 0, \\
\| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\end{cases}
\]

and
\[
\| \exp\{-A_t^{(\theta)}\} X_t^* \|_{1,F_0} \leq \| H_t \|_{1,F_0} & \text{if } \mathbb{E}[H_t] < \infty.
\]

The constants $\alpha_1, \alpha_2$ and $\beta$ only depend on $p$, see (S).

Remark 3.14. If $A$ is continuous, then $A_t^{(\theta)} = A_t^{1/\theta}$ for $\theta < 1$. If $A$ is not continuous, we have e.g. by Remark 10.3 that $A_t^{(\theta)} \geq A_t^{1/\theta}$.

3.3 Sharpness of constants

Recall the following definition from (S) for $p \in (0,1)$:
\[
\beta = (1-p)^{-1}, \quad \alpha_1 = (1-p)^{-1/p}, \quad \alpha_2 = p^{-1}.
\]

The key idea of the proof of the following theorem is due to Michael Scheutzow.

Theorem 3.15 (Scheutzow: Sharpness of the constant $\beta$). Let $p \in (0,1)$ and assume that $\tilde{\alpha}$ and $\beta$ are positive constants (depending on $p$) such that for any non-negative adapted continuous process $(X_t)_{t \geq 0}$ which satisfies
\[
X_t \leq \int_0^t X_s^* ds + M_t + H, \quad \forall t \geq 0.
\]

for some continuous local martingale $(M_t)_{t \geq 0}$ and some constant $H > 0$, we have for all $t \geq 0$
\[
\| X_t^* \|_p \leq \tilde{\alpha} H \exp(\tilde{\beta} t).
\]

Then, $\tilde{\beta} \geq \beta$ holds true.
Corollary 3.16. The constant $\beta$ in Theorem 3.7 a), Theorem 3.6 b) and Theorem 3.13 (in the case $\theta = 1$) is sharp. It is already sharp when $\eta(x) \equiv x$, $A_t \equiv t$, $H$ is a constant and $M$ and $X$ are continuous processes.

The next theorem studies the sharpness of $\alpha_1$ and $\alpha_1\alpha_2$. The constant $\alpha_1\alpha_2$ is the constant which also appears in Lenglart’s domination inequality. In particular, Theorem 3.17 a) is closely connected to [9, Theorem 2.1]. Assertion b) of Theorem 3.17 is known in the literature, see for example [26, Theorem 7.6, p. 300].

Theorem 3.17 (Sharpness of the constants $\alpha_1$ and $\alpha_1\alpha_2$). Let $X$ be an adapted non-negative continuous process, $H$ an adapted non-negative non-decreasing continuous process, and $M$ a càdlàg local martingale with no negative jumps and $M_0 = 0$. Assume that for all $t \geq 0$ the inequality

$$X_t \leq H_t + M_t$$

holds true. Then:

a) For all $t \geq 0$ $\|X_t\|^p \leq \alpha_1\alpha_2\|H_t\|^p$ holds and the constant $\alpha_1\alpha_2$ is optimal.

b) For all $t \geq 0$ $\|X_t\|^p \leq \alpha_1\|H_t\|_1$ and the constant $\alpha_1$ is optimal. It is already optimal when $H$ is a constant and $M$ is continuous.

corollary 3.18. a) The constants $\alpha_1\alpha_2$ and $\alpha_1$ in Theorem 3.1 a) are sharp.

b) The constant $\alpha_1$ in Theorem 3.1 b) is sharp. If $\eta(x) \equiv x$, then $G(x) = \log(x)$ and $G^{-1}(x) = \exp(x)$. In particular, $\alpha_1\alpha_2$ also appears in the estimate of b) in this case and it is sharp.

c) The constants $\alpha_1\alpha_2$ and $\alpha_1$ in Theorem 3.6 are sharp.

d) If $\theta \leq 1$, then both constants $\alpha_1\alpha_2$ and $\alpha_1$ are sharp in Theorem 3.13. Both constants are in all above listed cases already sharp when $A \equiv 0$.

4 Stochastic Gronwall inequalities

In this section we summarize the results in the literature for the linear case $\eta(x) = x$ and compare them with the inequalities of this paper.

Von Renesse and Scheutzow [33, Lemma 5.4] developed a stochastic Gronwall inequality for continuous martingales to study stochastic functional differential equations. This result was further generalized by Mehri and Scheutzow [24, Theorem 2.1], who applied Lenglart’s domination inequality in the proof.

Theorem 4.1 (Mehri and Scheutzow: A stochastic Gronwall inequality for $A_{\sup}$). Assume Assumption $A_{\sup}$, that $A$ is deterministic and $\eta(x) \equiv x$. Then, the following estimates hold for $p \in (0, 1)$ and $T > 0$.

$$\|X_T\|^p_{p,F_0} \leq \begin{cases} p^{-1/p}c_p\|H_T\|^p_{p,F_0}e^{p^{-1}c_p A_T} & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } H \text{ is predictable,} \\ p^{-1/p}(c_p^p + 1)^{1/p}\|H_T\|^p_{1,F_0}e^{-1}(c_p^{p+1})^{1/p}A_T & \text{if } \mathbb{E}[H_T^p] < \infty \text{ and } \Delta M \geq 0, \\ p^{-1/p}c_p\|H_T\|^p_{1,F_0}e^{-1}c_pA_T & \text{if } \mathbb{E}[H_T] < \infty, \end{cases}$$

where $c_p = \alpha_1\alpha_2$. 

15
The following is a corollary of Theorem 3.13 (b), Theorem 3.17 and Theorem 3.15. It slightly sharpens the result above and extends it to predictable integrators $A$:

**Corollary 4.2** (A sharp stochastic Gronwall inequality for $\mathcal{A}_{\sup}$). Assume Assumption $\mathcal{A}_{\sup}$ (see Definition 2.1) and $p \in (0, 1)$. Then, the following estimates hold:

$$
\|e^{-\beta A_t}X^*_t\|_{p,F_0} \leq \begin{cases} 
\alpha_1\|H_t\|_{1,F_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1\alpha_2\|H_t\|_{p,F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1\alpha_1\|H_t\|_{p,F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } H \text{ predictable}.
\end{cases}
$$

If $\|e^{\beta A_t}\|_{q/(p-q),F_0}$ is integrable, we have for $0 < q < p < 1$

$$
\|X^*_t\|_q \leq \begin{cases} 
\alpha_1\|H_t\|_{1,F_0}e^{\beta A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1\alpha_2\|H_t\|_{p,F_0}e^{\beta A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1\alpha_1\|H_t\|_{p,F_0}e^{\beta A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } H \text{ predictable}.
\end{cases}
$$

The constants $\alpha_1, \alpha_2$ and $\beta$ only depend on $p$ and they are sharp, see (8).

A stochastic Gronwall lemma (in setting nearly identical to $\mathcal{A}_{\nosup}$ with $\eta(x) \equiv x$) was proven for continuous martingales $M$ by Scheutzow [30, Theorem 4]. This result was extended by Xie and Zhang [36, Lemma 3.8] to càdlàg martingales:

**Theorem 4.3** (Xie and Zhang: A stochastic Gronwall inequality for $\mathcal{A}_{\nosup}$). Assume Assumption $\mathcal{A}_{\nosup}$. Furthermore, assume that $\eta(x) \equiv x$ and $A$ is continuous. Then, for any $0 < q < \hat{p} < 1$ and $t \geq 0$, we have:

$$
\|X^*_t\|_q \leq \left(\frac{\hat{p}}{\hat{p} - q}\right)^{1/q} \|H_t\|_1 e^{\tilde{A}_t\|_{\hat{p}/(1-\hat{p})}}.
$$

The following corollary of Theorem 3.6 slightly extends and marginally sharpens Theorem 4.3.

**Corollary 4.4** (A sharp stochastic Gronwall inequality for $\mathcal{A}_{\nosup}$). Assume Assumption $\mathcal{A}_{\nosup}$ (see Definition 2.2) and $p \in (0, 1)$. Then, the following estimates hold:

$$
\|e^{-A_t}X^*_t\|_{p,F_0} \leq \begin{cases} 
\alpha_1\|H_t\|_{1,F_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1\alpha_2\|H_t\|_{p,F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1\alpha_1\|H_t\|_{p,F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } H \text{ predictable}.
\end{cases}
$$

If $\|e^{\beta A_t}\|_{q/(p-q),F_0}$ is integrable, we have for $0 < q < p < 1$

$$
\|X^*_t\|_q \leq \begin{cases} 
\alpha_1\|H_t\|_{1,F_0}e^{A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H_T] < \infty, \\
\alpha_1\alpha_2\|H_t\|_{p,F_0}e^{A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } \Delta M \geq 0, \\
\alpha_1\alpha_1\|H_t\|_{p,F_0}e^{A_t}\|_{q/(p-q),F_0} & \text{if } \mathbb{E}[H^p_T] < \infty \text{ and } H \text{ predictable}.
\end{cases}
$$

The constants $\alpha_1, \alpha_2$ and $\beta$ only depend on $p$, see (8).

**Remark 4.5** (Comparison of constants). Choose any $0 < q < p < 1$ and set $\hat{p} := qp(p + q - q)^{-1}$. Then, $q < \hat{p} < 1$ holds and Theorem 4.3 implies:

$$
\|X^*_t\|_q \leq \left(\frac{1}{q} \frac{1}{1 - p}\right)^{1/q} \|H_t\|_1 e^{A_t}\|_{q/(p-q)}.
$$
Noting that, due to $0 < q < p < 1$ we have
\[
\left( \frac{p}{q} \frac{1}{1-p} \right)^{1/q} \geq \left( \frac{1}{1-p} \right)^{1/q} \geq \left( \frac{1}{1-p} \right)^{1/p},
\]
the constant in Corollary 4.4 is slightly sharper than that in Theorem 4.3. For deterministic $A$ (i.e. $\hat{p} \to 1$ or $p \to q$), both theorems yield the same constant $(1 - q)^{-1/q}$ which is optimal in this case.

5 Applications: A Lévy driven path-dependent SDE

The stochastic Bihari-LaSalle inequalities are, like the stochastic Gronwall inequalities, useful tools to study SDEs (see also the references of stochastic Gronwall applications listed in the introduction). We demonstrate this for the following path-dependent SDE driven by a Lévy process with bounded jumps, more general SDEs could be studied (see e.g. [24]).

Assume an underlying filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ satisfying the usual conditions. Let $L$ be an $\mathbb{R}^d$-valued Lévy process with bounded jumps and its Lévy-Itô decomposition given by
\[
L_t = b t + \sigma B_t + \int_{|\xi| \leq c} x \tilde{N}(t, d\xi) \quad \forall t \geq 0
\]
where $b \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d \times m}$, $B$ a Brownian motion, $N$ an independent Poisson random measure on $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ with Lévy measure $\nu$ and $c > 0$ some constant. We denote by $\tilde{N}$ the corresponding compensated Poisson random measure. For details, see for example [11] p. 126. We denote by $\| \cdot \|_F$ the Frobenius norm on $\mathbb{R}^{d \times m}$. Denote by $U := \{ \xi \in \mathbb{R}^d \mid 0 < |\xi| \leq c \}$ and set $\mathcal{U} := B(U)$, where $B(U)$ denotes the Borel $\sigma$-algebra on $U$.

We study the path-dependent SDE with random coefficients driven by $L$:
\begin{equation}
\begin{aligned}
&dX_t = f(t, X_t)dt + \int_{|\xi| \leq c} g(t, X_t, \xi) \tilde{N}(dt, d\xi) + h(t, X_t)dB_t \\
&X_t = z_t, t \in [-\tau, 0],
\end{aligned}
\end{equation}
where $\tau > 0$ is some constant and the initial condition $(z_t)_{t \in [-\tau, 0]}$ has càdlàg paths and is $\mathcal{F}_0$ measurable. Denote by $\mathcal{P}$ the predictable $\sigma$-algebra on $[0, \infty) \times \Omega$. We endow the space $C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d)$ with the topology generated by the uniform convergence on compact sets and denote by $\mathcal{B}(C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d))$ the ball Borel $\sigma$-algebra on $C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d)$. Assume that the coefficients
\[
f : ([0, \infty) \times \Omega \times C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d), \mathcal{P} \otimes \overline{\mathcal{B}(C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d)))} \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)),
\]
\[
g : ([0, \infty) \times \Omega \times C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d) \times U, \mathcal{P} \otimes \overline{\mathcal{B}(C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d))) \otimes \mathcal{U} \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)),
\]
\[
h : ([0, \infty) \times \Omega \times C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d), \mathcal{P} \otimes \overline{\mathcal{B}(C_\text{cadlag}([-\tau, \infty); \mathbb{R}^d))) \to (\mathbb{R}^{d \times m}, \mathcal{B}(\mathbb{R}^{d \times m}))
\]
are measurable mappings. For every $t \in [0, \infty)$, $\omega \in \Omega$, $\xi \in U$ assume that $f(t, \omega, x)$, $g(t, \omega, x, \xi)$ and $h(t, \omega, x)$ only depend on the path segment $x(s), s \in [-\tau, t)$. We denote by $f(t, x)$, $g(t, x, \xi)$ and $h(t, x)$ the corresponding random variables.

We study a less general SDE than [24 Theorem 3.3], but we have weaker assumptions on the coefficients, as we allow in (C1) and (C2) nonlinear upper bounds $\eta_1, \eta_2$. 17
Hypothesis 5.1. There exist non-negative functions $K \in L^1_{\text{loc}}([0, \infty), dt)$ and, for all $R > 0$, $K_R \in L^1_{\text{loc}}([0, \infty), dt)$ such that for all $x, y \in \text{Càdlàg}([-\tau, \infty); \mathbb{R}^d)$:

(C1) for $\sup_{s \in [-\tau, t]} |x(s)|$, $\sup_{s \in [-\tau, t]} |y(s)| \leq R$

$$2\langle x(t^-) - y(t^-), f(t, \omega, x) - f(t, \omega, y) \rangle + \int_U |g(t, \omega, x, \xi) - g(t, \omega, y, \xi)|^2 \nu(d\xi)$$

$$+ \|h(t, \omega, x) - h(t, \omega, y)\|_F^2 \leq L_R(t) \eta_1(\sup_{s \in [-\tau, t]} |x(s) - y(s)|^2),$$

(C2) $$2\langle x(t^-), f(t, \omega, x) \rangle + \int_U |g(t, \omega, x, \xi)|^2 \nu(d\xi) + \|h(t, \omega, x)\|_F^2 \leq K(t) \eta_2(1 + \sup_{s \in [-\tau, t]} |x(s)|^2),$$

(C3) $x \mapsto f(t, \omega, x)$ as a function from $\text{Càdlàg}([-\tau, \infty); \mathbb{R}^d)$ to $\mathbb{R}^d$ is continuous,

(C4) for $\sup_{s \in [-\tau, t]} |x(s)| \leq R$

$$|f(t, \omega, x)| + \int_U |g(t, \omega, x, \xi)|^2 \nu(d\xi) + \|h(t, \omega, x)\|_F^2 \leq K_R(t),$$

(C5) $\mathbb{E} [\sup_{s \in [-\tau, 0]} |z(s)|^2] < \infty,$

where $\eta_1, \eta_2 : [0, \infty) \to [0, \infty)$ are non-decreasing continuous functions with $\eta_i(x) > 0$ for all $x > 0$, $i = 1, 2$. Define

$$G_i(x) := \int_1^x \frac{du}{\eta_i(u)}, \quad \forall x > 0, i = 1, 2.$$

Assume:

$$k_{\eta_i} := \sup_{x > 0} \frac{x}{\eta_i(x)} < \infty \quad \text{for } i = 1, 2,$$

$$\int_0^\varepsilon \frac{du}{\eta_1(u)} = +\infty \quad \forall \varepsilon > 0 \quad \text{i.e. } \lim_{\varepsilon \to 0} G_1(\varepsilon) = -\infty,$$

$$\int_1^\infty \frac{du}{\eta_2(u)} = +\infty \quad \text{i.e. } \lim_{x \to \infty} G_2(x) = +\infty.$$

Remark 5.2. The assumption $\lim_{\varepsilon \to 0} G_1(\varepsilon) = -\infty$ implies that $\lim_{\varepsilon \to 0} \eta_1(\varepsilon) = 0$. Hence, (C1) and (C3) imply that for any fixed $\omega \in \Omega$ and $t \geq 0$ we have that the mappings

$$\text{Càdlàg}([-\tau, \infty); \mathbb{R}^d) \to L^2(U, \mathcal{U}, \nu), \quad x \mapsto g(t, \omega, x, \xi),$$

$$\text{Càdlàg}([-\tau, \infty); \mathbb{R}^d) \to \mathbb{R}^{d \times m}, \quad x \mapsto h(t, \omega, x)$$

are continuous.
5.1 Existence and uniqueness of strong solutions

Von Renesse and Scheutzow [33] studied existence and uniqueness of local and global solutions to stochastic functional differential equations driven by Brownian motion by using stochastic Gronwall inequalities and a nonlinear generalisation of a stochastic Gronwall lemma. In particular, they proved the existence of global solutions under a condition similar to (C2) of Hypothesis 5.1. Mehri and Scheutzow [24] generalized the stochastic Gronwall inequality of [33] to càdlàg martingales and studied a path-dependent SDE with càdlàg paths.

The following theorem is a nonlinear extension of [24, Theorem 3.3], our assumptions (C1) and (C2) of Hypothesis 5.1 include a nonlinear right-hand side. The proof is similar except that we use a stochastic Bihari-LaSalle inequality (Theorem 3.9) instead of a stochastic Gronwall inequalities and a nonlinear generalisation of a stochastic Gronwall inequality of [33] to càdlàg martingales and we also have a nonlinear monotonicity condition.

**Corollary 5.3** (Existence and uniqueness of global of solutions). Assume Hypothesis 5.1 holds. Then the SDE (16) has a unique strong global solution.

**Proof.** Proof of uniqueness of global solutions: Let $X$ and $Y$ be two global solutions of the SDE (16) with the same initial condition $(z_{t})_{t \in [-\tau, 0]}$. Define

$$
\tau(R) := \inf \{ t \geq 0 \mid |X_{t}| > R \text{ or } |Y_{t}| > R \}
$$

where we set $\inf \emptyset := +\infty$. Then, by Itô’s formula we have:

$$
|X_{t \wedge \tau(R)} - Y_{t \wedge \tau(R)}|^{2} = \int_{0}^{t \wedge \tau(R)} 2\langle X_{s^-} - Y_{s^-}, f(s, X) - f(s, Y) \rangle ds
$$

$$
+ \int_{0}^{t \wedge \tau(R)} \| h(s, X) - h(s, Y) \|_{F}^{2} ds
$$

$$
+ \int_{(0, t \wedge \tau(R))} \int_{U} |g(s, X, \xi) - g(s, Y, \xi)|^{2} \nu(\xi) ds + M_{t \wedge \tau(R)} \tag{17}
$$

where (C1) includes a nonlinear monotonicity condition.

$$
L_{R}(s) \eta_{1} \left( \sup_{u \in [0, s]} |X_{u} - Y_{u}|^{2} \right) ds + M_{t \wedge \tau(R)}
$$

We add a parameter $\varepsilon > 0$ to (17) and weaken it slightly:

$$
Z_{t} := |X_{t \wedge \tau(R)} - Y_{t \wedge \tau(R)}|^{2} + \varepsilon \leq \int_{0}^{t} \eta_{1}(Z_{s}^{*}) L_{R}(s) ds + M_{t \wedge \tau(R)} + \varepsilon.
$$
Applying [Theorem 3.9] to the process \((Z_t)_{t \geq 0}\) implies for any \(T > 0\):

\[
\mathbb{E}[G_1(\sup_{t \in [0,T]} |X_{t\wedge \tau(R)} - Y_{t\wedge \tau(R)}|^2 + \varepsilon)] \leq G_1(\varepsilon) + \mathbb{E}[\int_0^T L_R(s)ds] + k_{\eta,1}.
\]

Set \(p_\delta := \mathbb{P}[\sup_{t \in [0,T]} |X_{t\wedge \tau(R)} - Y_{t\wedge \tau(R)}|^2 > \delta]\). Due to \(x \rightarrow G_1(x)\) being increasing, we have for any \(\delta > 0\) and \(\varepsilon > 0\) that

\[
p_\delta G_1(\delta + \varepsilon) + (1 - p_\delta)G_1(\varepsilon) \leq \mathbb{E}[G_1(\sup_{t \in [0,T]} |X_{t\wedge \tau(R)} - Y_{t\wedge \tau(R)}|^2 + \varepsilon)] \leq G_1(\varepsilon) + \mathbb{E}[\int_0^T L_R(s)ds] + k_{\eta,1},
\]

which implies by rearranging (noting that \(G_1(\delta + \varepsilon) - G_1(\varepsilon) > 0\) and \(\lim_{\varepsilon \searrow 0} G_1(\varepsilon) = -\infty\):

\[
p_\delta \leq \frac{\mathbb{E}[\int_0^T L_R(s)ds] + k_{\eta,1}}{G_1(\delta + \varepsilon) - G_1(\varepsilon)} \rightarrow 0 \text{ for } \varepsilon \searrow 0.
\]

Hence, \(\mathbb{P}[\sup_{t \in [0,T]} |X_{t\wedge \tau(R)} - Y_{t\wedge \tau(R)}|^2 = 0] = 1\) for an \(T > 0, R > 0\), which implies the claim.

**Proof of the existence of strong global solutions:** We show the existence of a solution using the Euler method, see e.g. [19], [33], [24]. We define for \(n \in \mathbb{N}\) the Euler approximates as

\[
X_t^{(n)} := z_t, \quad t \in [-\tau, 0]
\]

\[
X_t^{(n)} := X_{\frac{k}{n}, t}^{(n)} + \int_{\left(\frac{k}{n}, t\right]} f(s, X_{\wedge k/n}^{(n)}) ds + \int_{\left(\frac{k}{n}, t\right]} \int_U g(s, X_{\wedge k/n}^{(n)}, \xi) \tilde{N}(ds, d\xi) + \int_{\left(\frac{k}{n}, t\right]} h(s, X_{\wedge k/n}^{(n)}) dB_s \quad \text{for } t \in \left(\frac{k}{n}, \frac{k+1}{n}\right], k \geq 0
\]

(18)

The Euler approximate \(X^{(n)}\) has càdlàg paths and is adapted. In particular, the stochastic integrals are well-defined. Define \(k(n, t) := \frac{k}{n}\) for \(t \in \left(\frac{k}{n}, \frac{k+1}{n}\right], k \geq 0\) and \(k(n, t) := t\) for \(t \in [-\tau, 0]\). Note that (18) can be rewritten as

\[
X_t^{(n)} = z_0 + \int_0^t f(s, X_{\wedge k(n,s)}^{(n)}) ds + \int_{[0,t]} \int_U g(s, X_{\wedge k(n,s),\xi}^{(n)}) \tilde{N}(ds, d\xi) + \int_0^t h(s, X_{\wedge k(n,s)}^{(n)}) dB_s
\]

(19)

Define the remainder

\[
p_t^{(n)} := X_{k(n,t)}^{(n)} - X_t^{(n)}, \quad t \in [-\tau, \infty).
\]

The stochastic process \(p^{(n)}\) is adapted and \(p^{(n)}(\frac{k}{n}+) = 0\) for every \(k \in \mathbb{N}_0\). Note that \(t \mapsto k(n, t)\) is càglàg and \(X_t\) is càdlàg. Therefore, \(p_t^{(n)}\) is neither càglàg nor càdlàg. Fix \(T \geq 0\) and define for \(R > 0\) the stopping times

\[
\tau_R^{(n)} := (\inf \{t \geq 0 \mid |X_t^{(n)}| > \frac{R}{3}\} \wedge T) \mathbb{1}_{\{R > \sup_{s \in [-\tau, 0]} |z_s|\}}.
\]

20
Then
\[ |p_t^{(n)}| \leq \frac{2R}{3}, \quad |X_t^{(n)}| \leq \frac{R}{3}, \quad t \in (0, \tau_R^{(n)}). \] (20)

On \{R > 3 \sup_{s \in [-\tau, 0]} |z_s| \} the inequalities extend to \([-\tau, \tau_R^{(n)}]\) and we have \(\tau_R^{(n)} > 0\) (due to the right-continuity of \(X^{(n)}\)).

As \(T\) was arbitrary, it suffices to prove the existence of a strong solution on \([0, T]\). The existence proof is done in the following steps:

a) For every \(t \geq 0\), \(1_{(0, \tau_R^{(n)}]}(t) \sup_{u \in (k(n), t]} |p_u^{(n)}| \to 0\) in probability as \(n \to \infty\).

b) \(E[\varphi(1 + \sup_{u \in [0, T]} |X_u^{(n)}| | 2)] \leq C(T, R, n)\) for some \(C(T, R, n)\) satisfying
\[ \lim_{n \to \infty} C(T, R, n) = \tilde{C}(T) \] for all \(R > 0\).

c) \(\lim_{R \to \infty} \limsup_{n \to \infty} \mathbb{P}[\tau_R^{(n)} < T] = 0\).

d) \(\lim_{n, m \to \infty} \mathbb{P}\left[ \sup_{t \in [0, T]} |X_t^{(n)} - X_t^{(m)}| > \varepsilon\right] = 0\) for all \(\varepsilon > 0\).

e) There exists a càdlàg adapted process \((X_t)_{t \geq 0}\) such that
\[ \lim_{n \to \infty} \mathbb{P}\left[ \sup_{t \in [0, T]} |X_t^{(n)} - X_t| > \varepsilon\right] = 0\] for all \(\varepsilon > 0\) and \(X\) is a strong solution of equation (16) on \([0, T]\).

We only prove the steps [b] [c] [d] The remaining steps are proven as in [24] Theorem 3.3.

**Proof of [b]** Using Itô’s formula, we have for \(t \in [0, T]\):
\[
|X_t^{(n)}|^2 = |z_0|^2 + \int_0^t 2\langle X_{s^{-}}^{(n)}, f(s, X_{\wedge k(n,s)}) \rangle ds + \int_0^t \int_U |g(s, X_{\wedge k(n,s)}, \xi)|^2 \nu(d\xi) ds
+ \int_0^t ||h(s, X_{\wedge k(n,s)})||^2_T ds + M_t^{(n)},
\]
where \(M^{(n)}\) is a local martingale defined by
\[ M_t^{(n)} := \int_{[0, t]} \int_U 2\langle X_{s^{-}}^{(n)}, g(s, X_{\wedge k(n,s)}, \xi) \rangle d\tilde{N}(ds, d\xi)
+ \int_{[0, t]} \int_U |g(s, X_{\wedge k(n,s)}, \xi)|^2 \tilde{N}(ds, d\xi)
+ \int_0^t 2\langle X_{s^{-}}^{(n)}, h(s, X_{\wedge k(n,s)}) dB_s \rangle.\]

Using (C2) and (C4), we have (using \(\lim_{r \nearrow s} X_{r \wedge k(n,s)}^{(n)} = X_{k(n,s)}^{(n)} = \lim_{r \nearrow s} X_{k(n,r)}^{(n)}\)):
\[
|X_{t \wedge \tau_R^{(n)}}^{(n)}|^2 \leq |z_0|^2 + \int_0^{t \wedge \tau_R^{(n)}} 2\langle X_{s^{-}}^{(n)} - \lim_{r \nearrow s} X_{r \wedge k(n,s)}^{(n)}, f(s, X_{\wedge k(n,s)}) \rangle ds
+ \int_0^t K(s) \eta_2 (1 + \sup_{u \in [-\tau, s]} |X_{u \wedge \tau_R^{(n)}}^{(n)}|) ds + M_t^{(n)}
\leq |z_0|^2 + \int_0^{t \wedge \tau_R^{(n)}} 2|p_s^{(n)}| \tilde{K}_R(s) ds
+ \int_0^t K(s) \eta_2 (1 + \sup_{u \in [-\tau, 0]} |z_u|^2 + \sup_{u \in [-\tau, s]} |X_{u \wedge \tau_R^{(n)}}^{(n)}|^2) ds + M_t^{(n)}
\]
Set 

\[ H_t^{n,R} := 1 + \sup_{u \in [-\tau,0]} |z_u|^2 + |z_0|^2 + \int_0^{t \wedge \tau_R(n)} 2|p_{s}^{(n)}|K_R(s)\,ds, \]

and note that by dominated convergence, (20), (a) and (C5):

\[ \limsup_{n \to \infty} \mathbb{E}[H_t^{n,R}] \leq 1 + 2\mathbb{E}[\sup_{u \in [-\tau,0]} |z_u|^2] < \infty. \]  

(21)

The application of Theorem 3.9 to \( Y_t := 1 + \sup_{u \in [-\tau,0]} |z_u|^2 + |X_{t \wedge \tau_R(n)}^{(n)}|^2 \), which satisfies \( Y_t \leq \int_0^t K(s)\eta_2(Y_s^*)\,ds + M_t^{(n)} + H_t^{n,R} \), implies:

\[ \mathbb{E}[G_2(1 + \sup_{t \in [0,T]} |X_{t \wedge \tau_R(n)}^{(n)}|^2)] \leq \mathbb{E}[G_2(Y_T^*)] \leq G_2(\mathbb{E}[H_T^{n,R}]) + \mathbb{E}[\int_0^T K(t)\,dt] + k_{\eta_2}. \]  

(22)

Combining (21) with (22) implies the assertion.

**Proof of (c):** We have due to \( G_2(x) \geq 0 \) for \( x \geq 1 \) and \( \lim_{x \to \infty} G_2(x) = +\infty \):

\[ \limsup_{n \to \infty} \limsup_{R \to \infty} \mathbb{P}[\tau_R^{(n)} < T] \leq \limsup_{n \to \infty} \limsup_{R \to \infty} \mathbb{P}[\sup_{t \in [0,T]} |X^{(n)}_t| \geq R/3] + \lim_{R \to \infty} \mathbb{P}[\sup_{s \in [-\tau,0]} |z_s| \geq R/3] \]

\[ \leq \limsup_{n \to \infty} \sup_{R \to \infty} \mathbb{E}[G_2(1 + \sup_{t \in [0,T]} |X^{(n)}_t|^2)] \]

\[ = \limsup_{R \to \infty} \frac{G_2(1 + (R/3)^2)}{G_2(1 + (R/3)^2)} \]

\[ b \]

\[ \leq \limsup_{R \to \infty} \frac{C(T, R, n)}{G_2(1 + (R/3)^2)} \]

\[ b \]

\[ = 0. \]

**Proof of (d):** Set \( \tau^{n,m}_{R} := \tau^{(n)}_{R} \land \tau^{(m)}_{R} \). Itô’s formula gives for \( t \in [0,T] \):

\[ |X^{(n)}_{t \wedge \tau^{n,m}_{R}} - X^{(m)}_{t \wedge \tau^{n,m}_{R}}|^2 \leq M_{t \wedge \tau^{n,m}_{R}} + \int_0^{t \wedge \tau^{n,m}_{R}} 2(p^{(n)}_{s} - p^{(m)}_{s}, f(s, X^{(n)}_{s \land k(m,s)}) - f(s, X^{(m)}_{s \land k(m,s)}))\,ds \]

\[ + \int_0^{t \wedge \tau^{n,m}_{R}} |L_R(s)| \left( \sup_{u \in [-\tau,s]} |X^{(n)}_{u \land k(m,s)} - X^{(m)}_{u \land k(m,s)}|^2 \right)\,ds \]

\[ \leq M_{t \wedge \tau^{n,m}_{R}} + H_{t \wedge \tau^{n,m}_{R}} + \int_0^T L_R(s) \left( \sup_{u \in [-\tau,s]} |X^{(n)}_{u \wedge \tau^{n,m}_{R}} - X^{(m)}_{u \wedge \tau^{n,m}_{R}}|^2 \right)\,ds \]  

(23)
where
\[
P_t^{n,m,R} := 4 \mathbb{1}_{(0,\tau_{R}^{(n)})}(s) \sup_{u \in [k(n)s,s]} |p_u^{(n)}|^2 + 4 \mathbb{1}_{(0,\tau_{R}^{(m)})}(s) \sup_{u \in [k(m)s,s]} |p_u^{(m)}|^2,
\]
\[
H_t^{n,m,R} := \int_0^t \mathbb{1}_{(0,\tau_{R}^{n,m})}(s) \left(4|p_s^{(n)}| + |p_s^{(m)}|\right) \tilde{K}_R(s) ds + \int_0^t L_R(s) \left( 2 \sup_{u \in [-\tau,s]} |X_u^{n,m} - X_u^{m,m} |^2 + p_s^{n,m,R} \right) ds.
\]

This implies by \(\varepsilon\) and the references therein. Using the stochastic Bihari-LaSalle inequality for see e.g. by Cox, Hutzenthaler and Jentzen [7], Hudde, Hutzenthaler and Mazzonetto 5.2 Exponential moment estimates for path-dependent SDEs

\[
\text{We have due to dominated convergence, (20), a)} \text{ and (C5):}
\]
\[
\limsup_{n,m \to \infty} \mathbb{E}[H_T^{n,m,R}] = 0.
\]
We multiply (23) by 2 and add an \(\varepsilon > 0\). Then, Theorem 3.9 implies:
\[
\mathbb{E}[G_1(2 \sup_{t \in [0,T]} |X_t^{(n)} - X_t^{(m)}|^2 + \varepsilon)] \leq G_1(2\mathbb{E}[H_T^{n,m,R} + \varepsilon]) + \mathbb{E}[2 \int_0^T L_R(t) ds].
\]
Therefore, we have for any fixed \(a > 0\) (using the same argument as in the proof of uniqueness)
\[
\mathbb{P}[ \sup_{t \in [0,\tau_{R}^{n,m} \land T]} |X_t^{(n)} - X_t^{(m)}|^2 > a] \leq \frac{G_1(\mathbb{E}[H_T^{n,m,R} + \varepsilon]) - G_1(\varepsilon)}{G_1(a + \varepsilon) - G_1(\varepsilon)} \int_0^T L_R(s) ds \to \frac{\mathbb{E}[\int_0^T L_R(s) ds]}{G_1(a + \varepsilon) - G_1(\varepsilon)} \text{ for } n, m \to \infty.
\]
This implies by \(\varepsilon \to 0\) that \(\lim_{n,m \to \infty} \mathbb{P}[\sup_{t \in [0,\tau_{R}^{n,m} \land T]} |X_t^{(n)} - X_t^{(m)}|^2 > a] = 0 \text{ for any } a > 0\). By [C] we have:
\[
\limsup_{n,m \to \infty} \mathbb{P}[\sup_{t \in [0,T]} |X_t^{(n)} - X_t^{(m)}| > a] \leq \limsup_{R \to \infty} \limsup_{n,m \to \infty} \left( \mathbb{P}[T > \tau_{R}^{(n)}] + \mathbb{P}[T > \tau_{R}^{(m)}] + \mathbb{P}[\sup_{t \in [0,\tau_{R}^{n,m} \land T]} |X_t^{(n)} - X_t^{(m)}| > a] \right) = 0,
\]
which implies the claim. □

5.2 Exponential moment estimates for path-dependent SDEs

Exponential integrability bounds for solutions of non-path-dependent SDEs are known, see e.g. by Cox, Hutzenthaler and Jentzen [7], Hudde, Hutzenthaler and Mazzonetto [13] and the references therein. Using the stochastic Bihari-LaSalle inequality for \(\eta(x) = x(\log(x) + c)\) we provide a similar result for path-dependent SDEs driven by Brownian motion.

For \(U = (U(s,y))_{s \in [0,T], y \in \mathbb{R}^d} \in C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})\) we define for all \(x \in \text{Càdlàg}([-\tau, \infty); \mathbb{R}^d)\) and all \(t \geq 0\)
\[
(\mathcal{G}_{f,h} U)(t,x) := \left( \frac{\partial}{\partial s} U \right)(t,x(t)) + \left( \frac{\partial}{\partial y} U \right)(t,x(t)) f(t,x) + \frac{1}{2} \text{trace}(h(t,x)h(t,x)^T \text{Hess}_y U(t,x(t))).
\]
Here, $f$ and $h$ denote the coefficients of the SDE (16).

One difference of the following corollary to corresponding results for non-path-dependent SDEs, see [7, Corollary 2.4] and [13, Corollary 3.3], is, that we assume that $U$ is non-negative.

**Corollary 5.4** (Exponential moment estimates for path-dependent SDEs). Let $X$ be a solution of the SDE (16) and assume that $g \equiv 0$. Let $U = (U(s, y))_{s \in [0, T], y \in \mathbb{R}^d} \in C^{1,2}[0, T] \times \mathbb{R}^d, [0, \infty))$, and let $\gamma \geq 0$ and $\kappa \geq 0$. Assume that for all $x \in \mathcal{C}^\alpha(\mathbb{R}^d)$, $t \geq 0$

$$(G_{f,h}U)(t, x) + \frac{1}{2}\|\frac{\partial U}{\partial y}(t, x(t))h(t, x)\|^2 \leq \gamma \sup_{s \in [0,t]} U(s, x(s)) + \gamma \kappa.$$  

(24)

Then, for all $p \in (0, 1)$ and all $T \geq 0$

$$\mathbb{E}_{\mathcal{F}_0}\left[ \sup_{t \in [0,T]} \exp(pU(t, X_t)e^{-\gamma \beta T}) \right]^{1/p} \leq \alpha_1 \alpha_2 \|\exp(U(0, X_0))\|_{p, \mathcal{F}_0} e^{(\kappa + U_0)(1 - e^{-\gamma \beta T})},$$

where $U_0 := \sup_{s \in [0, t]} U(s, z_s)$ and $\alpha_1, \alpha_2$ and $\beta$ only depend on $p$ (see (8)).

**Proof of Corollary 5.4.** We apply Itô’s formula to compute $(Y_t)_{t \geq 0} := (U(t, X_t))_{t \geq 0}$:

$$dY_t = dU(t, X_t) = (G_{f,h}U)(t, X_t)dt + \left(\frac{\partial U}{\partial y}(t, X_t)h(t, X)\right)dB_t.$$  

We apply Itô’s Formula to compute $(Z_t)_{t \geq 0} := (\exp(Y_t))_{t \geq 0}$:

$$dZ_t = d\exp(Y_t) = \exp(Y_t)dY_t + \frac{1}{2}\exp(Y_t)d\langle Y \rangle_t$$

$$= \exp(Y_t)\left((G_{f,h}U)(t, X_t) + \frac{1}{2}\|\frac{\partial U}{\partial y}(t, X_t)h(t, X)\|^2\right)dt + M_t$$

$$\leq \gamma \exp(Y_t)(Y_t^* + \kappa + \sup_{s \in [0, t]} U(s, z_s))dt + \tilde{M}_t$$

where $M := \int_0^t V(Y_s)\left(\frac{\partial U}{\partial y}(s, X_s)h(t, X)dB_s$ is a local martingale which starts in $0$. By assumption $U \geq 0$, and therefore $Z \geq 1$. Hence, we have for $\eta(x) = x(\log(x) + \kappa + U_0)$ for $x \geq 1$, where $U_0 := \sup_{s \in [0, t]} U(s, z_s)$:

$$Z_t \leq Z_0 + \int_0^t \eta(Z_s^*)\gamma ds + \tilde{M}_t.$$  

Using that

$$G(x) = \log(\kappa + U_0 + \log(x)), \quad G^{-1}(x) = \exp(e^x - \kappa - U_0),$$

implies

$$\|G^{-1}(G(Z_T^*) - \beta A_T)\|_{p, \mathcal{F}_0} = \|\exp(\exp(\log(\kappa + U_0 + \log(Z_T^*)) - \beta \gamma T)) - \kappa - U_0\|_{p, \mathcal{F}_0}$$

$$= \|\exp(\exp(\log(\kappa + U_0 + \log(Z_T^*))e^{-\beta \gamma T}) - \kappa - U_0\|_{p, \mathcal{F}_0}$$

$$= \|\exp(\{\kappa + U_0 + \log(Z_T^*)\}e^{-\beta \gamma T}) - \kappa - U_0\|_{p, \mathcal{F}_0}$$

$$= \|(Z_T^*)^e^{-\beta \gamma T}\exp(\{\kappa + U_0\}e^{-\beta \gamma T})) - \kappa - U_0\|_{p, \mathcal{F}_0}.$$  

Applying Theorem 3.6 and rearranging the terms yields:

$$\sup_{t \in [0, T]} \exp(U(t, X_t)e^{-\gamma \beta T}) \leq \alpha_1 \alpha_2 \|Z_0\|_{p, \mathcal{F}_0} e^{-(\kappa + U_0)e^{-\gamma \beta T} e^{\kappa + U_0}}.$$  

□
The constants $\alpha_1$, $\alpha_2$ and $\beta$ occuring in the following examples are defined in (6) and they only depend on $p$.

**Example 5.5.** Let $X$ be a solution of the path-dependent SDE (16) and assume $g \equiv 0$. Choosing $U(t, x) := r \|x\|^2$ for some $r > 0$ in Corollary 5.4 implies: Assume there exist constants $\gamma_1 \geq 0$, $\gamma_2 \geq 0$ such that for all $t \in [0, T]$; $x \in \mathbb{C}$
\[
\langle x(t), f(t, x) \rangle \leq \gamma_1 \sup_{s \in [-\tau, t]} \|x(s)\|^2, \quad \|h(t, x)\|^2 \leq \gamma_2.
\]
Then, we have
\[
\mathbb{E}_{\mathcal{F}_0} \left[ \sup_{t \in [0, T]} \exp(pr\|X_t\|^2e^{-\gamma T}) \right]^{1/p} \leq \alpha_1 \alpha_2 \exp(U(0, X_0)) \|p, \mathcal{F}_0 e^{r(\kappa_U)0(1-e^{-rT})}
\]
where $\gamma := 2\gamma_1 + 2r\gamma_2$, $\kappa := \frac{r\gamma_2}{2\gamma_1 + 2\gamma_2}$.

This can be seen as follows: For this choice of $U$ (24) can be strengthened to:
\[
2r \langle x(t), f(t, x) \rangle + r\|h(t, x)\|^2 + 2r^2\|x(t)\|^2\|h(t, x)\|^2 \leq \gamma r \sup_{s \in [-\tau, t]} \|x(s)\|^2 + \gamma \kappa
\]
Choosing $\gamma$ and $\kappa$ such that $2r\gamma_1 + 2r^2\gamma_2 = \gamma r$ and $r\gamma_2 = \gamma \kappa$ implies the assertion.

**Example 5.6.** Let $X$ be a solution of the path-dependent SDE (16) and assume $g \equiv 0$. Choosing $U(t, x) := r(\|x\|^2 + 1)^{1/2}$ for some $r > 0$ in Corollary 5.4 implies: Assume there exist constants $\gamma_1 \geq 0$, $\gamma_2 \geq 0$ such that for all $t \in [0, T]$; $x \in \mathbb{C}$
\[
\langle x(t), f(t, x) \rangle \leq \gamma_1 \sup_{s \in [-\tau, t]} (\|x(s)\|^2 + 1)^{1/2} (1 + \|x(t)\|^2)^{1/2},
\]
\[
\|h(t, x)\|^2 \leq \gamma_2 \sup_{s \in [-\tau, t]} (\|x(s)\|^2 + 1)^{1/2}.
\]
Then, we have
\[
\mathbb{E}_{\mathcal{F}_0} \left[ \sup_{t \in [0, T]} \exp(pr(\|X_t\|^2 + 1)^{1/2}e^{-\gamma T}) \right]^{1/p} \leq \alpha_1 \alpha_2 \exp(U(0, X_0)) \|p, \mathcal{F}_0 e^{U_0(1-e^{-rT})}
\]
where $\gamma := \gamma_1 + \frac{r}{2} \gamma_2 + \frac{r}{2} \gamma_2$.

This can be seen as follows: For this choice of $U$ and using that trace(Hess$_g U(t, x(t))) \leq rd(1 + \|x(t)\|^2)^{-1/2}$ and trace($A_1A_2$) $\leq$ trace($A_1A_2$) for positive semidefinite matrices $A_1$, $A_2$, the assumption (24) to strengthened to
\[
(1 + \|x(t)\|^2)^{-1/2}r \langle x(t), f(t, x) \rangle + \frac{1}{2} \text{trace}(h(t, x)h(t, x)^T)rd(1 + \|x(t)\|^2)^{-1/2}
\]
\[
+ \frac{1}{2}(1 + \|x(t)\|^2)^{-1}r^2\|h(t, x)\|^2\|x(t)\|^2 \leq \gamma r \sup_{s \in [-\tau, t]} (\|x(s)\|^2 + 1)^{1/2}.
\]
Multiplying with $(1 + \|x(t)\|^2)^{1/2}r^{-1}$ implies:
\[
\langle x(t), f(t, x) \rangle + \frac{r}{2}\|h(t, x)\|^2 + \frac{1}{2}r(1 + \|x(t)\|^2)^{-1/2}\|h(t, x)\|^2\|x(t)\|^2
\]
\[
\leq \gamma \sup_{s \in [-\tau, t]} (\|x(s)\|^2 + 1)^{1/2}(1 + \|x(t)\|^2)^{1/2}.
\]
Choosing $\gamma := \gamma_1 + \frac{r}{2} \gamma_2 + \frac{r}{2} \gamma_2$ ensures, that the coefficients satisfy (24).
6 Main tool for the proofs

6.1 Key idea of the proofs

Repeatedly, we prove results of this paper by extending the proof of Lenglart’s inequality (see e.g. [17, Corollaire II] or [24, Lemma 2.2 (ii)] or the references listed in [24]).

**Lemma 6.1** (Lenglart’s domination inequality). Let $X$ be a non-negative adapted right-continuous process and let $H$ be a non-negative right-continuous non-decreasing predictable process such that $\mathbb{E}[X_\tau \mid \mathcal{F}_0] \leq \mathbb{E}[H_\tau \mid \mathcal{F}_0] \leq \infty$ for any bounded stopping time $\tau$. Then, for all $p \in (0, 1)$:

$$\left\| \sup_{t \geq 0} X_t \right\|_{p, \mathcal{F}_0} \leq \alpha_1 \alpha_2 \left\| \sup_{t \geq 0} H_t \right\|_{p, \mathcal{F}_0},$$

where the constants $\alpha_1$ and $\alpha_2$ only depend on $p$ and are defined in [8].

We will frequently use the formulas of the subsequent remark (see also the proof of [27, Proposition 1.2]):

**Remark 6.2** (Calculation of $Z^p$, $p \in (0, 1)$). Let $Z$ be a non-negative random variable and $p \in (0, 1)$. Then $Z^p$ can be calculated using the three formulas below.

$$Z^p = p \int_0^\infty 1_{\{Z \geq u\}} u^{p-1} du$$

$$Z^p = (1 - p) \int_0^\infty Z 1_{\{Z \leq u\}} u^{p-2} du$$

$$Z^p = p(1 - p) \int_0^\infty (Z \wedge u) u^{p-2} du \quad (25)$$

The third equality follows e.g. by using the first and second equality. In particular, we also have $Z^{p-1} = (1 - p) \int_0^\infty 1_{\{Z \leq u\}} u^{p-2} du$.

We provide a short sketch of the proof of Lemma 6.1 as we generalize this procedure in the subsequent proofs: First one proves for all $u > 0, \lambda > 0, t > 0$:

$$\mathbb{P}[X_t^* > u] \leq \frac{1}{u} \mathbb{E}[H_t \wedge (\lambda u)] + \mathbb{P}[H_t \geq \lambda u], \quad (26)$$

see [17, Théorème I] or [24, Lemma 2.2 (ii)]. Lenglart’s domination inequality follows by integrating equation (26) w.r.t. $u^{p-1} du$, applying the formulas (25) and optimizing over $\lambda$, i.e. choosing $\lambda = p$.

A generalization of (26) is provided in Lemma 6.3 for the cases $A_{\sup}, A_{nosup}$ and $A_{\sup, \theta}$. In the proofs of the results of this paper we often integrate at some point inequality (27) of Lemma 6.3 w.r.t. $d\mathbb{P} u^{p-1} du$.

6.2 Another Lenglart type inequality

**Lemma 6.3** below is the key step of the proofs of this paper. It is a generalization of Lenglart’s estimate (26), noting that

$$1_{\{X_t^* > u\}} \leq \frac{1}{u} (X_{t \wedge \sigma} \wedge u) \quad \text{for } \sigma := \inf\{s \geq 0 \mid X_s > u\},$$

where $X$ is some non-negative right-continuous process. As in the proof of [27, Proposition 1.2], we will sometimes work with $X_{t \wedge \sigma} \wedge u$ instead of $u 1_{\{X_t^* > u\}}$. 

26
Lemma 6.3 (Lenglart type estimates). Fix some $T > 0$ and $p \in (0, 1)$. We consider the following 9 cases, which arise from combining $\mathcal{A}_{\sup}$, $\mathcal{A}_{\nosup}$ or $\mathcal{A}_{\sup,\theta}$ with one of the following three assumptions:

a) $H$ is predictable and $\mathbb{E}[H_T^p] < \infty$,

b) $M$ has no negative jumps and $\mathbb{E}[H_T^p] < \infty$,

c) $\mathbb{E}[H_T] < \infty$.

Fix arbitrary $u, \lambda > 0$ and set:

$$
\tau := \inf\{s \geq 0 \mid H_s \geq \lambda u\}, \quad \sigma := \inf\{s \geq 0 \mid X_s > u\},
$$

where $\inf \emptyset := +\infty$. Then, the following estimate holds true for all $t \in [0, T]$:

$$
\mathbb{1}_{\{X_t^* > u\}} u \leq X_t \wedge \sigma \wedge u \leq I_t^L + M_t^L + H_t^L,
$$

where $(I_t^L)_{t \geq 0}$ is a non-decreasing process containing the integral term with an additional indicator function

$$
I_t^L := I_t^{L,u} := \begin{cases} 
\int_{(0,t]} \eta(X_s^*) \mathbb{1}_{\{X_s^* \leq u\}} dA_s & \text{for } \mathcal{A}_{\sup}, \\
\int_{(0,t]} \eta(X_s^-) \mathbb{1}_{\{X_s^- \leq u\}} dA_s & \text{for } \mathcal{A}_{\nosup}, \\
\left(\int_{(0,t]} (X_s^*)^\theta \mathbb{1}_{\{X_s^* \leq u\}} dA_s\right)^{1/\theta} & \text{for } \mathcal{A}_{\sup,\theta}, 
\end{cases}
$$

the process $(M_t^L)_{t \geq 0}$ is the local martingale defined by

$$
M_t^L := M_t^{L,u} := \begin{cases} 
\lim_{n \to \infty} M_{t \wedge \tau^{(n)} \wedge \sigma} & \text{if } H \text{ is predictable and } \mathbb{E}[H_T^p] < \infty, \\
M_{t \wedge \tau \wedge \sigma} & \text{if } M \text{ has no negative jumps and } \mathbb{E}[H_T^p] < \infty, \\
M_{t \wedge \sigma} \mathbb{1}_{\{E_T[H_T] \leq u\}} & \text{if } \mathbb{E}[H_T] < \infty, 
\end{cases}
$$

(where $\tau^{(n)}$ denotes a localizing sequence of $\tau$ and $\tilde{M}_t := M_t + \mathbb{E}[H_T \mid F_t] - \mathbb{E}[H_T \mid F_0]$ for $t \in [0, T]$), and $(H_t^L)_{t \geq 0}$ is a non-decreasing process depending on $H$:

$$
H_t^L := H_t^{L,u} := \begin{cases} 
H_t \wedge (\lambda u) + u \mathbb{1}_{\{H_t \geq \lambda u\}} & \text{if } H \text{ is predictable and } \mathbb{E}[H_T^p] < \infty, \\
H_t \wedge (\lambda u) + u \mathbb{1}_{\{H_t \geq \lambda u\}} & \text{if } M \text{ has no negative jumps and } \mathbb{E}[H_T^p] < \infty, \\
E_{\mathcal{F}_0[H_T]} \wedge u & \text{if } \mathbb{E}[H_T] < \infty. 
\end{cases}
$$

Proof of Lemma 6.3. We denote by $Y$ the upper bound for $X$ which we have be $\mathcal{A}_{\sup}$, $\mathcal{A}_{\nosup}$ or $\mathcal{A}_{\sup,\theta}$ respectively; i.e. for all $t \geq 0$

$$
X_t \leq Y_t := M_t + H_t + \begin{cases} 
\int_{(0,t]} \eta(X_s^*) dA_s & \text{for } \mathcal{A}_{\sup}, \\
\int_{(0,t]} \eta(X_s^-) dA_s & \text{for } \mathcal{A}_{\nosup}, \\
\left(\int_{(0,t]} (X_s^*)^\theta dA_s\right)^{1/\theta} & \text{for } \mathcal{A}_{\sup,\theta}. 
\end{cases}
$$
**Step (a):** We first prove the inequality for the case that $H$ is predictable and $\mathbb{E}[H_T^p] < \infty$. Fix some $t \in [0, T]$. Because $H$ is predictable there exists a sequence of stopping times $(\tau^{(n)})_{n \in \mathbb{N}}$ that announces $\tau$. In particular, due to $H$ being non-decreasing, we have on $\{H_0 < \lambda u\}$ the inequality $H_{t \wedge \tau^{(n)} \wedge \sigma} \leq H_t \wedge (\lambda u)$. Moreover, by definition of $\sigma$ the equality

$$\{s \leq \sigma\} = \{s > \sigma\}^c = \{\exists \tau < s \mid X_\tau > u\}^c = \{X_\tau^* > u\}^c = \{X_\tau^* \leq u\}$$

holds true for all $s > 0$. Therefore, using (11), (12) or (13) respectively, we have on $\{H_0 < \lambda u\}$:

$$X_{t \wedge \tau^{(n)} \wedge \sigma} \leq I_t^L + M_{t \wedge \tau^{(n)} \wedge \sigma} + H_t \wedge (\lambda u).$$

Moreover, note that due to non-negativity of $X$, we have

$$X_{t \wedge \sigma} \wedge u - X_{t \wedge \tau^{(n)} \wedge \sigma} \wedge u \leq u 1_{\{\tau^{(n)} < t\}}.$$

The previous two inequalities imply on $\{H_0 < \lambda u\}$:

$$X_{t \wedge \sigma} \wedge u \leq \limsup_{n \to \infty} X_{t \wedge \tau^{(n)} \wedge \sigma} \wedge u + \limsup_{n \to \infty} \left(X_{t \wedge \sigma} \wedge u - X_{t \wedge \tau^{(n)} \wedge \sigma} \wedge u\right)$$

$$\leq I_t^L + \lim_{n \to \infty} M_{t \wedge \tau^{(n)} \wedge \sigma} + H_t \wedge (\lambda u) + \lim_{n \to \infty} u 1_{\{\tau^{(n)} < t\}},$$

$$\leq I_t^L + \lim_{n \to \infty} M_{t \wedge \tau^{(n)} \wedge \sigma} + H_t \wedge (\lambda u) + u 1_{\{H_t \geq \lambda u\}}.$$

On $\{H_0 \geq \lambda u\}$ we have

$$X_{t \wedge \sigma} \wedge u \leq u 1_{\{H_t \geq \lambda u\}}.$$

Noting that on $\{H_0 \geq \lambda u\}$ we have $\tau^{(n)} = \tau = 0$ and $M_0 = 0$, this implies by non-negativity of $I_t^L$ (27).

**Step (b):** Next we prove the inequality for the case that $M$ has no negative jumps and $\mathbb{E}[H_T^p] < \infty$. Fix again some $t \in [0, T]$. In the proof of the previous assertion we used $\lim_{n \to \infty} H_{t \wedge \tau^{(n)}} \leq H_t \wedge (\lambda u)$ on $\{H_0 < \lambda u\}$. As the existence of an announcing sequence of $\tau$ is not guaranteed if $H$ is not predictable, we need to take into account that $H$ might jump at time above $\lambda u$. To this end, we define for all $t \geq 0$:

$$\tilde{X}_t := X_{t \wedge \tau} 1_{\{t < \tau\}} = X_t 1_{\{t < \tau\}},$$

$$\tilde{Y}_t := M_{t \wedge \tau} + H_t \wedge (\lambda u) + \begin{cases} \int_{(0,t \wedge \tau]} \eta(X_s^*) dA_s & \text{for } A_{\text{sup}}, \\ \int_{(0,t \wedge \tau]} \eta(X_s^-) dA_s & \text{for } A_{\text{no sup}}, \\ \left(\int_{(0,t \wedge \tau]} (X_s^*)^\theta dA_s\right)^{1/\theta} & \text{for } A_{\text{sup, } \theta}. \end{cases}$$

It can be seen that $\tilde{X}$ is an adapted right-continuous process. Furthermore, $\tilde{Y}$ is non-negative: The local martingale $M$ having no negative jumps implies that $\tilde{Y}$ has no negative jumps, and hence $0 \leq \tilde{Y}_{t^-} = \tilde{Y}_t - \tilde{Y}_{t^-}$. This implies

$$\tilde{X}_t = \begin{cases} X_t & \text{on } \{t < \tau\} \\ 0 & \text{on } \{t \geq \tau\} \leq \begin{cases} Y_t & \text{on } \{t < \tau\} \\ \tilde{Y}_t & \text{on } \{t \geq \tau\} \leq \tilde{Y}_t \end{cases}.$$
for all \( t \geq 0 \). By construction we have \( \tilde{X}_s = X_s \) for \( s \in [0, t] \) on \( \{ \tau > t \} = \{ H_t < \lambda u \} \). Therefore, on \( \{ H_0 < \lambda u \} \) we have for all \( t \in [0, T] \)

\[
X_{t \land \sigma} \land u = \tilde{X}_{t \land \tau \land \sigma} \land u + \left( X_{t \land \sigma} \land u - \tilde{X}_{t \land \tau \land \sigma} \land u \right)
\]

\[
\leq \tilde{Y}_{t \land \sigma \land \tau} + u 1_{\{ \tau \leq t \}}
\]

\[
\leq I_t^L + M_{t \land \tau \land \sigma} + H_t \land (\lambda u) + u 1_{\{ H_t \geq \lambda u \}}.
\]

On \( \{ H_0 \geq \lambda u \} \) we have as before: \( X_{t \land \sigma} \land u \leq u 1_{\{ H_t \geq \lambda u \}} \). Combining the inequalities for \( \{ H_0 < \lambda u \} \) and \( \{ H_0 \geq \lambda u \} \) implies the claim.

**Step (c):** Now we prove the inequality for the case that \( \mathbb{E}[H_T] < \infty \). We weaken (11), (12) and (13) as follows:

\[
X_t \leq \tilde{M}_t + \mathbb{E}_{\mathcal{F}_0}[H_T] + \begin{cases} 
\int_{(0,t]} \eta(X^*_s) dA_s & \text{for } A_{\text{sup}}, \\
\int_{(0,t]} \eta(X^*_s) dA_s & \text{for } A_{\text{nosup}}, \\
\left( \int_{(0,t]} (X^*_s)^\theta dA_s \right)^{1/\theta} & \text{for } A_{\text{sup}, \theta},
\end{cases}
\]

where \( \tilde{M}_t := M_t + \mathbb{E}[H_T | F_t] - \mathbb{E}[H_T | \mathcal{F}_0] \). As our filtration satisfies by assumption the usual conditions, we may assume w.l.o.g. that \( \tilde{M} \) is càdlàg. We obtain for all \( u > 0 \) on \( \{ \mathbb{E}_{\mathcal{F}_0}[H_T] \leq u \} \):

\[
X_{t \land \sigma} \land u \leq I_t^L + \tilde{M}_{t \land \sigma} + \mathbb{E}_{\mathcal{F}_0}[H_T] \land u.
\]

Noting that \( X_{t \land \sigma} \land u \leq \mathbb{E}[H_T | \mathcal{F}_0] \land u \) on \( \{ \mathbb{E}_{\mathcal{F}_0}[H_T] \geq u \} \) implies the claim.

## 7 Proofs of the stochastic Bihari-LaSalle inequalities

**Remark 7.1.** In the proofs of the stochastic Bihari-LaSalle inequalities, we will frequently assume \( X \geq \varepsilon \) and \( H \geq \varepsilon \) for some \( \varepsilon > 0 \). We do this to ensure that terms like \( G(X_t) \) or \( G(H_t) \) are well-defined and to ensure that we may apply the deterministic Bihari-LaSalle inequality, see [Lemma 1.1]. We may do this without loss of generality because we can add an arbitrary \( \varepsilon > 0 \) to (11) (or similarly (12)) and slightly weaken it to obtain for all \( t \in [0, T] \):

\[
(X_t + \varepsilon) \leq \int_{[0,t]} \eta(X^*_s + \varepsilon) dA_s + M_t + (H_t + \varepsilon) \quad \mathbb{P}\text{-a.s.}
\]

Proving the assertions of the theorems for the processes \( (X_t + \varepsilon)_{t \geq 0} \), \( (A_t)_{t \geq 0} \), \( (M_t)_{t \geq 0} \) and \( (H_t + \varepsilon)_{t \geq 0} \) and then taking the limit \( \varepsilon \to 0 \) will imply the assertions for the general case \( X \geq 0, H \geq 0 \).

### 7.1 Proofs for concave \( \eta \) ([Theorem 3.1])

**Proof of [Theorem 3.1]**. We assume w.l.o.g. that the local martingale \( M^L \) from [Lemma 6.3] is a martingale.
Proof for predictable $H$ and $\Delta M \geq 0$: Fix some $t \in [0, T]$ and assume $\mathbb{E}[H_T^p] < \infty$. Assume that either $H$ is predictable or that $M$ has no negative jumps. Then Lemma 6.3 implies (under the assumption $\mathcal{A}_{\text{sup}}$) for all $u > 0, \lambda > 0, t > 0$:

$$1_{\{X_t^* > u\}} \leq \frac{1}{u} \int_{(0,t]} \eta(X_{s^-}^*) 1_{\{X_{s^-} \leq u\}} \, dA_s + \frac{1}{u} M_t^L + \frac{1}{u} H_t \wedge (\lambda u) + 1_{\{H_t \geq \lambda u\}},$$

We multiply the equation above with $pu^{p-1}$, take the conditional expectation given $\mathcal{F}_0$ and integrate over $u$. This implies (using Fubini):

$$\int_0^\infty \mathbb{P}_{\mathcal{F}_0}[X_t^* > u] pu^{p-1} \, du \leq \mathbb{E}_{\mathcal{F}_0} \left[ \int_{(0,t]} \eta(X_{s^-}^*) \int_0^\infty 1_{\{X_{s^-} \leq u\}} pu^{p-2} \, du \, dA_s \right] + \lambda \int_0^\infty \mathbb{E}_{\mathcal{F}_0}[(H_t \lambda^{-1}) \wedge u] pu^{p-2} \, du + \int_0^\infty \mathbb{P}_{\mathcal{F}_0}[H_t \lambda^{-1} \geq u] pu^{p-1} \, du.$$

Using (25) we have:

$$\mathbb{E}_{\mathcal{F}_0}[(X_t^*)^p] \leq \frac{p}{1-p} \mathbb{E}_{\mathcal{F}_0} \left[ \int_{(0,t]} \eta(X_{s^-}^*) (X_{s^-}^*)^{p-1} \, dA_s \right] + (\lambda(1-p))^{-1} \lambda^{-p} + \lambda^{-p} \mathbb{E}_{\mathcal{F}_0}[H_T^p].$$

An easy calculation gives that the choice $\lambda = p$ is optimal. Set $\eta_p(x) := \frac{p}{1-p} \eta(x^{1/p})x^{1-1/p}$ for all $x > 0$. Furthermore, by assumption, $\eta_p$ is concave and $A$ is deterministic. Therefore, we obtain by Jensen’s inequality for all $t \in [0, T]$:

$$
\mathbb{E}_{\mathcal{F}_0}[(X_t^*)^p] \leq \int_{(0,t]} \eta_p(\mathbb{E}_{\mathcal{F}_0}[(X_{s^-}^*)^p]) \, dA_s + (\alpha_1 \alpha_2)^p \mathbb{E}_{\mathcal{F}_0}[H_T^p]. \tag{29}
$$

Since $\eta_p$ is non-decreasing and $s \mapsto \mathbb{E}_{\mathcal{F}_0}[(X_{s^-}^*)^p]$ is càdlàg, we may apply the deterministic Bihari-LaSalle inequality (see Lemma 1.1) to the previous inequality. Recall the following definition and its properties (see (8)):

$$\tilde{G}_p(x) := \int_{cp}^x \frac{du}{\eta_p(u)}, \quad \text{satisfying} \quad \tilde{G}_p(x) = (1-p)G(x^{1/p}), \quad \tilde{G}_p^{-1}(x) = \left(G^{-1}(\frac{x}{1-p})\right)^p.$$

The deterministic Bihari-LaSalle inequality implies:

$$\|X_t^*\|_{p, \mathcal{F}_0} \leq \left\{ \tilde{G}_p^{-1} \left( \tilde{G}_p(\alpha_1^p \alpha_2^p \mathbb{E}_{\mathcal{F}_0}[H_T^p]) + A_T \right) \right\}^{1/p} = \left\{ \tilde{G}_p^{-1} \left( (1-p)G(\alpha_1 \alpha_2 \|H_T\|_{p, \mathcal{F}_0}) + A_T \right) \right\}^{1/p} = G^{-1} \left( G(\alpha_1 \alpha_2 \|H_T\|_{p, \mathcal{F}_0}) + \beta A_T \right) \cdot$$

Proof for $\mathbb{E}[H_T] < \infty$: The proof is very similar to the previous case. We take the conditional expectation of (27) given $\mathcal{F}_0$ and integrating w.r.t. $pu^{p-2} \, du$, which implies using (25):

$$\mathbb{E}_{\mathcal{F}_0}[(X_t^*)^p] \leq \frac{p}{1-p} \mathbb{E}_{\mathcal{F}_0} \left[ \int_{(0,t]} \eta(X_{s^-}^*) (X_{s^-}^*)^{p-1} \, dA_s \right] + \int_0^\infty (\mathbb{E}_{\mathcal{F}_0}[H_t] \wedge u) pu^{p-2} \, du$$

$$\quad = \mathbb{E}_{\mathcal{F}_0} \left[ \int_{(0,t]} \eta_p((X_{s^-}^*)^p) \, dA_s \right] + \alpha_1^p \mathbb{E}_{\mathcal{F}_0}[H_T^p] \tag{30}$$

The same arguments as in the proof for predictable $H$ imply the claim. \qed

30
Proof of Theorem 3.1 [b]. We assume w.l.o.g. that $M$ is a martingale. As before, we define the following stopping times for some fixed $u, \lambda > 0$:

$$\tau_u := \inf\{s \geq 0 \mid H_s \geq \lambda u\}, \quad \sigma_u := \inf\{s \geq 0 \mid X_s > u\}.$$

We first prove the assertion for the case $\mathbb{E}[H_T] < \infty$, which is a simple extension of the proof given in [36, Lemma 3.8].

Proof for $\mathbb{E}[H_T] < \infty$: As $\eta$ is concave and $H$ is non-decreasing, we have for all $t \in [0, T]$:

$$\mathbb{E}_{\mathcal{F}_t}[X_{t \wedge \sigma_u} \wedge u] \leq \mathbb{E}_{\mathcal{F}_t}\left[\int_{(0,t] \wedge \sigma_u} \eta(X_s^-)dA_s + \mathbb{E}_{\mathcal{F}_t}[H_T]\right]$$

$$\leq \mathbb{E}_{\mathcal{F}_t}\left[\int_{(0,t]} \eta(X_{s^-} \wedge u) dA_s + \mathbb{E}_{\mathcal{F}_t}[H_T]\right]$$

$$\leq \int_{(0,t]} \eta(\mathbb{E}_{\mathcal{F}_t}[X_{s^-} \wedge u]) dA_s + \mathbb{E}_{\mathcal{F}_t}[H_T]$$

Recall that $X$ is a non-negative process. Therefore, dominated convergence implies that $t \mapsto \mathbb{E}[X(t \wedge \sigma_u) \wedge u]$ is càdlàg. Thus, we may apply the deterministic Bihari-LaSalle inequality [Lemma 1.1] which gives:

$$\mathbb{P}_{\mathcal{F}_0}[X_t^+ > u]u \leq \mathbb{E}_{\mathcal{F}_0}[X_{t \wedge \sigma_u} \wedge u] \leq G^{-1}(G(\mathbb{E}_{\mathcal{F}_0}[H_T]) + A_T) =: \delta$$

and hence

$$\|X_t^+\|_{p,\mathcal{F}_0}^p = p \int_0^\infty \mathbb{P}[X_t^+ > u]u^{p-1}du \leq p \int_0^\infty (\delta \wedge u)u^{p-2}du$$

$$= \frac{1}{1-p} \delta^p = \frac{1}{1-p} G^{-1}(G(\mathbb{E}_{\mathcal{F}_0}[H_T]) + A_T)^p$$

which implies the claim.

Proof for predictable $H$ and $\Delta M \geq 0$: Assume that $\mathbb{E}[H_T^p] < \infty$ and that either $H$ is predictable or $\Delta M \geq 0$. By [Lemma 6.3] (using the case $A_{\text{nasup}}$) we have

$$\mathbb{E}_{\mathcal{F}_t}[X_{t \wedge \sigma_u} \wedge u] \leq \mathbb{E}_{\mathcal{F}_t}\left[\int_{(0,t]} \eta(X_s^-)\mathbb{1}_{\{X_s^- \leq u\}}dA_s\right]$$

$$+ \lambda \mathbb{E}_{\mathcal{F}_t}[H_t \lambda^{-1} \wedge u] + u \mathbb{P}_{\mathcal{F}_0}[H_t \lambda^{-1} \geq u].$$

Integrating w.r.t to $p(1-p)u^{p-2}du$, using [Remark 6.2] and choosing $\lambda = p$ gives:

$$y(t) := p(1-p) \int_0^\infty \mathbb{E}_{\mathcal{F}_0}[X_{t \wedge \sigma_u} \wedge u]u^{p-2}du$$

$$\leq p \mathbb{E}_{\mathcal{F}_0}\left[\int_{(0,t]} \eta(X_s^-)(X_s^-)^{p-1}dA_s + p^{-p} \mathbb{E}_{\mathcal{F}_0}[H_T^p]\right]$$

$$\leq (1-p) \int_{(0,t]} \mathbb{E}_{\mathcal{F}_t}[\eta_p(X_{s^-}^p)]dA_s + \alpha_2^p \mathbb{E}_{\mathcal{F}_0}[H_T^p].$$

Due to $\eta_p$ being concave and non-decreasing and $\mathbb{E}[X_T^p] \leq y_t$, we obtain the inequality:

$$y(t) \leq (1-p) \int_{(0,t]} \eta_p(y(s^-))dA_s + \alpha_2^p \mathbb{E}_{\mathcal{F}_0}[H_T^p].$$
which implies by $\tilde{G}_p(x) = (1-p)G(x^{1/p})$ and $\tilde{G}^{-1}_p(x) = (G^{-1}\left(\frac{1}{1-p}\right))^p$ (see Section 2.1):

$$y(T) \leq \tilde{G}^{-1}_p(\tilde{G}_p(\alpha_0^p \mathbb{E}_{\mathcal{F}_0}[H^p_T]) + (1-p)A_T) = G^{-1}(G(\alpha_2\|H_T\|_{\mathcal{F}_0} + A_T)^p$$

which implies by $\mathbb{P}_{\mathcal{F}_0}[X^*_T > u] \leq \mathbb{E}_{\mathcal{F}_0}[X_{T\wedge \sigma_u} \wedge u]u^{-1}$

$$\|X^*_T\|_{\mathcal{F}_0}^p = \int_0^\infty p\mathbb{P}[X_T > u]u^{p-1}du \leq p \int_0^\infty \mathbb{E}_{\mathcal{F}_0}[X_{T\wedge \sigma_u} \wedge u]u^{p-2}du \leq \frac{1}{1-p}y(T) \leq \frac{1}{1-p}G^{-1}(G(\alpha_2\|H_T\|_{\mathcal{F}_0} + A_T)^p)$$

which implies the assertion.

\[\square\]

### 7.2 Proofs for convex $\eta$ (Theorem 3.6)

**Proof of Theorem 3.6** We assume w.l.o.g. that $H \geq c_0 + \varepsilon$ and $X \geq c_0 + \varepsilon$ on $\Omega$ for some constant $\varepsilon > 0$, see Remark 7.1.

**Proof of a):** Denote by $(Y_t)_{t \geq 0}$ to be the right-hand side of (12). Due to convexity of $\eta$ and $\eta(c_0) = 0$ there exists some $K > 0$ s.t. $\eta(x) \leq K(x - c_0)$ for $x \in [c_0, c]$ where $c$ denotes the constant from the definition of $G$, see (8). This implies $\lim_{x \to c_0} G(x) = -\infty$, in particular, $\text{domain}(G^{-1}) = \text{range}(G) = (-\infty, \lim_{x \to c_0} G(x))$. Moreover, we have $\text{range}(G^1) = \text{domain}(G) = (c_0, \infty)$. Therefore, the following definition is well-defined:

$$f: (c_0, \infty) \times (0, \infty) \mapsto (c_0, \infty), \quad (x, a) \mapsto G^{-1}(G(x) - a).$$

We have for all $x \in (c_0, \infty), a \in [0, \infty)$:

$$\frac{\partial}{\partial x} f(x, a) = \frac{G'(x)}{G'(G^{-1}(G(x) - a))} = \frac{\eta(f(x, a))}{\eta(x)},$$

$$\frac{\partial}{\partial a} f(x, a) = -\eta(f(x, a)),$$

$$\frac{\partial^2}{\partial a^2} f(x, a) = \frac{\eta(f(x, a))}{(\eta(x))^2}(\eta'(f(x, a)) - \eta'(x)).$$

We want to obtain an upper bound on $f(Y, A)$ using Itô’s formula. We first show that the jump term that occurs in the Itô formula is non-positive: By assumption $\eta'$ is non-decreasing. Due to $f(x, a) \leq x$, we have $\frac{\partial^2}{\partial a^2} f(x, a) \leq 0$. This implies that for all fixed $a > 0$ and for all $x, x + \Delta x \in (c_0, \infty)$

$$f(x + \Delta x, a) - f(x, a) - \frac{\partial}{\partial a} f(x, a)\Delta x \leq 0,$$

and therefore the jump term of $f(Y, A)$ in Itô’s formula is non-positive. Therefore, Itô’s formula implies:

$$df(Y_t, A_t) \leq \frac{\partial}{\partial x} f(Y_t, A_t) dY_t + \frac{\partial}{\partial a} f(Y_t, A_t) dA_t + \frac{1}{2} \frac{\partial^2}{\partial a^2} f(Y_t, A_t) [\gamma^c(Y^c, Y^c)]_t$$

$$\leq \eta(f(Y_t, A_t)) \frac{\eta'(Y_t)}{\eta(Y_t)} dA_t + \frac{\eta(f(Y_t, A_t))}{\eta(Y_t)} dH_t + \frac{\eta(f(Y_t, A_t))}{\eta(Y_t)} dM_t - \eta(f(Y_t, A_t)) dA_t$$

$$\leq \tilde{M}_t + dH_t,$$
for some local martingale $\hat{M}$ starting in 0. Note that $\Delta M \geq 0$ implies $\Delta \hat{M} \geq 0$. Due to $f \geq 0$, the family of processes $(f(Y, A))$, the process which is constant 0, $\hat{M}$, $H$) satisfy Assumption $A_{\sup}$ e.g. for $\tilde{\eta}(x) \equiv x$. Therefore, we obtain by Theorem 3.1

$$\|G^{-1}(G(X^*_T) - A_T)\|_p \leq \| \sup_{t \in [0,T]} G^{-1}(G(X_t) - A_t)\|_p$$

$$\leq \| \sup_{t \in [0,T]} G^{-1}(G(Y_t) - A_t)\|_p \leq \begin{cases} a_1 \|H_T\|_1 & \text{if } \mathbb{E}[H_T] < \infty, \\ a_1 \alpha_2 \|H_T\|_p & \text{if } \Delta M \geq 0, \mathbb{E}[H_T^p] < \infty, \\ a_1 \|H_T\|_p & \text{if } H \text{ predictable, } \mathbb{E}[H_T^p] < \infty. \end{cases}$$

Note that if $\eta(x) \equiv x$, then $f(x, a) = xe^{-a}$. Hence, $df(Y_t, A_t)$ can be calculated using Remark 10.2 implying the assertion for non-continuous $A$.

Proof of b): We first sketch the idea of the proof, then we provide the details. Fix some $p \in (0,1)$. We would like to transform inequality (11) of Assumption $A_{\sup}$ to an inequality for $((X^*_t)^p)_{t \in [0,T]}$ which satisfies Assumption $A_{\no sup}$ for some suitable $\tilde{\eta}_p$. Then assertion a) could be applied to obtain an estimate for $((X^*_t)^p)_{t \in [0,T]}$.

Heuristically, this transformation from $((X_t)_{t \in [0,T]}$ to $((X^*_t)^p)_{t \in [0,T]}$ could be achieved e.g. for the case that $\mathbb{E}[H_T] < \infty$ by integrating the following inequality which is given by Lemma 6.3 for $t \in [0,T]$

$$\mathbb{1}_{\{X_t > u\}} u \leq \int_{(0,t]} \eta(x) \mathbb{1}_{\{X_{s-} \leq u\}} dA_s + M^{L,u}_t + \mathbb{E}_{\mathcal{F}_0}[H_T] \wedge u,$$  \hspace{1cm} (32)

w.r.t. to $pu^{p-2}du$ to obtain for $Y_t \in [0,T] : = ((X^*_t)^p)_{t \in [0,T]}$ by the formulas (25) and (recalling $\eta_p(x) := \frac{1}{1-p} \eta(x^{1/p})x^{1-1/p}$)

$$Y_t \leq \int_{(0,t]} \eta_p(Y_{s-}) dA_s + \alpha_2 \mathbb{E}_{\mathcal{F}_0}[H_T]^p \wedge \hat{M}_t,$$  \hspace{1cm} (33)

where $\hat{M}_t := \int_0^t M^{L,u}_s pu^{p-2}du.$

However, it is not clear if e.g. $\hat{M}_t$ is a well-defined local martingale. To avoid technical problems, we change some details of this approach.

Fix some $T > 0$ and some $\hat{u} > \varepsilon$. Since $A$ is predictable, we may assume w.l.o.g. that $A_T$ is bounded by a finite constant. We will derive an inequality similar to that of Assumption $A_{\no sup}$ for $((X^*_t \wedge \hat{u}))_{t \in [0,T]}$. We use similar notation as in Lemma 6.3 for $u > 0$

$$\tau_u := \inf\{s \geq 0 \mid H_s \geq pu\}, \quad \sigma_u := \inf\{s \geq 0 \mid X_s > u\}.$$  

Recall from Lemma 6.3 for $t \in [0,T]$, $\lambda = p$ the definitions

$$H^{L,u}_t := \begin{cases} H(t \wedge \tau_u) + u \mathbb{1}_{\{H_{t} \geq pu\}} & \text{if } H \text{ is predictable and } \mathbb{E}[H_T^p] < \infty, \\ H(t \wedge \tau_u) + u \mathbb{1}_{\{H_{t} \geq pu\}} & \text{if } M \text{ has no negative jumps and } \mathbb{E}[H_T^p] < \infty, \\ \mathbb{E}_{\mathcal{F}_t}[H_T] \wedge u & \text{if } \mathbb{E}[H_T] < \infty, \end{cases}$$

$$M^{L,u}_t := \begin{cases} \lim_{n \to \infty} M^{(\lambda \wedge \sigma_u)}_{t \wedge \lambda \wedge \sigma_u} & \text{if } H \text{ is predictable, } \mathbb{E}[H_T^p] < \infty, \\ M_{t \wedge \tau_u \wedge \sigma_u} & \text{if } \Delta M \geq 0, \mathbb{E}[H_T^p] < \infty, \\ (M_{t \wedge \sigma_u} + \mathbb{E}[H_T | \mathcal{F}_{t \wedge \sigma_u}] - \mathbb{E}_{\mathcal{F}_0}[H_T]) \mathbb{1}_{\{\mathbb{E}_{\mathcal{F}_0}[H_T] \leq u\}} & \text{if } \mathbb{E}[H_T] < \infty, \end{cases}$$
where $\tau_u^{(n)}$ denotes an announcing sequence of $\tau_u$ if $H$ is predictable. It can be seen without further calculations that in the estimate (27) given in Lemma 6.3 we may replace $\tau_u$ by $\bar{\tau}_u$. By rearranging the previous inequality, we obtain that the local supermartingale $(M_t^{L,u})_{t \in [0,T]}$ has lower and upper bounds for $t \in [0,T]$:

$$1_{\{X^*_t > u\}} A_t \leq \int_{\{0,t\}} \eta(X^*_s) 1_{\{X^*_s \leq u\}} \mathrm{d}A_s + M_t^{L,u} \wedge u + H_t^{L,u}. \quad (34)$$

By rearranging the previous inequality, we obtain that the local supermartingale $(M_t^{L,u} \wedge u)_{t \in [0,T]}$ has lower and upper bounds for $t \in [0,T]$:

$$-(1+p)u - \eta(u)A_T \leq M_t^{L,u} \wedge u \leq u,$$

hence in particular, we have:

$$\sup_{t \in [0,T]} \sup_{u \in [0,\bar{u}]} |M_t^{L,u} \wedge u| \leq (1+p)\bar{u} + \eta(\bar{u})A_T.$$

Note that for $r > 0$ it holds $\{\sigma_u < r\} = \{\exists s < r \mid X_s > u\} = \{X^*_s > u\}$. Similarly, due to $H$ being càdlàg and non-decreasing, we have for $r > 0$ that $\{\tau_u \leq r\} = \{H_r \geq pu\}$.

This implies measurability of the following mappings for $\delta \in (0,1)$

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\Omega \times [0,\infty), F_t \otimes B([0,\infty)))$$

$$(\omega, u) \mapsto (\omega, \sigma_u \wedge t),$$

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\Omega \times [0,\infty), F_t \otimes B([0,\infty)))$$

$$(\omega, u) \mapsto (\omega, \tau_u \wedge t),$$

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\Omega \times [0,\infty), F_t \otimes B([0,\infty)))$$

$$(\omega, u) \mapsto (\omega, (\tau_u \wedge t - \delta)_+).$$

Therefore, also

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\mathbb{R}, B(\mathbb{R}))$$

$$(\omega, u) \mapsto \lim_{\delta \to 0} M_t^{\wedge \sigma_u \wedge (\tau_u \wedge t - \delta)_+}(\omega),$$

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\mathbb{R}, B(\mathbb{R}))$$

$$(\omega, u) \mapsto M_t^{\wedge \sigma_u \wedge \tau_u}(\omega),$$

$$(\Omega \times [0,\infty), F_t \otimes B([0,\infty))) \to (\mathbb{R}, B(\mathbb{R}))$$

$$(\omega, u) \mapsto M_t^{\wedge \sigma_u}(\omega)$$

are measurable mappings, where we use the notation $\tilde{M}_t := (M_t^{\wedge \sigma_u} + \mathbb{E}[H_T \mid F_t^{\wedge \sigma_u}]) - \mathbb{E}_{\bar{F}_0}[H_T] \mathbb{1}_{(\bar{F}_0[H_T]) \leq u}$ for all $t \in [0,T]$. Having discussed measurability and boundedness, we may now define for $t \in [0,T]$

$$\tilde{M}_t := \int_{\varepsilon}^{\bar{u}} (M_t^{L,u} \wedge u)pu^{p-2}du.$$

Dominated convergence ensures that $(\tilde{M}_t)_{t \in [0,T]}$ is a bounded càdlàg supermartingale. Now we may integrate (34) w.r.t. $1_{(\varepsilon,\bar{u})}pu^{p-2}du$:

$$(X^*_t \wedge \bar{u})^p - \varepsilon^p = \int_{\varepsilon}^{\bar{u}} 1_{\{X^*_t > u\}} pu^{p-1} du$$

$$\leq \int_{\varepsilon}^{\bar{u}} \int_{\{0,t\}} \eta(X^*_s) 1_{\{X^*_s \leq u\}} \mathrm{d}A_s pu^{p-2} du$$

$$+ \int_{\varepsilon}^{\bar{u}} (M_t^{L,u} \wedge u)pu^{p-2} du + \int_{\varepsilon}^{\bar{u}} H_t^{L,u} pu^{p-2} du$$

$$= \frac{p}{1-p} \int_{\{0,t\}} \eta(X^*_s \wedge \bar{u})((X^*_s \wedge \bar{u})^p - \bar{u}^{p-1}) \mathrm{d}A_s + \tilde{M}_t + \int_{\varepsilon}^{\bar{u}} H_t^{L,u} pu^{p-2} du.$$
We set $Y_t := (X_t^* \wedge \hat{u})^p$, $\hat{H}_t := \int_0^t H^L_{t,u} pu^{p-2} du$ for $t \in [0, T]$. Noting that $\int_0^\varepsilon H^L_{t,u} pu^{p-2} du \geq \varepsilon^p$ we have for $t \in [0, T]$:

$$Y_t \leq \int_{(0,t]} \eta_p(Y_{s^-}) dA_s + \hat{M}_t + \hat{H}_t.$$ 

We define similarly as in the proof of assertion a)

$$f: (0, \infty) \times (0, \infty) \mapsto [0, \infty), \quad (x, a) \mapsto \tilde{G}^{-1}_p(\tilde{G}_p(x) - a),$$

(see (9) for the definition and properties of $\tilde{G}_p$). By nearly identical arguments as in the proof of a), using that the bounded supermartingale $\hat{M}$ has a Doob Meyer decomposition, we obtain by an application of Itô’s formula

$$\mathbb{E}_{\mathcal{F}_0}[\tilde{G}^{-1}_p(\tilde{G}_p(Y_T) - A_T)] \leq \mathbb{E}_{\mathcal{F}_0}[\hat{H}_T].$$

In [Theorem 3.1] we calculated

$$\mathbb{E}_{\mathcal{F}_0}[\hat{H}_T] = \begin{cases} \alpha^p_1 \alpha^p_2 \mathbb{E}_{\mathcal{F}_0}[H^p_T] & \text{if either } H \text{ predictable or } \Delta M \geq 0, \mathbb{E}[H^p_T] < \infty, \\ \alpha^p_1 \mathbb{E}_{\mathcal{F}_0}[H^p_T] & \text{if } \mathbb{E}[H_T] < +\infty. \end{cases}$$

Using the relations between $\tilde{G}_p$ and $G$,

$$\tilde{G}_p(x) = (1 - p)G(x^{1/p}), \quad \tilde{G}^{-1}_p(x) = (G^{-1}(\frac{x}{1-p}))^p,$$

see (10), and taking the limit $\hat{u} \to \infty$ implies by monotone convergence the assertion:

$$\mathbb{E}_{\mathcal{F}_0}[G^{-1}(G(X_t^*) - \beta A_t)] \leq \begin{cases} \alpha^p_1 \alpha^p_2 \mathbb{E}_{\mathcal{F}_0}[H^p_T] & \text{if either } H \text{ predictable or } \Delta M \geq 0, \mathbb{E}[H^p_T] < \infty, \\ \alpha^p_1 \mathbb{E}_{\mathcal{F}_0}[H^p_T] & \text{if } \mathbb{E}[H_T] < +\infty. \end{cases}$$

\[ \square \]

7.3 Proof for general $\eta$ (Theorem 3.9)

Proof. We assume without loss of generality that $M$ is a martingale. Lemma 6.3 implies:

$$\mathbb{E}_{\mathcal{F}_0}[\mathbb{1}_{(X_t^* > u)} u] \leq \mathbb{E}_{\mathcal{F}_0}\left[ \int_{(0,t]} \eta(X_{s^-}^*) \mathbb{1}_{(X_{s^-}^* \leq u)} dA_s \right] + \mathbb{E}_{\mathcal{F}_0}[H_T] \wedge u. \tag{35}$$

We will integrate (35) w.r.t. $d\frac{1}{\eta(u)}$. Note that $\lim_{x \to \infty} \eta(x) = +\infty$. Furthermore, by integration by parts for Lebesgue-Stieltjes integrals, we have (see e.g. [10] Theorem A) and note that $\eta$ is continuous:

$$\int_{(a,b]} ud\frac{1}{\eta(u)} + \int_{(a,b]} \frac{1}{\eta(u)} du = \left. \frac{u}{\eta(u)} \right|_a^b.$$ 

We first calculate the single terms of (35) integrated w.r.t. to $d\frac{1}{\eta(u)}$ over $(\mathbb{E}_{\mathcal{F}_0}[H_T], \infty)$:

$$- \int_{\mathbb{E}_{\mathcal{F}_0}[H_T]} \mathbb{1}_{(X_t^* > u)} ud\frac{1}{\eta(u)} = \int_{\mathbb{E}_{\mathcal{F}_0}[H_T]} \frac{1}{\eta(u)} du - \left. \frac{u}{\eta(u)} \right|_{\mathbb{E}_{\mathcal{F}_0}[H_T]}^{X_t^*} = G(X_t^*) - G(\mathbb{E}_{\mathcal{F}_0}[H_T]) - \left. \frac{u}{\eta(u)} \right|_{\mathbb{E}_{\mathcal{F}_0}[H_T]}^{X_t^*}$$

$$- \eta(X_{t^-}^*) \int_{X_{t^-}^* \vee \mathbb{E}_{\mathcal{F}_0}[H_T]} d\frac{1}{\eta(u)} \leq 1$$

$$- \int_{\mathbb{E}_{\mathcal{F}_0}[H_T]} \mathbb{E}_{\mathcal{F}_0}[H_T] \wedge ud\frac{1}{\eta(u)} = \frac{\mathbb{E}_{\mathcal{F}_0}[H_T]}{\eta(\mathbb{E}_{\mathcal{F}_0}[H_T])}.$$
Combining the calculations, we obtain:

$$
\mathbb{E}_{\mathcal{F}_0}[G(X^*_T)] - G(\mathbb{E}_{\mathcal{F}_0}[H_T]) - \mathbb{E}_{\mathcal{F}_0}\left[\frac{u}{\eta(u)}|X^*_T\right] \leq \mathbb{E}_{\mathcal{F}_0}[A_T] + \frac{\mathbb{E}_{\mathcal{F}_0}[H_T]}{\eta(\mathbb{E}_{\mathcal{F}_0}[H_T])}.
$$

Rearranging the terms implies:

$$
\mathbb{E}_{\mathcal{F}_0}[G(X^*_T)] \leq G(\mathbb{E}_{\mathcal{F}_0}[H_T]) + \mathbb{E}_{\mathcal{F}_0}\left[\frac{X^*_T}{\eta(X^*_T)}\right] + \mathbb{E}_{\mathcal{F}_0}[A_T],
$$

which implies the claim. \(\square\)

8 Proof of the stochastic \(L^\theta\) Gronwall inequality \(\text{(Theorem 3.1)}\)

**Proof of Theorem 3.13** Denote by \(M^L\) the local martingale from Lemma 6.3 and assume w.l.o.g. that \(e^{-\mu A}dM^L\) is a martingale.

**Proof of \(b\)** We first prove the claim for the special case \(\theta = 1\) under the assumption that \(\mathbb{E}[H_T^p] < \infty\) and that either \(H\) is predictable or \(M\) has no negative jumps. Fix some \(u > 0\). Denote by \((Y_t)_{t \geq 0}\) the right-hand side of (27) divided by \(u\), i.e.

$$
Y_t := \frac{1}{u} \int_{(0,t]} X^*_s \mathbb{1}_{\{X^*_s \leq u\}} dA_s + \frac{1}{u} M^L_t + \frac{1}{u} H_t \wedge (\lambda u) + \mathbb{1}_{\{H_t \geq \lambda u\}}.
$$

(36)

Note that \(Y\) and \(M^L\) both depend on \(u\). Choose some \(\mu > 0\). We apply Itô’s formula to calculate \(e^{-\mu A}Y_t\), making use of \(A\) being predictable. Due to Remark 10.2 we obtain:

$$
e^{-\mu A}Y_t \leq Y_0 + \int_{(0,t]} e^{-\mu A_s} dY_s - \int_{(0,t]} \mu e^{-\mu A_s} Y_s dA_s
$$

Noting that by Lemma 6.3 \(\mathbb{1}_{\{X^*_s > u\}} \leq Y_s\) and \(0 \leq e^{-\mu A_s} \leq 1\) for all \(s \geq 0\), yields:

$$
e^{-\mu A} \mathbb{1}_{\{X^*_t > u\}} \leq Y_0 + \int_{(0,t]} e^{-\mu A_s} dY_s - \int_{(0,t]} \mu e^{-\mu A_s} \mathbb{1}_{\{X^*_s > u\}} dA_s
$$

$$
\leq \frac{1}{u} \int_{(0,t]} e^{-\mu A_s} X^*_s \mathbb{1}_{\{X^*_s \leq u\}} dA_s + \frac{1}{u} \int_{(0,t]} e^{-\mu A_s} dM^L_s
$$

$$
+ \frac{1}{u} H_t \wedge (\lambda u) + \mathbb{1}_{\{H_t \geq \lambda u\}} - \int_{(0,t]} \mu e^{-\mu A_s} \mathbb{1}_{\{X^*_s > u\}} dA_s.
$$

We will use the following formulas from Remark 6.2 for positive random variables \(Z\) and \(p \in (0, 1)\):

$$
Z^p = p \int_0^\infty \mathbb{1}_{\{Z > u\}} w^{p-1} \, du = p(1 - p) \int_0^\infty (Z \wedge u) \, w^{p-2} \, du
$$

(37)

$$
= (1 - p) \int_0^\infty Z \mathbb{1}_{\{Z \leq u\}} w^{p-2} \, du.
$$

We take the conditional expectation given \(\mathcal{F}_0\) and integrate w.r.t. \(pw^{p-1}du\). Using the previous formulas for \(Z^p\) gives:
\[\mathbb{E}_F_0[e^{-\mu A_t}(X_t^*)^p] \leq \frac{p}{1-p}\mathbb{E}_F_0[\int_{(0,t]} e^{-\mu A_s}(X_{s+}^*)^p dA_s]
- \mu\mathbb{E}_F_0[\int_{(0,t]} e^{-\mu A_s}(X_{s+}^*)^p dA_s]
+ \frac{\lambda}{1-p}\mathbb{E}_F_0[H_T^p \lambda^{-p}] + \mathbb{E}[H_T^p \lambda^{-p}]. \]

Choosing \(\mu = \frac{p}{1-p}\), \(\lambda = p\) implies:
\[\mathbb{E}_F_0[e^{-p/(1-p)A_T}(X_T^*)^p] \leq \frac{p}{1-p}\mathbb{E}_F_0[H_T^p].\]

Now we prove the assertion for \(\theta = 1\) and \(\mathbb{E}[H_T] < \infty\) by small modifications of the proof of the previous case. Fix some \(u > 0\) and denote by \(Y_t, t \in [0,T]\) the right-hand side of (27) divided by \(u\). By the same steps as before, we obtain:
\[e^{-\mu A_s}1_{\{X_t^*>u\}} \leq Y_0 + \int_{(0,t]} e^{-\mu A_s}dY_s - \int_{(0,t]} \mu e^{-\mu A_s}1_{\{X_t^*>u\}} dA_s \leq \frac{1}{u} \int_{(0,t]} e^{-\mu A_s}1_{\{X_t^*>u\}} dA_s + \frac{1}{u} \mathbb{E}_F_0[H_T] \wedge 1 \]
\[\leq \frac{1}{u} \int_{(0,t]} \mu e^{-\mu A_s}1_{\{X_t^*>u\}} dA_s + \frac{1}{u} \int_{(0,t]} e^{-\mu A_s} dM_s^L. \]

We choose again \(\mu = \frac{p}{1-p}\). We take the conditional expectation given \(\mathcal{F}_0\) and integrate w.r.t. \(p\mu^{-1}du\). This gives using (37):
\[\mathbb{E}_F_0[e^{-\frac{p}{1-p}A_T}(X_T^*)^p] \leq \frac{1}{1-p}\|H_T\|_{1,F_0}^p. \]

Now we prove the assertion for \(\theta < 1\): Set \(Y_t := \int_0^t (X_s^*)^{\theta} dA_s\) and \(\alpha = \theta^{-1}\). Then, by remark [Remark 10.3]
\[(\int_{(0,t]} (X_{s+}^*)^\theta dA_s)^{1/\theta} \leq \theta^{-1} \int_{(0,t]} (\int_{(0,s]} (X_t^* r_0)^\theta dA(r))^{1/\theta-1} (X_{s+}^*)^\theta dA_s \]
\[\leq \theta^{-1} \int_{(0,t]} (X_{s+}^*)^{1-\theta} \int_{(0,s]} A_s^{1/\theta-1} (X_{s+}^*)^\theta dA_s \]
\[= \theta^{-1} \int_{(0,t]} (X_{s+}^*) A_s^{1/\theta-1} dA_s. \]

The assertion follows from the case \(\theta = 1\) by setting \(A_{t}^{(\theta)} := \theta^{-1} \int_{(0,t]} A_s^{1/\theta-1} dA_s\) and insertiong (40) into (13).

**Proof of [a]**: Let \(\mathbb{E}[H_T^p] < \infty\) and assume that \(H\) is predictable or \(\Delta M \geq 0\). The following estimate holds true for any \(u > 0, \theta > 1, \epsilon > 0, t \geq 0\) by Young’s inequality:
\[(\int_{(0,t]} (X_{s+}^*)^\theta 1_{\{X_{s+}^* \leq u\}} dA_s)^{1/\theta} \]
\[\leq c\theta^{-1} \int_{(0,t]} (X_{s+}^*)^\theta 1_{\{X_{s+}^* \leq u\}} dA_s + c^{1/(\theta-1)}(1 - \theta^{-1})(X_{s+}^* (t^-) \wedge u). \]

37
We set
\[ Z_t^{(1)} := 1_{\{X_t^* > u\}} - \frac{1}{u} e^{-1/(\theta - 1)} (1 - \theta^{-1}) (X_t^* \wedge u). \]
\[ Z_t^{(2)} := \frac{1}{u} e^{\theta - 1} \int_{[0,t]} X^*_s 1_{\{X^*_s \leq u\}} dA_s + \frac{1}{u} M_t^L + \frac{1}{u} H_t \wedge (\lambda u) + 1_{\{H_t \geq \lambda u\}}. \]

By Lemma 6.3 and (41) we have
\[ Z_t^{(1)} \leq Z_t^{(2)} \quad \forall t \geq 0. \]

As \( Z_2 \) is non-negative, we cannot apply Remark 10.2. Instead, we need the additional assumption that \( A \) is continuous and apply Itô’s formula. This implies for any constant \( \mu \in \mathbb{R} \):
\[ e^{-\mu A_t} Z_t^{(2)} = Z_0^{(2)} e^{-\mu A_0} + \int_{(0,t]} e^{-\mu A_s} dZ_s^{(2)} - \mu \int_{(0,t]} e^{-\mu A_s} Z_s^{(2)} dA_s. \]

Similar as in proof of Theorem 3.13, we have:
\[ e^{-\mu A_t} Z_t^{(1)} \leq e^{-\mu A_t} Z_t^{(2)} \leq Z_0^{(2)} + \int_{(0,t]} e^{-\mu A_s} dZ_s^{(2)} - \mu \int_{(0,t]} e^{-\mu A_s} Z_s^{(1)} dA_s. \]

We have \( p \int_0^{\infty} \mathbb{E}_{\mathcal{F}_0} \big[ |Z_t^{(1)}|^pu^{\gamma - 1} du \big] = c_{\theta} \mathbb{E}_{\mathcal{F}_0} \big[ |X_t^*|^p \big] \) where \( c_{\theta} := (1 - \frac{1}{1-p} e^{-1/(\theta - 1)} (1 - \theta^{-1})). \)

Hence, taking the conditional expectation given \( \mathcal{F}_0 \), integrating w.r.t. \( pu^{\gamma - 1} du \) and a very similar calculation as in the proof of the case \( \theta = 1 \) (see (38)) imply:
\[ c_{\theta} \mathbb{E}_{\mathcal{F}_0} \big[ e^{-\mu A_t} (X_t^*)^p \big] \leq \frac{p}{1 - p} \mathbb{E}_{\mathcal{F}_0} \bigg[ \int_{(0,t]} e^{-\mu A_s} (X_s^*)^p dA_s \bigg] - \mu c_{\theta} \mathbb{E}_{\mathcal{F}_0} \bigg[ \int_{(0,t]} e^{-\mu A_s} (X_s^*)^p dA_s \bigg] + \frac{p - p}{1 - p} \mathbb{E}[H_t^p]. \]

Hence, choosing \( \mu = \frac{p \theta - 1}{(1 - p) c_{\theta}} \) implies:
\[ c_{\theta} \mathbb{E}_{\mathcal{F}_0} \big[ \exp \{- \frac{p \theta - 1}{(1 - p) c_{\theta}} A_t\} (X_t^*)^p \big] \leq \frac{p - p}{1 - p} \mathbb{E}[H_t^p]. \]

Choosing \( c = (1 - p)^{-2(\theta - 1)} \) implies \( c_{\theta} = \theta^{-1} \), and therefore:
\[ \theta^{-1} \mathbb{E}_{\mathcal{F}_0} \big[ \exp \{- \frac{p}{(1 - p) \theta} A_t\} (X_t^*)^p \big] \leq \frac{p - p}{1 - p} \mathbb{E}[H_t^p]. \]

Now we assume instead, that \( \mathbb{E}[H_T] < \infty \). By redefining for \( t \in [0,T] \):
\[ Z_t^{(2)} := \frac{1}{u} e^{\theta - 1} \int_{[0,t]} X^*_s 1_{\{X^*_s \leq u\}} dA_s + \frac{1}{u} M_t^L + \left( u^{-1} \mathbb{E}_{\mathcal{F}_0} [H_T] \right) \wedge 1. \]

and the same arguments as before, we obtain for \( t \in [0,T] \):
\[ \theta^{-1} \mathbb{E}_{\mathcal{F}_0} \big[ \exp \{- \frac{p}{(1 - p) \theta} A_t\} (X_t^*)^p \big] \leq \frac{1}{1 - p} \mathbb{E}_{\mathcal{F}_0} [H_T]^p. \]
**Remark 10.2** implies (by the same arguments as in the proof of a)) for predictable \( E \). Similar arguments imply in the case \( \Delta M \geq 0 \), implies for \( \mu = p \), setting \( \theta > 1 \) and setting \( \mu = p \) and \( c = 1 \) and setting \( \theta > 1 \):

\[
\tilde{Z}_t^{(1)} := \frac{1}{u}X_t^* \wedge u - \frac{1}{u}(1-\theta^{-1})(X_t^* \wedge u) = \theta^{-1} \frac{1}{u}X_t^* \wedge u.
\]

Since \( \tilde{Z}_t \) is non-negative, we do not need the continuity assumption of \( A \). Applying [Remark 10.2](#) implies (by the same arguments as in the proof of a)) for predictable \( H \) or \( \Delta M \geq 0 \):

\[
\theta^{-1}EX_0[\exp\{pA_T\}(X_T^*)^p] \leq p^{-p}EX_0[H_T^p].
\]

Similar arguments imply in the case \( E[H_T] < \infty \) for all \( t \in [0, T] \):

\[
\theta^{-1}EX_0[\exp\{pA_T\}(X_T^*)^p] \leq EX_0[H_T^p].
\]

\( \square \)

## 9 Proof of sharpness

### 9.1 Proof of the sharpness of \( \beta \) (Theorem 3.15)

**Proof of Theorem 3.15.** We prove the assertion by defining families of processes \( \{X_{t}^{\varepsilon, \delta}, \varepsilon, \delta \in (0, 1)\} \) and \( \{M_{t}^{\varepsilon, \delta}, \varepsilon, \delta \in (0, 1)\} \) satisfying the assumptions of Theorem 3.15 such that

\[
\beta \geq \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \lim_{t \to \infty} \frac{1}{t} \log(\alpha^{-1}H^{-1}\|(X_{t}^{\varepsilon, \delta})^*\|_{p}) \geq \frac{1}{1-p} = \beta.
\]

We use the notation

\[
X_{t}^{\varepsilon, \delta, p} := \sup_{0 \leq s \leq t} X_{s}^{p} \quad \forall t \geq 0,
\]

for \( p \in (0, 1) \) and non-negative continuous processes \( X \).

**Step 1:** General construction procedure

We fix some \( \varepsilon, \delta \in (0, 1) \). We define \( \{X_{t}^{\varepsilon, \delta}, t \geq 0\} \) by choosing a suitable continuous local martingale \( M_{t}^{\varepsilon, \delta} = \{M_{t}^{\varepsilon, \delta}, t \geq 0\} \) and defining for all \( t \geq 0 \):

\[
X_{t}^{\varepsilon, \delta} := \int_{0}^{t} (X_{s}^{\varepsilon, \delta})^* g_{\varepsilon, \delta}(s) ds + M_{t}^{\varepsilon, \delta} + 1
\]

\[
\leq \int_{0}^{t} (X_{s}^{\varepsilon, \delta})^* ds + M_{t}^{\varepsilon, \delta} + 1,
\]

39
where \( g_{\varepsilon, \delta}(s) := \sum_{k=0}^{\infty} \mathbb{1}_{[(k+\varepsilon)\delta, (k+1)\delta]}(s) \) is a function which only takes values in \{0, 1\}. The local martingale \( M^{\varepsilon, \delta} \) will be chosen such that the non-negativity of \( X^{\varepsilon, \delta} \) is ensured.

**Step 2: Inductive construction of \( M^{\varepsilon, \delta} \)**

We define \( M^{\varepsilon, \delta} \) using a family of sped up Brownian motions: Define the following increasing homeomorphism:

\[
h : [0, \varepsilon, \delta) \to [0, \infty), \quad t \mapsto \log \left( \frac{2}{1 - (\varepsilon \delta)^{-1} t - 1} \right).
\]

Let \( \{(B^i_t)_{t \geq 0}, i \in \mathbb{N}_0\} \) be a family of independent Brownian motions on some suitable underlying probability space. We call \( \{B^i_{h(t)}, t \in [0, \varepsilon \delta)\} \) a sped up Brownian motion. We define the filtration for the local martingale \( M^{\varepsilon, \delta} \) as follows: We set \( \mathcal{F}_0 = \{0, \Omega\} \) and

\[
\tilde{\mathcal{F}}_t := \tilde{\mathcal{F}}_{k\delta} \vee \sigma(\{B^k_{h(r)} | r \in [0, (t - k\delta)] \cap [0, \delta \varepsilon)\}) \quad \forall t \in (k\delta, (k + 1)\delta], \quad k \in \mathbb{N}_0.
\]

We will denote by \( (\mathcal{F}_t)_{t \geq 0} \) the smallest augmented filtration containing \( (\tilde{\mathcal{F}}_t)_{t \geq 0} \). We choose \( (\mathcal{F}_t)_{t \geq 0} \) as our filtration for \( M^{\varepsilon, \delta} \). We define the local martingale \( M^{\varepsilon, \delta} \) inductively on \([0, \delta], (\delta, 2\delta], (2\delta, 3\delta], \ldots \) as follows.

**Step 2a: Definition of \( M^{\varepsilon, \delta}_t \) for \( t \in [0, \delta] \):** On \([0, \delta] \) we have \( g_{\varepsilon, \delta} = 0 \). We define \( M^{\varepsilon, \delta}_t \) such that \( X^{\varepsilon, \delta} \) remains non-negative. We choose:

\[
\tau_0 := \inf\{t \in [0, \varepsilon\delta) \mid B^0_{h(t)} + 1 = 0\},
\]

\[
M^{\varepsilon, \delta}_t := B^0_{h(t \wedge \tau_0)} \quad \text{on } \{\tau_0 < \varepsilon\delta\} \quad \forall t \in [0, \delta],
\]

where we set \( \inf \emptyset := 0 \), noting that this only happens on a set of measure zero and \( (\mathcal{F}_t)_{t \geq 0} \) is complete. Hence, we have defined \( M^{\varepsilon, \delta}_t, t \in [0, \delta] \). Note, that \( \tau_0 < \varepsilon\delta \), hence \( M^{\varepsilon, \delta}_t \) is constant on \( t \in [\varepsilon\delta, \delta] \). By construction, the following equalities hold \( \mathbb{P}\text{-a.s.}: \)

\[
X^{\varepsilon, \delta}_{\tau_0} = 0, \quad X^{\varepsilon, \delta}_t = \alpha(X^{\varepsilon, \delta}_{\tau_0})^*,
\]

where \( \alpha := (1 - \varepsilon)\delta < 1 \).

**Step 2b: Definition of \( M^{\varepsilon, \delta}_t \) for \( t \in (k\delta, (k + 1)\delta] \):** Now assume we have defined \( M^{\varepsilon, \delta} \) on \([0, k\delta] \) for some \( k \in \mathbb{N} \) and define it in \((k\delta, (k + 1)\delta]\). By definition, we have \( g_{\varepsilon, \delta} = 0 \) on \([k\delta, k\delta + \varepsilon\delta] \). Define:

\[
\tau_k := \inf\{t \in [k\delta, k\delta + \varepsilon\delta) \mid B^k_{h(t - k\delta)} + X^{\varepsilon, \delta}_{k\delta} = 0\},
\]

\[
M^{\varepsilon, \delta}_t := B^k_{h(t \wedge \tau_k - k\delta)} + M^{\varepsilon, \delta}_{k\delta} \quad \forall t \in (k\delta, (k + 1)\delta],
\]

setting \( \inf \emptyset := k\delta \), noting that this only happens on a set of measure 0. Again, \( M^{\varepsilon, \delta} \) is by construction constant on \([k + \varepsilon)\delta, (k + 1), \delta \] and we have \( \mathbb{P}\text{-a.s.}: \)

\[
X^{\varepsilon, \delta}_{\tau_k} = 0, \quad X^{\varepsilon, \delta}_{(k+1)\delta} = \alpha(X^{\varepsilon, \delta}_{\tau_k})^*.
\]

By definition, \( X^{\varepsilon, \delta}_t \) is non-negative and satisfies

\[
X^{\varepsilon, \delta}_t \leq \int_0^t (X^{\varepsilon, \delta}_s)^* \, ds + M^{\varepsilon, \delta}_t + 1
\]
and $M^{\varepsilon,\delta}$ is a local martingale with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$. The localizing sequence can be constructed as follows: Let $\tilde{\sigma}^{(n)}_k, n \in \mathbb{N}$ be a localizing sequence of $(B^k_{h((t-\kappa\delta)_+)})_{t \geq 0}$ taking values in $[k\delta, k\delta + \varepsilon\delta)$. Set
\[
\sigma^{(n)}_k = \begin{cases} 
\tilde{\sigma}^{(n)}_k & \text{on } \{\tilde{\sigma}^{(n)}_k < \tau_k\} \\
\infty & \text{else,} 
\end{cases}
\]
which is a localizing sequence of $(B^k_{h((t-\kappa\tau_k-\kappa\delta)_+)})_{t \geq 0}$. Then $\sigma^{(n)} := (\min_{k \in \mathbb{N}_0} \sigma^{(n)}_k) \land (n\delta)$ is a localizing sequence of $M^{\varepsilon,\delta}$.

It remains to find a good lower bound of $\|(X^{\varepsilon,\delta})^*\|_p$.

**Step 3:** Proof of $\|(X^{\varepsilon,\delta})^*\|_p \geq \frac{1-p+p\alpha}{1-p} \|(X^{\varepsilon,\delta})_{\tau_k-1}^*\|_p$ for all $p \in (0, 1), k \in \mathbb{N}$: Define for some fixed $c > 0$:
\[
\begin{align*}
\sigma_1 & := \inf\{t \geq 0 \mid (B^k_t + X^\varepsilon_{\tau_k})^p = (X^\varepsilon_{\tau_k-1})^p + c\}, \\
\sigma_2 & := \inf\{t \geq 0 \mid B^k_t + X^\varepsilon_{\tau_k} = 0\}.
\end{align*}
\]
Denote by $\mathcal{F}_k$ the $\sigma$-algebra generated by the set $\{B^k_t, i = 0, ..., k, t \geq 0\}$. Due to $X^\varepsilon_{\tau_k} = \alpha(X^\varepsilon_{\tau_k})^*$ we have $((X^\varepsilon_{\tau_k-1})^* + c)^{1/p} - X^\varepsilon_{\tau_k} \geq 0$. Moreover, we have
\[
\{(X^\varepsilon_{\tau_k})^* \geq (X^\varepsilon_{\tau_k-1})^* + c\} = \{\sigma_1 < \sigma_2\},
\]
and if $(X^\varepsilon_{\tau_k})^* > (X^\varepsilon_{\tau_k-1})^*$, then the supremum of $X^\varepsilon$ on $[0, \tau_k]$ must occur in $[k\delta, \tau_k]$, since $X^\varepsilon$ takes the values in $[0, \alpha(X^\varepsilon_{\tau_k-1})^*]$ and $\alpha < 1$. Therefore, by the independence of the Brownian motions $B_k, k \in \mathbb{N}_0$:
\[
P[(X^\varepsilon_{\tau_k})^* \geq (X^\varepsilon_{\tau_k-1})^* + c \mid \mathcal{F}_{k-1}] = P[\sigma_1 < \sigma_2 \mid \mathcal{F}_{k-1}] = \frac{X^\varepsilon_{\tau_k}}{((X^\varepsilon_{\tau_k-1})^* + c)^{1/p}},
\]
as $P[\sigma_1 < \sigma_2 \mid \mathcal{F}_{k-1}]$ is the conditional probability that $B^k$ hits $((X^\varepsilon_{\tau_k-1})^* + c)^{1/p} - X^\varepsilon_{\tau_k}$ before $X^\varepsilon_{\tau_k}$. Applying the previous equation gives:
\[
\begin{align*}
E[(X^\varepsilon_{\tau_k})^*]\ &= \int_0^\infty P[(X^\varepsilon_{\tau_k-1})^* \geq u] \, du \\
&= E\left[(X^\varepsilon_{\tau_k-1})^* + \int_0^\infty P[(X^\varepsilon_{\tau_k-1})^* \geq u] \, du\right] \\
&= E\left[(X^\varepsilon_{\tau_k-1})^* + \int_0^\infty P[(X^\varepsilon_{\tau_k})^* \geq (X^\varepsilon_{\tau_k-1})^* + c \mid \mathcal{F}_{k-1}] \, dc\right] \\
&= E[(X^\varepsilon_{\tau_k})^*] + \int_0^\infty \frac{X^\varepsilon_{\tau_k}}{((X^\varepsilon_{\tau_k-1})^* + c)^{1/p}} \, dc \\
&= E[(X^\varepsilon_{\tau_k})^*] + \frac{p}{1-p} E[(X^\varepsilon_{\tau_k-1})^* - 1 + p].
\end{align*}
\]
Applying, that $X^\varepsilon = \alpha(X^\varepsilon_{\tau_k})^*$ implies
\[
E[(X^\varepsilon_{\tau_k})^*] \geq \frac{1-p+p\alpha}{1-p} E[(X^\varepsilon_{\tau_k})^*].
\]

Step 4: Lower bound of $\| (X_{\epsilon,\delta}^X)^{*} \|_p$ for $p \in (0, 1)$: Noting that $\mathbb{E}[(X_{\tau_0}^\epsilon)^{*} p] = \frac{1}{1-p}$ and iterating (16) yields:

$$\mathbb{E}[(X_{\epsilon,\delta}^X)^{*} p] \geq \frac{1}{1-p} \left( \frac{1-p + p\alpha}{1-p} \right)^k = \frac{1}{1-p} \exp \left\{ k \delta \log \left( \frac{1-p + p\alpha}{1-p} \right)^{1/\delta} \right\}.$$ 

(47)

By assumption, we have (due to $\| X_t^* \|_p \leq \tilde{\alpha} H \exp(\tilde{\beta} t)$ and $H := 1$) for all $k \in \mathbb{N}$:

$$\tilde{\beta} \geq (k\delta)^{-1} \log(\| (X_{k\delta}^\epsilon)^* \|_p) - (k\delta)^{-1} \log(\tilde{\alpha}).$$

Hence, for all $k \in \mathbb{N}$:

$$\tilde{\beta} \geq p^{-1} \log \left\{ \left( \frac{1-p + p\alpha}{1-p} \right)^{1/\delta} \right\} - (k\delta)^{-1} p^{-1} \log(1-p) - (k\delta)^{-1} \log(\tilde{\alpha}),$$

i.e.

$$\tilde{\beta} \geq \lim_{\delta \to 0} \lim_{\epsilon \to 0} \frac{1}{p\delta} \log \left( \frac{1-p + p\delta(1-\epsilon)}{1-p} \right) = \frac{\partial}{\partial x} \log(1 + \frac{1-p}{1-p} x) \bigg|_{x=0} = \frac{1}{1-p} = \beta$$

which implies the assertion.

9.2 Proof of the sharpness of $\alpha_1$ and $\alpha_1 \alpha_2$ (Theorem 3.17)

**Sketch of proof of** Theorem 3.17. **The inequalities follow e.g. from Theorem 3.1 by choosing** $A = 0$. **Here we only discuss the sharpness of the constants:**

Proof of $[\text{b}]$ Let $B$ be a Brownian motion on a suitable underlying filtered probability space. Choose $H_t \equiv 1$. Let $\tau$ be the time $B$ first hits $-1$ and set $M_t := B_{t \wedge \tau}$ and $X_t := M_t + H_t$ for all $t \geq 0$. The stopping times $\tau$ ensures $X \geq 0$. An easy calculation gives $\| \sup_{t \geq 0} X_t \|_p = \alpha_1$, which implies the assertion of $[\text{b}]$.

Proof of $[\text{a}]$. Fix some $p \in (0, 1)$. In the proof of [9, Theorem 2.1]) families of continuous processes $X^{(n)}, n \in \mathbb{N}$ and $H^{(n)}, n \in \mathbb{N}$ are defined, satisfying:

(i) $X^{(n)}$ is non-negative, adapted, continuous,

(ii) $H^{(n)}$ is non-negative, adapted, continuous, non-decreasing,

(iii) $\mathbb{E}[X^{(n)}] \leq \mathbb{E}[H^{(n)}]$ for all bounded stopping times $\tau$,

(iv) and

$$\alpha_1 \alpha_2 = \lim_{n \to \infty} \frac{\sup_{t \geq 0} X^{(n)}_t}{\sup_{t \geq 0} H^{(n)}_t}.$$ 

To prove assertion $[\text{a}]$ it remains to show the existence of a family of local martingales with non-negative jumps $M^{(n)}, n \in \mathbb{N}$ starting in 0 such that $X^{(n)}_t \leq H^{(n)}_t + M^{(n)}_t$ for all $t \geq 0, n \in \mathbb{N}$. 

42
To this end, we first shortly recall the definition of $X^{(n)}$ and $H^{(n)}$ from [10, Theorem 2.1]): Let $Z$ be an exponentially distributed random variable on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[Z] = 1$. Set
$$a : [0, \infty) \rightarrow [0, \infty), \quad t \mapsto \exp(t/p).$$
Define for all $t \geq 0$
$$\tilde{X}_t := a(Z)1_{(Z, \infty)}(t), \quad \tilde{H}_t := \int_0^{t \wedge \tau} a(s)ds.$$
Choose $\tilde{\tau}_t := \sigma(\{Z \leq r\} \mid 0 \leq r \leq t)$ for all $t \geq 0$. The compensator of $\tilde{X}$ is $\tilde{H}$ due to $Z$ being exponentially distributed. Now we use $\tilde{X}$ and $\tilde{H}$ to construct the families of processes $X^{(n)}, n \in \mathbb{N}$ and $H^{(n)}, n \in \mathbb{N}$. Assume w.l.o.g. that there exists a Brownian motion $B$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $B$ is independent of $Z$. Denote by $(\mathcal{F}_t)_{t \geq 0}$ the smallest filtration satisfying the usual conditions which contains $(\mathcal{F}_t)$ and with respect to which $B$ is a Brownian motion. Denote by $g_{n+1} : [0, \infty) \rightarrow [0, 1]$ a continuous non-decreasing function such that
$$g_{n,n+1}(t) = 0 \quad \forall t \in [0, n], \text{ and } g_{n,n+1}(t) = 1 \quad \forall t \in [n + 1, \infty).$$
Define:
$$\tau^{(n)} := \inf\{t \geq n + 1 \mid \tilde{X}_n + (B_t - B_{n+1})1_{(t \geq n+1)} = 0\},$$
$$X^{(n)}_t := g_{n,n+1}(t)\tilde{X}_n + (B_{t \wedge \tau^{(n)}} - B_{t \wedge (n+1)}),$$
$$H^{(n)}_t := \tilde{H}_{t \wedge n}.$$The stopping time $\tau^{(n)}$ ensures that $X^{(n)}$ is non-negative. Define
$$M^{(n)}_t := \tilde{X}_{t \wedge n} - \tilde{H}_{t \wedge n} + B_{t \wedge \tau^{(n)}} - B_{t \wedge (n+1)},$$
which is a local martingale (recall that $\tilde{H}$ is the compensator of $\tilde{X}$). Moreover,
$$X^{(n)}_t \leq \tilde{X}_{t \wedge n} + (B_{t \wedge \tau^{(n)}} - B_{t \wedge (n+1)}) \leq H^{(n)}_t + M^{(n)}_t.$$It is easily seen that $M$ only has non-negative jumps, as $\tilde{X}$ only has non-negative jumps.

10 Appendix

10.1 Proof of the (deterministic) Bihari-LaSalle inequality

Proof of Lemma 1.1. Denote by $y(t)$ the right-hand side of (6), which is a (non-decreasing) function of finite variation. One can apply for example Ito’s formula to calculate the deterministic function $G(y(t))$, noting that $G$ only needs to be once continuously differentiable, because $y$ is a process of bounded variation. We obtain, noting that the jump terms of Ito’s formula are non-positive in this case, that:
$$G(y(t)) - G(y(0)) = \int_{[0,t]} G'(y(s^-))dy(s) + \text{Jump terms of Ito’s formula} \leq \int_{[0,t]} \frac{dy(s)}{\eta(y(s^-))} \leq \int_{[0,t]} \frac{dy(s)}{\eta(x(s^-))} = A(t).$$Rearranging the terms, noting that $y(0) = H$ and applying $G^{-1}$ (if possible) implies the assertion.
10.2 A counterexample for convex \( \eta \)

We provide a counterexample why bounds of the type (15) are not possible for general \( \eta \).

**Counterexample 10.1.** Let \((W_t)_{t \geq 0}\) be a Brownian motion on a suitable underlying filtered probability space and let \(x_0 > 0\) and \(\gamma > 0\) be constants. Define:

\[
X_t := e^{\gamma(x_0 + W_t)^2} \quad \forall t \geq 0.
\]

An application of Itô’s formula implies that \(X\) satisfies the following equation:

\[
dX_t = 2\gamma(x_0 + W_t)X_t dW_t + (\gamma + 2\gamma^2(x_0 + W_t)^2)X_t dt
\]

\[
= \eta(X_t) dt + dM_t
\]

where \(dM_t := 2\gamma(x_0 + W_t)X_t dW_t, \ t \geq 0\) is a local martingale starting in 0, and \(\eta(x) := \gamma(1 + 2\log(x))x \) for \(x \geq 1\). Note, that \(X_t \geq 1\). Extend \(\eta\) to a convex non-decreasing function on \([0, \infty)\) with \(\eta(0) = 0\).

Hence, \(X\) as defined above satisfies

\[
X_t \leq \int_0^t \eta(X_s) ds + M_s + e^{\gamma x_0^2}, \quad \forall t \geq 0.
\]

Furthermore, we have

\[
\|X_t^p\|_p \in \begin{cases} [0, \infty) & \text{for } p\gamma T < \frac{1}{2} \\ \{\infty\} & \text{for } p\gamma T \geq \frac{1}{2} \end{cases}, \quad (48)
\]

which can be seen as follows: For the case \(p\gamma T > 1/2\), we have:

\[
\|X_t^p\|_p \geq \mathbb{E}[X_t^p] = \int_{-\infty}^{\infty} \exp(p\gamma(x_0 + \sqrt{t}w)^2) \exp(-w^2/2) dw
\]

\[
= \int_{-\infty}^{\infty} \exp((p\gamma T - 1/2)w^2 + 2x_0p\gamma \sqrt{t}w + p\gamma x_0^2) dw
\]

which is not finite for \(p\gamma T \geq \frac{1}{2}\). For the case \(p\gamma T < 1/2\) we use the estimate \(\|X_t^p\|_{\tilde{q}} \leq (1 - \tilde{q})^{1/\tilde{q}} \mathbb{E}[e^{\gamma x_0^2} + \int_0^T \eta(X_s) ds]\) (by e.g. [Theorem 3.1]). We have:

\[
\mathbb{E}[\eta(X_t)] = \int_{-\infty}^{\infty} \gamma(1 + 2\gamma(x_0 + \sqrt{t}w)^2) \exp(\gamma(x_0 + \sqrt{t}w)^2) \exp(-w^2/2) dw
\]

\[
\leq \int_{-\infty}^{\infty} \gamma(1 + 2\gamma(x_0 + \sqrt{t}w)^2) \exp((\gamma T - 1/2)w^2 + 2\gamma x_0 \sqrt{t}|w| + \gamma x_0^2) dw
\]

which is finite when \(\gamma T < \frac{1}{2}\). Hence, we have

\[
\mathbb{E}[\sup_{t \in [0, T]} \exp(\gamma q(x_0 + W_t)^2)] < \infty \text{ for all } q \in (0, 1), \gamma T < 1/2,
\]

which implies

\[
\mathbb{E}[\sup_{t \in [0, T]} \exp(\gamma p(x_0 + W_t)^2)] < \infty \text{ for all } pT < 1/2.
\]

Hence, we have shown (48).

Furthermore, we have for all \(x \geq 1, y \geq 0\):

\[
G(x) := \int_1^x \frac{dw}{\eta(w)} = \frac{1}{2\gamma} \log(2\log(x) + 1), \quad G^{-1}(y) = e^{2\gamma y/2 - 1/2}.
\]

As \(G\) and \(G^{-1}\) are bounded on bounded subintervals of \([1, \infty)\), (48) is a contradiction to \(X\) having an upper bound of the type (15), as (15) would imply that \(\|X_t^p\|_p\) remains finite for all \(t \geq 0\) and any fixed \(p \in (0, 1)\).
10.3 Two remarks on jump terms

In some of the proofs we use Itô’s formula. If the integrator $A$ has jumps, then jump terms occur in the Itô’s formulas we use. The next two remarks analyse the jump terms we encounter. Recall that $A$ is always assumed to be predictable.

**Remark 10.2** (Analysis of the jump terms in Itô’s formula). Let $Z$ be a non-negative semimartingale and $A$ a predictable non-decreasing process of finite variation. Then

$$e^{-A_t} Z_t \leq Z_0 e^{-A_0} + \int_{(0,t]} e^{-A_s} dZ_s - \int_{(0,t]} e^{-A_s} Z_{s^-} dA_s$$

holds true.

This can be seen as follows. Note that since $Z$ is a semimartingale and $e^{-A}$ is predictable and a process of finite variation, we have (see e.g. [14, Proposition I.4.49 (a) and (b)]):

$$\int_{(0,t]} \Delta e^{-A_s} dZ_s = [Z, e^{-A}]_t = \int_{(0,t]} \Delta Z_s e^{-A_s} = \sum_{s \leq t} \Delta Z_s \Delta e^{-A_s}.$$ 

Hence, due to $A$ being predictable, we may rearrange Itô’s formula as follows:

$$e^{-A_t} Z_t = Z_0 e^{-A_0} + \int_{(0,t]} e^{-A_s} dZ_s - \int_{(0,t]} e^{-A_s} Z_{s^-} dA_s$$

$$- \int_{(0,t]} \Delta e^{-A_s} dZ_s + \int_{(0,t]} Z_{s^-} \Delta e^{-A_s} dA_s$$

$$+ \sum_{s \leq t} \left( e^{-A_s} Z_s - e^{-A_s} Z_{s^-} - e^{-A_{s^-}} \Delta Z_s + e^{-A_{s^-}} Z_{s^-} \Delta A_s \right)$$

$$= Z_0 e^{-A_0} + \int_{(0,t]} e^{-A_s} dZ_s - \int_{(0,t]} e^{-A_s} Z_{s^-} dA_s$$

$$+ \sum_{s \leq t} Z_{s^-} e^{-A_s} (1 + \Delta A_s - e^{\Delta A_s})$$

Noting that the last term is non-positive since $\Delta A_s \geq 0$ for all $s > 0$, the claim follows.

**Remark 10.3** (Analysis of the jump terms in Itô’s formula). Let $\alpha > 1$ and let $Y$ be a predictable, non-decreasing process. By rearranging Itô’s formula (noting that we only need $x \mapsto x^\alpha$ to be $C^1$ since $Y$ is of finite variation), we obtain:

$$Y_s^\alpha = Y_0^\alpha + \alpha \int_{(0,t]} Y_s^{\alpha - 1} dY_s + \sum_{s \leq t} \left( \alpha (Y_s^{\alpha - 1} - Y_{s^-}^{\alpha - 1}) \Delta Y_s \right.$$  

$$+ Y_s^\alpha - Y_{s^-}^\alpha - \alpha Y_{s^-}^{\alpha - 1} \Delta Y_s \right)$$

$$= Y_0^\alpha + \alpha \int_{(0,t]} Y_s^{\alpha - 1} dY_s + \sum_{s \leq t} \left( Y_s^\alpha - Y_{s^-}^\alpha - \alpha Y_{s^-}^{\alpha - 1} \Delta Y_s \right)$$

$$\leq Y_0^\alpha + \alpha \int_{(0,t]} Y_s^{\alpha - 1} dY_s.$$ 

**Acknowledgements**

The author wishes to thank Michael Scheutzow for his valuable suggestions and comments, in particular for his idea on how to prove the sharpness of the constant $\beta$ in [Theorem 3.15].
References

[1] David Applebaum, *Levy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009. [MR2512800]

[2] Hussein K. Asker, *Well-posedness and exponential estimates for the solutions to neutral stochastic functional differential equations with infinite delay*, Journal of Systems Science and Information 8 (2020), no. 5, 434–446.

[3] Stefan Bachmann, *Well-posedness and stability for a class of stochastic delay differential equations with singular drift*, Stoch. Dyn. 18 (2018), no. 2, 1850019, 27. [MR3735415]

[4] Stefan Bachmann, *On the strong Feller property for stochastic delay differential equations with singular drift*, Stochastic Process. Appl. 130 (2020), no. 8, 4563–4592. [MR4108463]

[5] Jianhai Bao, Feng-Yu Wang, and Chenggui Yuan, *Asymptotic log-Harnack inequality and applications for stochastic systems of infinite memory*, Stochastic Process. Appl. 129 (2019), no. 11, 4576–4596. [MR4013873]

[6] Imre Bihari, *A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations*, Acta Math. Acad. Sci. Hungar. 7 (1956), 81–94. [MR79154]

[7] Sonja Cox, Martin Hutzenthaler, and Arnulf Jentzen, *Local lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations*, 2021, arXiv:1309.5595.

[8] Xiaoming Fu, *On invariant measures and the asymptotic behavior of a stochastic delayed SIRS epidemic model*, Phys. A 523 (2019), 1008–1023. [MR3942772]

[9] Sarah Geiss and Michael Scheutzow, *Sharpness of Lenglart’s domination inequality and a sharp monotone version*, Electron. Commun. Probab. 26 (2021), Paper No. 44, 8. [MR4284625]

[10] Edwin Hewitt, *Integration by parts for stieltjes integrals*, The American Mathematical Monthly 67 (1960), no. 5, 419–423.

[11] Xing Huang and Feng-Yu Wang, *Distribution dependent SDEs with singular coefficients*, Stochastic Process. Appl. 129 (2019), no. 11, 4747–4770. [MR4013879]

[12] Xing Huang and Feng-Yu Wang, *McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance*, Discrete Contin. Dyn. Syst. 41 (2021), no. 4, 1667–1679. [MR4211198]

[13] Anselm Hudde, Martin Hutzenthaler, and Sara Mazzonetto, *A stochastic Gronwall inequality and applications to moments, strong completeness, strong local Lipschitz continuity, and perturbations*, Ann. Inst. Henri Poincaré Probab. Stat. 57 (2021), no. 2, 603–626. [MR4260476]
[14] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR1943877

[15] Theresa Lange and Wilhelm Stannat, *Mean field limit of ensemble square root filters - discrete and continuous time*, Foundations of Data Science 3 (2021), no. 3, 563–588.

[16] Joseph Pierre LaSalle, *Uniqueness theorems and successive approximations*, Ann. of Math. (2) 50 (1949), 722–730. MR31165

[17] Érik Lenglart, *Relation de domination entre deux processus*, Ann. Inst. H. Poincaré Sect. B (N.S.) 13 (1977), 171–179. MR0471069

[18] Chengcheng Ling and Longjie Xie, *Strong solutions of stochastic differential equations with coefficients in mixed-norm spaces*, Potential Analysis (2021), 1–15.

[19] Wei Liu and Michael Röckner, *Stochastic partial differential equations: an introduction*, Universitext, Springer, Cham, 2015. MR3410409

[20] Khoa Lê and Chengcheng Ling, *Taming singular stochastic differential equations: A numerical method*, 2021, arXiv:2110.01343

[21] Cloud Makasu, *A stochastic Gronwall lemma revisited*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 22 (2019), no. 1, 1950007, 5. MR3942499

[22] Cloud Makasu, *Extension of a stochastic Gronwall lemma*, Bull. Pol. Acad. Sci. Math. 68 (2020), no. 1, 97–104. MR4165506

[23] Oliver Matte, *Continuity properties of the semi-group and its integral kernel in non-relativistic QED*, Rev. Math. Phys. 28 (2016), no. 5, 1650011, 90. MR3522448

[24] Sima Mehri and Michael Scheutzow, *A stochastic Gronwall lemma and well-posedness of path-dependent SDEs driven by martingale noise*, ALEA, Lat. Am. J. Probab. Math. Stat. 18 (2021), no. 4198874, 193–209.

[25] Slimane Mekki, Juan J. Nieto, and Abdelghani Ouahab, *Stochastic version of Henry type Gronwall’s inequality*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 24 (2021), no. 2, Paper No. 2150013, 10. MR4288354

[26] Adam Osekowski, *Sharp maximal inequalities for the martingale square bracket*, Stochastics 82 (2010), 589–605. MR2783421

[27] Maurizio Pratelli, *Sur certains espaces de martingales localement de carré intégrable*, Séminaire de Probabilités, X (Seconde partie: Théorie des intégrales stochastiques, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975) (P. Meyer, ed.), Springer, 1976, pp. 401–413. Lecture Notes in Math., Vol. 511. MR0438467

[28] Michael Röckner, Longjie Xie, and Xicheng Zhang, *Superposition principle for non-local Fokker-Planck-Kolmogorov operators*, Probab. Theory Related Fields 178 (2020), no. 3-4, 699–733. MR4168386

[29] Michael Röckner and Xicheng Zhang, *Well-posedness of distribution dependent SDEs with singular drifts*, Bernoulli 27 (2021), no. 2, 1131–1158. MR4255229
[30] Michael Scheutzow, *A stochastic Gronwall lemma*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 16 (2013), no. 2, 1350019, 4. MR3078830

[31] Michael Scheutzow and Susanne Schulze, *Strong completeness and semi-flows for stochastic differential equations with monotone drift*, J. Math. Anal. Appl. 446 (2017), no. 2, 1555–1570. MR3563048

[32] Renming Song and Longjie Xie, *Well-posedness and long time behavior of singular Langevin stochastic differential equations*, Stochastic Process. Appl. 130 (2020), no. 4, 1879–1896. MR4074685

[33] Max-K. von Renesse and Michael Scheutzow, *Existence and uniqueness of solutions of stochastic functional differential equations*, Random Oper. Stoch. Equ. 18 (2010), no. 3, 267–284. MR2718125

[34] D. Willett and J. S. W. Wong, *On the discrete analogues of some generalizations of Gronwall’s inequality*, Monatsh. Math. 69 (1965), 362–367. MR185175

[35] Pengcheng Xia, Longjie Xie, Xicheng Zhang, and Guohuan Zhao, *$L^q(L^p)$-theory of stochastic differential equations*, Stochastic Process. Appl. 130 (2020), no. 8, 5188–5211. MR4108486

[36] Longjie Xie and Xicheng Zhang, *Ergodicity of stochastic differential equations with jumps and singular coefficients*, Ann. Inst. Henri Poincaré Probab. Stat. 56 (2020), no. 1, 175–229. MR4058986

[37] Xicheng Zhang and Guohuan Zhao, *Singular Brownian diffusion processes*, Commun. Math. Stat. 6 (2018), no. 4, 533–581. MR3877717