Notes on the Comparison Lemma and Various Forms of Gronwall's Inequality

SooJean Han[1*]

September 14, 2021

Abstract

In this note, we present a review of the Gronwall-Bellman Inequality, then provide a few selected extensions seen in literature. Next, we discuss applications of the Gronwall-Bellman inequality to existence and uniqueness theorems of ordinary and stochastic differential equations. Finally, we describe a tangentially-related inequality known as the Comparison Lemma, which is a widely used result in the field of control theory.

1 Original Form for ODEs

The original statement of the *Gronwall-Bellman inequality* is as follows.

Lemma 1 (Gronwall-Bellman)**.** Let $y : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function, $\mu : \mathbb{R}^+ \to \mathbb{R}^{\geq 0}$ be a nonnegative, continuous function, and $\theta : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function. Suppose that for all $t \in [0, T)$ for some $T > 0$:

$$y(t) \leq \theta(t) + \int_0^t \mu(s)y(s)ds \tag{1}$$

Then

$$y(t) \leq \theta(t) + \int_0^t \mu(s)\theta(s)e^{\int_s^t \mu(r)dr}ds \tag{2}$$

In particular, when $\theta(t) \equiv \theta$ is constant, (2) reduces to

$$y(t) \leq \theta e^{\int_0^t \mu(r)dr} \tag{3}$$

Proof. Define $z(t) := \int_0^t \mu(s)y(s)ds$ and define $v(t) := z(t) + \theta(t) - y(t)$. By (1), we have that $v(t) \geq 0$. We can write an ODE for $z(t)$ as follows:

$$\dot{z}(t) = \mu(t)y(t) = \mu(t)z(t) - \mu(t)(\theta(t) - v(t)) \tag{4}$$

*[1]SooJean Han is with the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA. Email: soojean@caltech.edu

and using the integrating factor $e^{\int_0^t \mu(s)ds}$ and the fact that $z(0) = 0$ by construction, we get

$$z(t) = \int_0^t e^{\int_0^s \mu(r)dr} \mu(s)(\theta(s) - v(s))ds \tag{5}$$

Since $v(t) \geq 0$ by construction, $\mu(t) \geq 0$ by assumption, and $e^{\int_0^s \mu(r)dr} \geq 0$, we get that

$$z(t) \leq \int_0^t e^{\int_0^s \mu(r)dr} \mu(s)\theta(s)ds \tag{6}$$

Substituting in $z(t)$ to get $y(t)$ and using the fact that $v(t) \geq 0$, we get the final inequality (2).

For the special case where $\theta(t) \equiv \theta$ is constant, (2) becomes

$$\begin{aligned}
y(t) &\leq \theta \left(1 + \int_0^t \mu(s) e^{\int_s^t \mu(r)dr} ds\right) \\
&= \theta \left(1 + \int_0^t \mu(s) e^{-\int_t^s \mu(r)dr} ds\right) \\
&= \theta \left(1 - \int_0^t \frac{d}{ds} \left\{e^{\int_s^t \mu(r)dr}\right\} ds\right) \quad \text{by FTC} \\
&= \theta \left(1 - e^{\int_s^t \mu(r)dr}\Big|_0^t\right) \\
&= \theta e^{\int_0^t \mu(r)dr}
\end{aligned} \tag{7}$$

which is exactly (3). ■

2 For Fractional Differential Equations

We consider two variations to the Gronwall-Bellman inequality from Lemma 1 in order to be able to account for *fractional differential equations*. The first variation is adapted from [1].

Lemma 2 (Gronwall-Bellman for Fractional Differential Equations)**.** Suppose $y : \mathbb{R}^+ \to \mathbb{R}$ and $\theta : \mathbb{R}^+ \to \mathbb{R}$ are continuous, $\mu : \mathbb{R}^+ \to \mathbb{R}^{\geq 0}$ is nonnegative, nondecreasing with an upper bound $\mu(t) \leq M$ for some $M > 0$ and all $t > 0$. Further suppose $\beta \in \mathbb{N}^+, \beta \geq 1$. If, for a fixed $T > 0$, the following inequality holds for all $t \in [0, T)$:

$$y(t) \leq \theta(t) + \mu(t) \int_0^t (t-s)^{\beta-1} y(s)ds \tag{8}$$

then

$$y(t) \leq \theta(t) + \sum_{n=1}^{\infty} \frac{(\Gamma(\beta)\mu(t))^n}{\Gamma(n\beta)} \int_0^t (t-s)^{n\beta-1}\theta(s)ds \tag{9}$$

for all $t \in [0, T)$.

Proof. Define the operator $\mathcal{B}\phi(t) := \mu(t) \int_0^t (t-s)^{\beta-1}\phi(s)ds$ for any $\phi : \mathbb{R}^+ \to \mathbb{R}$ continuous. The proof proceeds by induction on the number of times $\mathcal{B}$ is applied. When the operator is not applied, (8) tells us

$$y(t) \leq \theta(t) + \mathcal{B}y(t) \tag{10}$$

Substituting the expression for $y(t)$ on the right side of the inequality (10) $N \in \mathbb{N}^+$ times, we get:

$$\begin{aligned} y(t) &\leq \theta(t) + \mathcal{B}y(t) \\ &\leq \theta(t) + \mathcal{B}\left(\theta(t) + \mathcal{B}y(t)\right) = \theta(t) + \mathcal{B}\theta(t) + \mathcal{B}^2 y(t) \\ &\leq \cdots \\ &\leq \sum_{n=0}^{N-1} \mathcal{B}^n \theta(t) + \mathcal{B}^N y(t) \end{aligned} \tag{11}$$

Now suppose that the inequality

$$\mathcal{B}^n \phi(t) \leq \mu^n(t) \int_0^t \frac{\Gamma(\beta)^n}{\Gamma(n\beta)} (t-s)^{n\beta-1} \phi(s) ds \tag{12}$$

holds true for any $\phi : \mathbb{R}^+ \to \mathbb{R}$ continuous when the operator has been applied $n = k$ times.

We show that (12) holds for $n = k+1$ applying $\mathcal{B}$ across (12). By definition of the operator $\mathcal{B}$:

$$\begin{aligned} \mathcal{B}^{k+1}\phi(t) = \mathcal{B}(\mathcal{B}^k \phi(t)) &\leq \mu(t) \int_0^t (t-s)^{\beta-1} \left(\mu(s)^k \int_0^s \frac{\Gamma(\beta)^k}{\Gamma(k\beta)} (s-r)^{k\beta-1} \phi(r) dr \right) ds \\ &\leq \mu^{k+1}(t) \int_0^t (t-s)^{\beta-1} \left(\int_0^s \frac{\Gamma(\beta)^k}{\Gamma(k\beta)} (s-r)^{k\beta-1} \phi(r) dr \right) ds \\ &= \mu^{k+1}(t) \frac{\Gamma(\beta)^k}{\Gamma(k\beta)} \left(\int_0^t \int_r^t (t-s)^{\beta-1} (s-r)^{k\beta-1} ds \right) \phi(r) dr \end{aligned} \tag{13}$$

where the second-to-last inequality comes from the assumption of μ being a nondecreasing function, and the last equality comes from interchanging the limits of integration from $[0,t] \times [0,s]$ to $[0,t] \times [r,t]$. Now we can simplify the following integral using change of variables:

$$\begin{aligned} \int_r^t (t-s)^{\beta-1} (s-r)^{k\beta-1} ds &= \int_0^{t-r} (t-u-r)^{\beta-1} u^{k\beta-1} du \\ &= (t-r)^{\beta+k\beta-2} \int_0^{t-r} \left(1 - \frac{u}{t-r}\right)^{\beta-1} \left(\frac{u}{t-r}\right)^{k\beta-1} du \\ &= (t-r)^{\beta+k\beta-1} \int_0^{t-r} (1-z)^{\beta-1} z^{k\beta-1} dz \end{aligned} \tag{14}$$

Recall that the beta probability distribution is given as follows

$$1 = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1} (1-z)^{\beta-1} dz$$

Therefore, we get

$$(14) = (t-r)^{(k+1)\beta-1} \frac{\Gamma(\beta)\Gamma(k\beta)}{\Gamma((k+1)\beta)} \tag{15}$$

Substituting (15) back into (13) yields

$$(13) \leq \mu^{k+1}(t) \frac{\Gamma(\beta)^{k+1}}{\Gamma((k+1)\beta)} \int_0^t (t-r)^{(k+1)\beta-1} \phi(r) dr \tag{16}$$

which is exactly the form of (12) with k replaced by $k+1$. Thus, the formula (12) holds for all $n = 1, 2, \cdots$. Applying this to $\phi \equiv \theta$ and $\phi \equiv y$ to simplify (11), we get:

$$\begin{aligned} y(t) &\leq \sum_{k=0}^{n-1} \mathcal{B}^k \theta(t) + \mathcal{B}^n y(t) \\ &= \theta(t) + \sum_{k=1}^{n-1} \mathcal{B}^k \theta(t) + \mathcal{B}^n y(t) \\ &= \theta(t) + \sum_{k=1}^{n-1} \mu^k(t) \int_0^t \frac{\Gamma(\beta)^k}{\Gamma(k\beta)} (t-s)^{k\beta-1} \theta(s) ds + \mathcal{B}^n y(t) \end{aligned} \tag{17}$$

Taking $n \to \infty$, we see that $\mathcal{B}^n y(t) \to 0$ since

$$\lim_{n\to\infty} \mathcal{B}^n y(t) \leq \lim_{n\to\infty} \frac{(M\Gamma(\beta))^n}{\Gamma(n\beta)} \int_0^t (t-s)^{n\beta-1} y(s) ds \leq \frac{(M\Gamma(\beta))^n}{\Gamma(n\beta)} T^{n\beta} \int_0^T y(s) ds \to 0$$

since M, T, β are all fixed positive constants, $\int_0^T y(s)ds$ is a constant, and $\Gamma(\beta)^n < \Gamma(n\beta)$ for all n. Overall, (17) becomes

$$y(t) \leq \theta(t) + \sum_{k=1}^{\infty} \mu^k(t) \int_0^t \frac{\Gamma(\beta)^k}{\Gamma(k\beta)} (t-s)^{k\beta-1} \theta(s) ds \tag{18}$$

which is exactly (9). ■

3 Relaxing Nonnegativity Assumptions

Pham 2009 [2] presents a variation of Gronwall's inequality to account for the case where $\mu(t)$ is a negative coefficient. However, [2] restricts the scope to the specific form of linear functions $\theta(t) \equiv \theta t$ for a constant $\theta \in \mathbb{R}$. Below, we present an extension to more general functions $\theta(t)$ with only two conditions: 1) $\theta(0) = 0$, and 2) θ is continuously-differentiable.

Lemma 3 (Gronwall-Bellman for Negative Coefficients)**.** Let $y : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function, $\mu(t) \equiv -\mu$ with $\mu > 0$ a negative constant, and $\theta : \mathbb{R}^+ \to \mathbb{R}$ be a continuously-differetiable function such that $\theta(0) = 0$. Suppose that for all $u, t \in [0, T)$ such that $u < t$ for some $T > 0$:

$$y(t) - y(u) \leq \theta(t) - \mu \int_u^t y(s) ds \tag{19}$$

Then

$$y(t) \leq \psi(t) + [y(0) - \psi(0)]^+ e^{-\mu t} \tag{20}$$

where $[a]^+ := \max\{0, a\}$ and

$$\psi(t) := \int_0^t \dot{\theta}(s) e^{-\alpha(t-s)} ds \tag{21}$$

where the dot notation denotes the derivative with respect to time.

Proof. We can consider four separate cases of θ and $y(0)$.

1. When $\theta \equiv 0$ and $y(0) \geq 0$: Define $z(t) = y(0)e^{-\alpha t}$. Note that $z(0) = y(0) > 0$. Consider

$$z(t) - z(u) = y(0)\left(e^{-\alpha t} - e^{-\alpha u}\right) = y(0)\int_u^t -\alpha e^{-\alpha s} ds = -\alpha \int_u^t z(s)ds \tag{22}$$

Hence, $z(t)$ satisfies (19) with equality and $\theta = 0$.

Define the set $\mathcal{S} := \{t > 0 : y(t) > z(t)\}$. Suppose that there exists an element $r \in \mathcal{S}$, and let $\tau := \inf\{r' < r : \forall s \in (r', r), y(s) > z(s)\}$. Then by continuity of y and z, $y(\tau) = z(\tau)$.

Denote $\phi(t) := y(\tau) - \alpha \int_\tau^t y(s)ds$ for $t \geq \tau$. By (19) and (22), we have that $\phi(t) \geq z(t)$ for all $t \in (\tau, r)$. This means that

$$-\alpha \int_\tau^t y(s)ds \geq -\alpha \int_\tau^t z(s)ds \implies \int_\tau^t y(s)ds \leq \int_\tau^t z(s)ds \tag{23}$$

Because both y and z are continuous, r and τ cannot be a singular points where $y(s) \geq z(s)$. This contradicts the existence of $r \in \mathcal{S}$.

Hence, $\mathcal{S} = \varnothing$. This implies that $y(t) \leq z(t)$ for all $t \in [0, T)$, and together with (19), they imply that $y(t) \leq y(0)e^{-\alpha t}$, which is equivalent to (20) for $\theta = 0$ and $y(0) \geq 0$.

2. When $\theta \equiv 0$ and $y(0) \leq 0$: Define the set $\mathcal{S} := \{t > 0 : y(t) > 0\}$. Suppose that there exists an element $r \in \mathcal{S}$, and let $\tau := \inf\{r' < r : \forall s \in (r', r), y(s) > 0\}$. Choose $t_0 \in (\tau, r)$. Then (19) implies that

$$y(t_0) \leq y(\tau) - \alpha \int_\tau^{t_0} y(s)ds = \alpha \int_\tau^{t_0} y(s)ds \tag{24}$$

where the equality comes from the fact that $g(\tau) = 0$. Since $y(s) > 0$ for all $s \in (\tau, t_0)$, (24) implies that $y(t_0) \leq 0$. This contradict the fact that $t_0 \in \mathcal{S}$.

Hence, $\mathcal{S} = \varnothing$. So (19) implies that

$$y(t) \leq y(0) - \int_0^t y(s)ds \leq 0 \tag{25}$$

which is equivalent to (20) for $\theta = 0$ and $y(0) \leq 0$.

3. When $\theta \equiv \theta t$ for constant $\theta \in \mathbb{R}$: Define $\hat{y}(t) := y(t) - (\theta/\alpha)$. Then for all $t > 0$:

$$\hat{y}(t) - \hat{y}(0) = y(t) - y(0) \leq \int_0^t -\alpha y(s)ds + \theta t = \int_0^t -\alpha\left(y(s) - \frac{\theta}{\alpha}\right) ds = \int_0^t -\alpha \hat{y}(s)ds \tag{26}$$

Hence, $\hat{y}(t)$ satisfies (19) for either of the above two cases where $\theta = 0$. This implies

$$\hat{y}(t) \leq [\hat{y}(0)]^+ e^{-\alpha t} \quad \implies \quad y(t) \leq \frac{\theta}{\alpha} + \left[y(0) - \frac{\theta}{\alpha}\right]^+ e^{-\alpha t}$$

which is exactly (20).

4. Original theorem hypothesis: Consider

$$\hat{y}(t) := y(t) - \psi(t) \tag{27}$$

where y is assumed to satisfy (19) and $\psi(t)$ is chosen such that $\psi(0) = 0$. We will show that a choice of ψ as in (21) allows for $\hat{y}$ to satisfy Lemma 3 with the constant $\theta \equiv 0$, from which we can use cases 1 or 2 above. We get:

$$\hat{y}(t) - \hat{y}(0) = y(t) - y(0) - \psi(t) \text{ by (22)}$$

$$\leq -\alpha \int_0^t y(s)ds + \theta(t) - \psi(t) \quad \text{since } y \text{ satisfies (19)} \tag{28}$$

We want to choose a ψ such that we can write (23) as

$$-\alpha \int_0^t y(s)ds + \theta(t) - \psi(t) = -\alpha \int_0^t (y(s) - \psi(s))\, ds \tag{29}$$

This implies that we need

$$\theta(t) - \psi(t) = \alpha \int_0^t \psi(s)ds \tag{30}$$

which can be solved as the following ODE:

$$\dot{\theta}(t) - \dot{\psi}(t) = \alpha\psi(t) \implies \dot{\psi}(t) - \alpha\psi(t) = \dot{\theta}(t) \implies d\left(\psi(t)e^{\alpha t}\right) = \dot{\theta}(t)e^{\alpha t} \tag{31}$$

and solving for ψ in terms θ yields the expression (21). Continuing from (23), we then get:

$$\hat{y}(t) - \hat{y}(0) \leq -\alpha \int_0^t (y(s) - \psi(s))\, ds = \alpha \int_0^t \hat{y}(s)ds \tag{32}$$

and thus, $\hat{y}$ satisfies the conditions of Lemma 3 with constant $\theta = 0$. The remainder of the proof follows from cases 1 and 2 above, and the application of the inequality (20) to (22):

$$\hat{y}(t) \leq [\hat{y}(0)]^+ e^{-\alpha t} \implies y(t) \leq \psi(t) + [y(0)]^+ e^{-\alpha t} \tag{33}$$

which is exactly the result of (20).

■

4 For Discrete Systems

Now we consider an extension of Lemma 1 to discrete difference equations, adapted from [3]. The statement of the discrete Gronwall lemma is presented below.

Theorem 1 (Gronwall-Bellman for Discrete Systems)**.** Let $\{y_k\}_{k=0}^{\infty}, \{\theta_k\}_{k=0}^{\infty}$, and $\{\mu_k\}_{k=0}^{\infty}$ be sequences of real numbers with $\mu_k \geq 0$ for all k. Take a fixed value of $k_0, N \in \mathbb{N}^+$ such that $N > k_0$. Define the set

$$\mathcal{S}(k_0, N) := \text{argmax}_{k \in \{k_0, \cdots, N\}} \left\{ y_k \prod_{j=k_0}^{k-1} \frac{1}{1 + \mu_j} \right\} \tag{34}$$

Now, suppose that the following condition holds

$$y_k \leq \theta_k + \sum_{j=k_0}^{k-1} \mu_j y_j, \quad \forall\, k \in \{k_0, \cdots, N\} \tag{35}$$

Then for any $i \in \mathcal{S}(k_0, N)$, the following holds:

$$y_k \leq \theta_i \prod_{j=k_0}^{k-1} (1 + \mu_j), \quad \forall\, k \in \{k_0, \cdots, N\} \tag{36}$$

Before we proceed with the proof, we provide a couple of examples demonstrating application of the lemma to certain discrete sequences.

Example 1 (Simple Sequence of Integers)**.** Consider a sequence of positive integers defined by

$$y_k = k,\ \theta_k = 1,\ \mu_k = \frac{1}{k} \tag{37}$$

for all $k \in \{0, 1, \cdots\}$. Fix $k_0 := 0$ and $N := 5$. Note that this sequence satisfies the condition of Theorem 1:

$$\begin{aligned}
&y_0 = 0 \le 1 = \theta_0 \\
&y_1 = 1 \le 1 + 0 = \theta_1 + \mu_0 y_0 \\
&y_2 = 2 \le 1 + (0 + 1) = \theta_2 + \sum_{j=0}^{1} \mu_j y_j \\
&y_3 = 3 \le 1 + (0 + 1 + 1) = \theta_3 + \sum_{j=0}^{2} \mu_j y_j \\
&y_4 = 4 \le 1 + (0 + 1 + 1 + 1) = \theta_4 + \sum_{j=0}^{3} \mu_j y_j \\
&y_5 = 5 \le 1 + (0 + 1 + 1 + 1 + 1) = \theta_5 + \sum_{j=0}^{4} \mu_j y_j
\end{aligned}$$

The set $\mathcal{S}(0, 5)$ is defined as follows.

$$\mathcal{S}(0, 5) = \operatorname{argmax} \left\{0, \frac{1}{2}, \frac{2}{3}, \cdots, \frac{5}{6}\right\} = \{5\}$$

We can then confirm the result of the lemma (36). Since the θ_k are constant, θ_i takes the same value for all $i \in \mathcal{S}(0, 5)$.

$$\begin{aligned}
&y_0 = 0 \le \theta_i(1 + \mu_0) = 1 \\
&y_1 = 1 \le \theta_i(1 + \mu_0)(1 + \mu_1) = 2 \\
&y_2 = 2 \le \theta_i(1 + \mu_0)(1 + \mu_1)(1 + \mu_2) = 3 \\
&\quad \vdots \\
&y_5 = 5 \le \theta_i(1 + \mu_0)(1 + \mu_1)(1 + \mu_2) \cdots (1 + \mu_5) = 6
\end{aligned}$$

and so the result of Theorem 1 holds true. □

Example 2 (Solution to a Nonlinear Difference Equation)**.** Suppose we are given a system of difference equations, one which describes the nominal system and one which describes a perturbed system. Let $\{\mathbf{x}_k\}_{k=0}^{\infty} \subset \mathbb{R}^n$ be the sequence of states for the perturbed system, $\{\mathbf{y}_k\}_{k=0}^{\infty} \subset \mathbb{R}^n$ be the states for the nominal system. Let $\{f_k\}_{k=0}^{\infty}$ where $f_k : \mathbb{R}^n \to \mathbb{R}^n$ for each $k \in \mathbb{N}^+$ be the sequence of discrete nominal dynamics which is assumed to be Lipschitz with constant L_k. Further let $\{\xi_k\}_{k=0}^{\infty} \subset \mathbb{R}^n$ describe the sequence of perturbations.

$$\mathbf{x}_{k+1} = \mathbf{x}_k + f_k(\mathbf{x}_k) + \xi_k, \quad \mathbf{y}_{k+1} = \mathbf{y}_k + f_k(\mathbf{y}_k) \tag{38}$$

Suppose that $\mathbf{x}_0 = \mathbf{y}_0$. We are interested in determining a bound on the difference between the trajectories of the nominal system state $\mathbf{y}_N$ and the perturbed system state $\mathbf{x}_N$ at some time $N \in \mathbb{N}^+$. Consider the norm-difference between $\mathbf{x}_N$ and $\mathbf{y}_N$:

$$\|\mathbf{x}_N - \mathbf{y}_N\| = \left\| \cancel{\mathbf{x}_0} + \sum_{k=0}^{N} f_k(\mathbf{x}_k) + \sum_{k=0}^{N} \xi_k - \cancel{\mathbf{y}_0} - \sum_{k=0}^{N} f_k(\mathbf{y}_k) \right\| \leq \sum_{k=0}^{N} L_k \|\mathbf{x}_k - \mathbf{y}_k\| + \sum_{k=0}^{N} \|\xi_k\| \tag{39}$$

Obtaining a closed-form bound on the norm-difference can be achieved by applying Theorem 1 to (39) with $y_k := \|\mathbf{x}_k - \mathbf{y}_k\|$, $\mu_k := L_k$, and $\theta_k := \sum_{j=0}^{k} \|\xi_j\|$. This yields

$$\|\mathbf{x}_N - \mathbf{y}_N\| \leq \max_{i \in [0, N-1]} \sum_{j=0}^{i} \|\xi_j\| \prod_{j=0}^{N-1} (1 + L_j) \tag{40}$$

and we are done. □

Now we are ready to prove the discrete Gronwall-Bellman inequality.

Proof of Theorem 1. For any $j, k \in \mathbb{N}^+$ such that $k_0 \leq j \leq k$, define

$$\beta_{j,k} := \prod_{\ell=j}^{k-1} \frac{1}{1 + \mu_\ell} \tag{41}$$

Then note that $\mathcal{S}(k_0, N) = \operatorname{argmax}_{k \in \{k_0, \cdots, N\}} y_k \beta_{k_0,k}$. Multiply both sides of (35) by $\beta_{k_0,k}$ to get:

$$y_k \beta_{k_0,k} \leq \theta_k \beta_{k_0,k} + \sum_{j=k_0}^{k-1} \mu_j \beta_{k_0,k} = \theta_k \beta_{k_0,k} + \sum_{j=k_0}^{k-1} \mu_j y_j \beta_{k_0,j} \beta_{j,k} \tag{42}$$

Note that

$$\mu_j \beta_{j,k} = \left(\prod_{\ell=j+1}^{k-1} \frac{1}{1 + \mu_\ell} \right) \left(\frac{\mu_j}{1 + \mu_j} \right) = \left(\prod_{\ell=j+1}^{k-1} \frac{1}{1 + \mu_\ell} \right) \left(1 - \frac{1}{1 + \mu_j} \right) = \beta_{j+1,k} - \beta_{j,k}$$

and so, (42) becomes

$$\begin{aligned} y_k \beta_{k_0,k} &\leq \theta_k \beta_{k_0,k} + \sum_{j=k_0}^{k-1} y_j \beta_{k_0,j} (\beta_{j+1,k} - \beta_{j,k}) \\ &\leq \theta_k \beta_{k_0,k} + y_i \beta_{k_0,i} \sum_{j=k_0}^{k-1} (\beta_{j+1,k} - \beta_{j,k}) \quad \text{for any chosen } i \in \mathcal{S}(k_0, N) \\ &= \theta_k \beta_{k_0,k} + y_i \beta_{k_0,i} (1 - \beta_{k_0,k}) \quad \text{by telescoping sum} \end{aligned} \tag{43}$$

Note that by setting $k = i$ in (43) and solving for $y_i \beta_{k_0,i}$, we get:

$$y_i \beta_{k_0,i} \leq \theta_i \beta_{k_0,i} + y_i \beta_{k_0,i} (1 - \beta_{k_0,i}) \implies y_i \beta_{k_0,i} \leq \theta_i \tag{44}$$

By construction of $i \in \mathcal{S}(k_0, N)$, we have that $y_k \beta_{k_0,k} \leq y_i \beta_{k_0,i}$ for any $k \in \mathbb{N}^+$. Thus:

$$y_k \beta_{k_0,k} \leq y_i \beta_{k_0,i} \leq \theta_i \implies y_k \leq \theta_i \beta_{k_0,k}^{-1} = \theta_i \prod_{j=k_0}^{k-1} (1 + \mu_j)$$

which is exactly the desired (36). ■

Now we consider an extension of the discrete Gronwall inequality as follows.

Theorem 2 (Discrete Gronwall Inequality Extension to Theorem 1)**.** Let $\{a_n\}, \{b_n\} \subset \mathbb{R}^{\geq 0}$ be nonnegative scalar, real-valued sequences. Further let $d : \mathbb{N}^+ \times \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ be a function which satisfies the following two inequalities:

$$d_n(z) - d_n(y) \geq 0, \quad \forall 0 \leq y \leq z \tag{45a}$$

$$d_n(z) - d_n(y) \leq L_n(y)(z - y), \quad \forall 0 \leq y \leq z \tag{45b}$$

Suppose that if $\{x_n\} \subset \mathbb{R}^{\geq 0}$ is a nonnegative sequence of real numbers such that the following inequality holds:

$$x_n \leq a_n + b_n \sum_{k=0}^{n-1} d_k(x_k), \quad \forall n \in \mathbb{N}^+ \tag{46}$$

then

$$x_n \leq a_n + b_n \sum_{k=0}^{n-1} d_k(a_k) \left(\prod_{i=k+1}^{n-1} (1 + L_i(a_i) b_i) \right) \tag{47}$$

Before we prove this theorem, we consider a couple of examples.

Example 3. Consider $a_n = b_n = x_n = n$ and

$$d_n(z) := \begin{cases} \frac{z}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}$$

Then clearly (45b) is satisfied with Lipschitz constant

$$L_n(z) = \begin{cases} \frac{1}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}$$

for all $z \in \mathbb{R}^{\geq 0}$.

We can show that (46) holds for all $n \in \mathbb{N}^+$ because:

$$n \leq n + n \sum_{k=1}^{n-1} \frac{1}{k} k = n + n(n-1)$$

Indeed, substituting the values into (47) shows that it too holds for all $n \in \mathbb{N}^+$:

$$n \leq n + n \sum_{k=1}^{n-1} \frac{1}{k} k \prod_{i=k+1}^{n-1} \left(1 + \frac{1}{i} i\right) = n + n \sum_{k=1}^{n-1} 2^{n-k-1} = n 2^n$$

This particular example is an instance where the result (47) of Theorem 2 is a looser inequality than (46).
$\square$

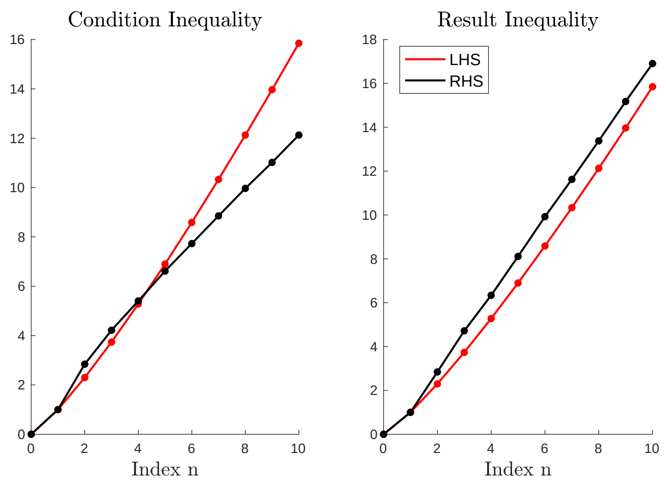

Figure 1: A comparison of the inequalities (46) (the Condition Inequality) and (47) (the Result Inequality) for Example 4 for up to $n = 10$. The red line refers to the left-hand side of each respective inequality, which is x_n, and the black line refers to the right-hand side of each inequality.

Example 4 (Sufficient, but not Necessary)**.** The satisfaction of (46) is sufficient for (47) to hold, but not necessary. We demonstrate this via the following counterexample.

Consider $x_n = n^{1.2}$, $a_n = n$, $b_n = 1$, and

$$d_n(z) = \begin{cases} \frac{1}{n}|\sin(z)| & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}$$

for any $z \in \mathbb{R}^{\geq 0}$. One way that (45b) can be satisfied is by using the Lipschitz constant

$$L_n(z) = \begin{cases} \frac{1}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}$$

for any $z \in \mathbb{R}^{\geq 0}$, since $|\sin(z)|$ is Lipschitz with constant 1.

The right side of (46) is computed as follows:

$$a_n + b_n \sum_{k=0}^{n-1} d_k(x_k) = n + \sum_{k=1}^{n-1} \frac{1}{k} \left|\sin\left(k^{1.2}\right)\right|$$

There exist indices $n \in \mathbb{N}^+$ where the inequality (46) is not satisfied. In particular, for $n \geq 5$:

$$\begin{aligned} n = 5: \quad & 5 + \sum_{k=1}^{4} \frac{1}{k} \left|\sin\left(k^{1.2}\right)\right| = 6.6133 < 6.8986 = 5^{1.2} \\ n = 6: \quad & 6 + \sum_{k=1}^{5} \frac{1}{k} \left|\sin\left(k^{1.2}\right)\right| = 7.7287 < 8.5858 = 6^{1.2} \\ & \ldots \end{aligned}$$

However, for some of these same indices, the inequality (47) holds. The right side is computed as follows:

$$a_n + b_n \sum_{k=0}^{n-1} d_k(a_k) \left(\prod_{i=k+1}^{n-1} (1 + L_k(a_k) b_k) \right) = n + \sum_{k=1}^{n-1} \frac{1}{k} |\sin(k)| \left(\prod_{i=k+1}^{n-1} 1 + \frac{1}{i} \right)$$

And for some values of $n \geq 5$, (47) is satisfied.

$$
\begin{aligned}
&n = 5: \; 5 + \sum_{k=1}^{4} \frac{1}{k} |\sin(k)| \left(\prod_{i=k+1}^{4} 1 + \frac{1}{i} \right) = 8.1094 > 5^{1.2} \\
&n = 6: \; 6 + \sum_{k=1}^{5} \frac{1}{k} |\sin(k)| \left(\prod_{i=k+1}^{5} 1 + \frac{1}{i} \right) = 9.9231 > 6^{1.2} \\
&\cdots
\end{aligned}
$$

We illustrate the relationship of x_n with the respective right sides of both inequalities (46) and (47) in Figure 1. Clearly, for $n \in \{0, \cdots, 4\}$ both inequalities (46) and (47) are satisfied. However, (47) is satisfied for larger values of n, all the way up to $n = 10$. This indicates that for truncated sequences up to some fixed value of the index n, (46) is not a necessary condition for (47) to hold. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Define $y_n := \sum_{k=0}^{n-1} d_k(x_k)$ for $n \geq 1$ with $y_0 = 0$. Consider the following difference

$$
\begin{aligned}
y_{n+1} - y_n &= d_n(x_n) \\
&\leq d_n(a_n + b_n y_n) \quad \text{by (46) and (45a)} \\
&\leq d_n(a_n) + L_n(a_n) b_n y_n \quad \text{y (45b)}
\end{aligned} \tag{48}
$$

Rearranging terms of (48) yields

$$
y_{n+1} \leq d_n(a_n) + y_n \left(1 + L_n(a_n) b_n\right) \tag{49}
$$

Define $\alpha_n := d_n(a_n)$, $\beta_n := 1 + L_n(a_n) b_n$, and

$$
\gamma_n := y_n \prod_{k=0}^{n-1} \beta_k^{-1}, \;\; \gamma_0 := 0
$$

Then we can rewrite (49) as

$$
\prod_{k=0}^{n} \beta_k \gamma_{n+1} \leq \alpha_n + \beta_n \gamma_n \prod_{k=0}^{n-1} \beta_k = \alpha_n + \gamma_n \prod_{k=0}^{n-1} \beta_k \implies \gamma_{n+1} - \gamma_n \leq \alpha_n \prod_{k=0}^{n-1} \beta_k^{-1} \tag{50}
$$

Summing (50) from 0 to $n-1$, then substituting the values of $\alpha_n, \beta_n, \gamma_n$ back in yields

$$
\gamma_n - \gamma_0 \leq \left(\prod_{k=0}^{n-1} \beta_k \right) \sum_{k=0}^{n-1} \alpha_k \left(\prod_{i=0}^{k} \beta_i^{-1} \right) \implies y_n \leq \sum_{k=0}^{n-1} d_k(a_k) \left(\prod_{i=k+1}^{n-1} (1 + L_i(a_i) b_i) \right) \tag{51}
$$

Substituting (51) into (46) then yields (47). ■

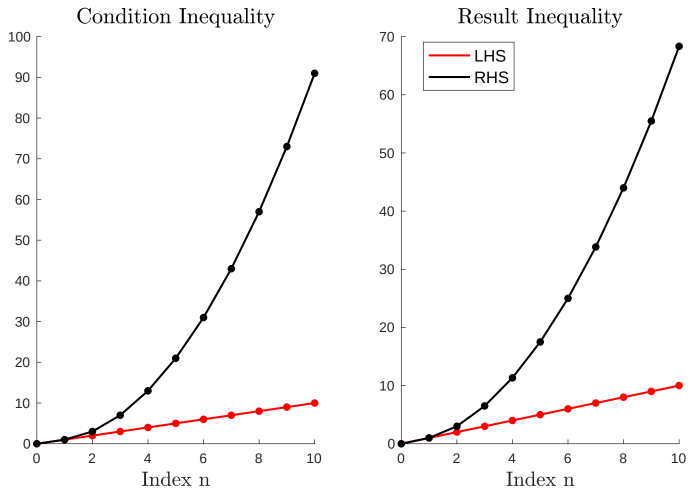

Figure 2: A comparison of the inequalities (46) (the Condition Inequality) and (47) (the Result Inequality) for Example 5 for up to $n = 10$. As in Figure 1, the red line refers to the left-hand side of each respective inequality, which is x_n, and the black line refers to the right-hand side of each inequality.

Example 5 (Relaxing Nonnegativity Assumptions)**.** One might be interested in developing theorems analogous to Theorem 1 or Theorem 2 for the cases where some of the sequences are allowed to be nonnegative. In particular, consider an example where $\{b_n\}$ from Theorem 2 is a nonpositive sequence. Let $x_n = n, a_n = n^2$, $b_n = -1$ for all $n \in \mathbb{N}^+$, and let $d_n(z) := z/n$ if $n \geq 1$ and 0 if $n = 0$ for all $z \in \mathbb{R}^{\geq 0}$. Then condition (46) is clearly satisfied for all $n \in \mathbb{N}$:

$$n \leq n^2 - \sum_{k=1}^{n-1} \frac{1}{k} k = n^2 - \sum_{k=1}^{n-1} 1 = n^2 - n + 1 = n(n-1) + 1$$

Condition (47) is also satisfied:

$$\begin{aligned}
& n \leq n^2 - \sum_{k=1}^{n-1} \frac{1}{k} k^2 \prod_{i=k+1}^{n-1} \left(1 - \frac{1}{i}\right) = n^2 - \sum_{k=1}^{n-1} k \prod_{i=k+1}^{n-1} \left(1 - \frac{1}{i}\right) \\
& \implies n = 0: \quad 0 \leq 0 \\
& \implies n = 1: \quad 1 \leq 1 \\
& \implies n = 2: \quad 2 \leq 4 - \sum_{k=1}^{1} k \prod_{i=k+1}^{1} \left(1 - \frac{1}{i}\right) = 4 - (0 + 1) = 3 \\
& \implies n = 3: \quad 3 \leq 9 - \sum_{k=1}^{2} k \prod_{i=k+1}^{2} \left(1 - \frac{1}{i}\right) = 9 - \left(\frac{1}{2} + 2\right) = 6.5 \\
& \ldots
\end{aligned}$$

We illustrate the relationship of x_n with the respective right sides of both inequalities (46) and (47) in Figure 2 for values of n up to 10. Clearly, both inequalities hold true, as the right-hand side is always greater than or equal to the left-hand side.

In fact, for any version of Theorem 2 which allows $\{b_n\}$ to take negative values, the only condition that needs to be imposed is that $\{b_n\}$ must be so that $L_k(a_n)b_n > -1$ for all $n \in \mathbb{N}^+$. We demonstrate this in the corollary below.

□

Corollary 1 (Theorem 2 for Relaxed Assumption on b_n)**.** Let $\{a_n\} \subset \mathbb{R}^{\geq 0}$ be a nonnegative scalar, real-valued sequence, and let $d : \mathbb{N}^+ \times \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ be a function which satisfies the two inequalities of (45). Further, let $\{b_n\} \subset \mathbb{R}$ be such that $L_n(a_n)b_n > -1$ for all $n \in \mathbb{N}^+$.

Suppose that if $\{x_n\} \subset \mathbb{R}^{\geq 0}$ is a nonnegative sequence of real numbers such that the inequality (46) holds. Then (47) holds as a result.

Proof of Corollary 1. The proof follows similarly to the proof to Theorem 2]. Define $y_n := \sum_{k=0}^{n-1} d_k(x_k)$ for $n \geq 1$ with $y_0 = 0$. Then the argument until (49) holds in very much the same way.

Define $\alpha_n := d_n(a_n)$, $\beta_n := 1 + L_n(a_n)b_n$, and $\gamma_n := y_n \prod_{k=0}^{n-1} \beta_k^{-1}$ with $\gamma_0 := 0$. Note that by the condition imposed on b_n, β_n and γ_n are positive for all $n \in \mathbb{N}^+$. We get the same (50) as before, and we can sum n equations from 0 to $n-1$. Because β_n and γ_n are positive, dividing the resulting cumulative sum inequality through by $\prod_{k=0}^{n-1} \beta_k$ does not reverse the inequality. Thus, we obtain exactly (47), and we are done. ■

A direct extension to Theorem 2 is made by modifying the condition (45) on the function $d : \mathbb{N} \times \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$. If instead, d was differentiable in its second argument with the conditions

$$\partial_x d_n(x) \geq 0, \quad \forall x \geq 0, n \in \mathbb{N} \tag{52a}$$
$$\exists L : \mathbb{N} \times \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0} \text{ s.t. } \partial_x d_n(x) \leq L_n(z), \quad \forall n \in \mathbb{N}, 0 \leq z \leq x \tag{52b}$$

Then for given nonnegative, real-valued sequences $\{a_n\}, \{b_n\}, \{x_n\}$, if (46) holds with (52) in place, the inequality (47) holds. This can be easily seen by first applying the Mean-Value Theorem on d: for any $x, y \in \mathbb{R}^{\geq 0}$ with $0 \leq x \leq y$, there exists a $z \in (x, y)$ such that

$$d_k(y) - d_k(x) = \partial_x d_k(z)(y - x)$$

Choosing $z = x$ in the above, then applying (52) yields exactly (45). We can then carry out the proof in Theorem 2.

An even further extension to Theorem 2 can be made based on the differentiability condition (52).

Corollary 2 (Theorem 2 with Differentiable d)**.** Let $\{a_n\}, \{b_n\} \subset \mathbb{R}^{\geq 0}$ be nonnegative scalar, real-valued sequences. Further let $d : \mathbb{N}^+ \times \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ be a function which is differentiable with respect to its second argument, and satisfies (52a) and:

$$\partial_x d_n(y) \leq \partial_x d_n(x), \quad \forall n \in \mathbb{N}, 0 \leq x \leq y \tag{53}$$

Suppose that if $\{x_n\} \subset \mathbb{R}^{\geq 0}$ is a nonnegative sequence of real numbers such that the following inequality holds:

$$x_n \leq a_n + b_n \sum_{k=0}^{n-1} d_k(x_k), \quad \forall n \in \mathbb{N}^+ \tag{54}$$

then

$$x_n \leq a_n + b_n \sum_{k=0}^{n-1} d_k(a_k) \left(\prod_{i=k+1}^{n-1} (1 + \partial_x d_i(x_i) b_i) \right) \tag{55}$$

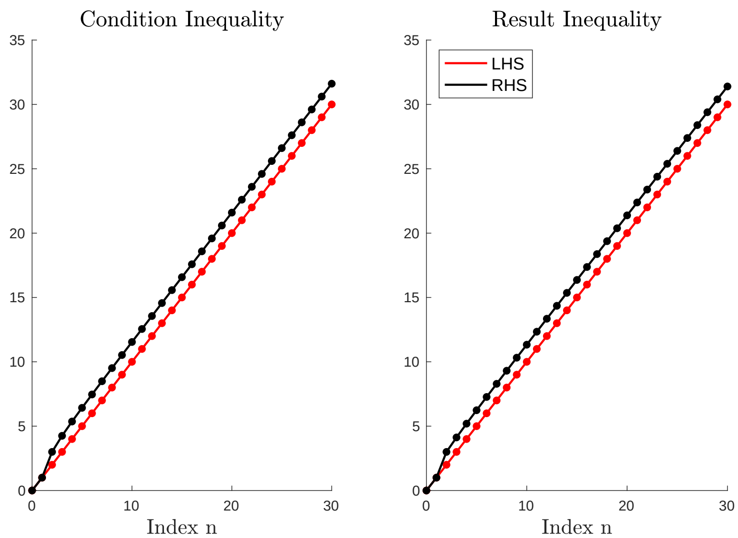

Figure 3: A comparison of the inequalities (54) (the Condition Inequality) and (55) (the Result Inequality) from Corollary 2 in Example 5 for up to $n = 30$. We demonstrate a choice of $f(z) = 1/z$ and $x_n = n$.

Moreover, a useful version of Corollary 2 occurs when $d_n(z) := c_n f(z)$ is a separable function of its arguments, where $\{c_n\}$ is a nonnegative sequence of real numbers and differentiable function $f : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ which is monotonic increasing with derivative f' nondecreasing on $\mathbb{R}^+$. The following example demonstrates cases of $d_n(z) := c_n f(z)$ in which the inequalities of Corollary 2 are satisfied even when the condition on f and f' are broken.

Example 6 (Corollary 2 when f is not Monotonic Increasing or f' is not Nondecreasing)**.** Let $a_n = n$ and $b_n = 1$. Further suppose $d_n(z) := c_n f(z)$ holds with

$$c_n = \begin{cases} \frac{1}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}$$

First, consider $x_n = n$ with $f(z) = 1/z$ for $z > 0$. This implies $f(z) = -1/z^2$. Note that f breaks the condition of being monotonic increasing. However, (54) and (55) are satisfied. Namely, the inequality (54) writes as:

$$n \leq n + \sum_{k=1}^{n-1} \frac{1}{k^2}$$

which is clearly true since $\sum_{k=1}^{n-1} 1/k^2 \geq 0$. Meanwhile, (55) writes as

$$n \leq n + \sum_{k=1}^{n-1} \frac{1}{k^2} \left(\prod_{i=k+1}^{n-1} \left(1 - \frac{1}{i^2} \right) \right)$$

which is true because the summation on the right-hand side is nonnegative. Further note that this choice of f ensures that $f'(z)b > 1$. Hence, the inequality is satisfied even though $f' < 0$. This is similar to the relaxed condition on b from which Corollary 1 was derived.

Next, consider $x_n = 1$ for all $n \geq 1$ with $x_0 = 0$, and $f(z) = -z^{1.2}$ for $z > 0$. This implies $f(z) = -1.2z^{0.2}$. Unlike the previous case, f' breaks the condition of being nondecreasing in addition to f not

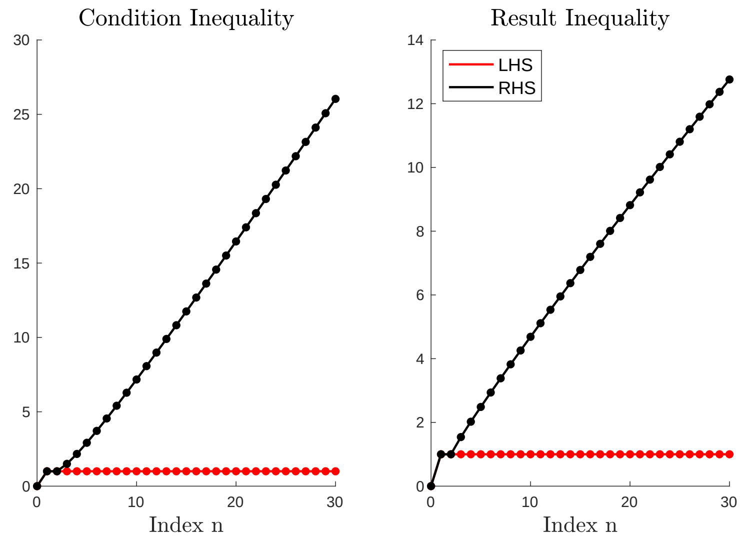

Figure 4: Figure 3 for a choice of $f(z) = -z^{1.2}$ and $x_n = 1$.

being monotonic increasing. Moreover, f is a nonpositive function. Despite these condition violations, (54) and (55) are satisfied. First, we have the condition inequality:

$$1 \leq n + \sum_{k=1}^{n-1} \frac{1}{k} \cdot (-k^{1.2}) = n - \sum_{k=1}^{n-1} k^{0.2}$$

and the result inequality

$$1 \leq n + \sum_{k=1}^{n-1} \frac{1}{k}(-k^{1.2}) \left(\prod_{i=k+1}^{n-1} \left(1 - \frac{1.2}{i}(i^{0.2})\right)\right) = n - \sum_{k=1}^{n-1} k^{0.2} \left(\prod_{i=k+1}^{n-1} \left(1 - \frac{1.2}{i^{0.8}}\right)\right)$$

Both examples are visualized in Figure 3 for n up to 30. $\square$

5 For Stochastic Systems

In this section, we present a version of the Gronwall inequality for stochastic systems. While the deterministic Gronwall inequalities in the sections prior can be applied to stochastic systems by taking the expected value, we present a version in which the inequality is satisfied 1) almost-surely and 2) with respect to higher-order moments. The result is adapted from [4], but we present a version of the proof which is simplified compared to what is shown in [4]. A version of the inequality for moments $p \in (0, 1)$ has been presented in [5].

Definition 1 ($\mathcal{L}^p$ Norm)**.** Suppose we are given any probability space $(\Omega, \mathcal{F}, \mathbb{P})$, moment $p \in (0, \infty]$, vector space $(V, \|\cdot\|_V)$, and any stochastic function $f : \mathcal{F} \to \mathcal{B}(V)$, where $\mathcal{B}(V)$ is the set of Borel-measurable subsets of V. Then, define the $\mathcal{L}^p(V; \mathbb{R})$ norm of f as follows:

$$\|f\|_{\mathcal{L}^p(V;\mathbb{R})} := \begin{cases} \mathbb{E}\left[\|f\|_V^p\right]^{\frac{1}{p}} & \text{if } p < \infty \\ \inf\{c \in [0, \infty) : \|f\|_V \leq c \ \ \mathbb{P}\text{- a.s.}\} & \text{if } p = \infty \end{cases} \tag{56}$$

$\square$

Definition 2 (HS-Norm)**.** Let $(U, \langle\cdot,\cdot\rangle_U, \|\cdot\|_U)$ and $(H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ be separable Hilbert spaces. Further let $\mathbb{U}$ denote the countable, orthonormal basis corresponding to U. Denote $\mathcal{L}(U,H)$ to be the set of all linear operators X mapping from U to H. Then for all $f \in \mathcal{L}(U,H)$, its HS-norm is defined as follows

$$\|f\|_{HS(U,H)} := \sum_{u\in\mathbb{U}} \|Xu\|_H^2 \tag{57}$$

Further denote the set

$$HS(U,H) := \left\{X \in \mathcal{L}(U,H) | \, \|X\|_{HS(U,H)} < \infty\right\} \tag{58}$$

□

A particularly useful result which will be used throughout the proof of the stochastic Gronwall inequality is *Hölder's inequality*, stated below.

Lemma 4 (Hölder's Inequality)**.** Let (S, Σ, μ) be a measure space and let $p, q \in [1, \infty]$ with

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then for all measurable real-valued functions f and g on S,

$$\|f \cdot g\|_{\mathcal{L}^1(S;\mathbb{R})} \leq \|f\|_{\mathcal{L}^p(S;\mathbb{R})} \|g\|_{\mathcal{L}^q(S;\mathbb{R})} \tag{59}$$

The stochastic Gronwall inequality is split into two parts. In the assumption below, we provide the universal setting and notation under which both parts operate.

Assumption 1. Suppose we are given two separable Hilbert spaces and $(U, \langle\cdot,\cdot\rangle_U, \|\cdot\|_U)$ and $(H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$. Fix $T \in \mathbb{R}^+$ and $\mathcal{O} \subseteq H$ an open subset of H. Let $X : [0,T] \times \Omega \to \mathcal{O}, \mu : [0,T] \times \Omega \to H$, $\sigma : [0,T] \times \Omega \to HS(U,H)$, and $\alpha, \beta : [0,T] \times \Omega \to [0,\infty]$ be adapted processes on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ with continuous sample paths. Let $\tau : \Omega \to [0,T]$ denote a stopping time. Let $p \in [1,\infty)$ be the moment, and let $V : [0,T] \times \mathcal{O} \to \mathbb{R}^+, V \in \mathcal{C}^{(1,2)}$ denote a Lyapunov function. Given this setup and notation, we assume the following two conditions:

1. The condition

$$\int_0^\tau \left(\|\mu(s)\|_H + \|\sigma(s)\|_{HS(U,H)}^2\right) ds < \infty \tag{60}$$

   holds almost-surely and X can be expressed as the following process:

$$X(t \wedge \tau) = X_0 + \int_0^t \mathbb{1}_{[0,\tau]}(s)\mu(s)ds + \int_0^t \mathbb{1}_{[0,\tau]}(s)\sigma(s)dW(s) \tag{61}$$

   for some initial state $X_0 \in \mathbb{R}^n$ and where $t \in [0,T]$ and $\{W(t)\}_{t\in[0,T]}$ is a standard Wiener process.

2. The condition

$$\int_0^\tau |\alpha(s)|ds < \infty \tag{62}$$

   holds almost-surely, and the Lyapunov function V satisfies

$$\begin{aligned} &\partial_t V(t, X(t)) + \partial_x V(t, X(t))\mu(t) + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T(t) H_x V(t, X(t))\right) \\ &\quad + \frac{p-1}{2V(t,X(t))} \|\partial_x V(t, X(t))\sigma(t)\|_{HS(U,\mathbb{R})}^2 \leq \alpha(t)V(t,X(t)) + \beta(t) \end{aligned} \tag{63}$$

   for all $t \in [0,T]$.

$\square$

Theorem 3 (Stochastic Gronwall Inequality for Moments $p \in [1,\infty)$: Result 1)**.** Suppose we are given the setup and notation in Assumption 1. Then, for all $q_1, q_2 \in (0,\infty]$ with $\frac{1}{q_1} = \frac{1}{q_2} + \frac{1}{p}$,

$$\|V(\tau, X(\tau))\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} \leq \left\|\phi(\tau)^{-1}\right\|_{\mathcal{L}^{q_2}(\mathbb{P};\mathbb{R})} \left(\|V(0,X_0)\|_{\mathcal{L}^{p}(\mathbb{P};\mathbb{R})} + \int_0^T \left\|\mathbb{1}_{[0,\tau]}(s)\beta(s)\phi(s)\right\|_{\mathcal{L}^{p}(\mathbb{P};\mathbb{R})} ds\right) \tag{64}$$

where

$$\phi(t) := e^{-\int_0^t \alpha(s)ds} \tag{65}$$

Theorem 4 (Stochastic Gronwall Inequality for Moments $p \in [1,\infty)$: Result 2)**.** Suppose we are given the setup and notation in Assumption 1. Then, for all $q_1, q_2, q_3 \in (0,\infty]$ with $q_3 < p$ and $\frac{1}{q_1} = \frac{1}{q_2} + \frac{1}{q_3}$,

$$\left\|\sup_{t\in[0,\tau]} V(t, X(s))\right\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} \leq K \left\|\phi(\tau)^{-1}\right\|_{\mathcal{L}^{q_2}(\mathbb{P};\mathbb{R})} \left\|V(0,X_0) + p\int_0^\tau \beta(s)\phi(s)ds\right\|_{\mathcal{L}^{q_3}(\mathbb{P};\mathbb{R})} \tag{66}$$

where

$$K := \left(\int_{\frac{p-q_3}{q_3}}^{\infty} \frac{s^{\frac{q_3}{p}-1}}{s+1} ds + 1\right)^{\frac{p}{q_3}}$$

and $\phi(t)$ defined as in (65).

First, we require the use of the following two results in order to prove both Theorem 3 and Theorem 4.

Lemma 5 (Itô Formula with Integrating Factor)**.** Suppose we are given the notation and setup of Assumption 1. Further define $\gamma : [0,T]\times\Omega \to \mathbb{R}$ and $\eta : [0,T]\times\Omega \to HS(U,\mathbb{R})$ such that

$$\int_0^\tau \left(|\gamma(s)| + \|\eta(s)\|^2_{HS(U,\mathbb{R})}\right) ds < \infty$$

Define

$$\phi(t) := e^{\int_0^t \left(-\gamma(s) + \frac{1}{2}\|\eta(s)\|^2_{HS(U,\mathbb{R})}\right)ds - \int_0^t \eta(s)dW(s)}$$

Then applying Itô's formula to $V(t\wedge\tau, X(t\wedge\tau))\phi(t\wedge\tau)$ yields:

$$\begin{aligned}
V(t\wedge\tau, X(t\wedge\tau))\phi(t\wedge\tau) - V(0,X_0) &= \int_0^t \mathbb{1}_{[0,\tau]}(s)\partial_s\left\{V(s,X(s))\phi(s)\right\} ds \\
&\quad + \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\partial_x V(s,X(s))dX(s) \\
&\quad + \frac{1}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)H_x V(s,X(s))d\langle X, X\rangle(s) \\
= \int_0^t \mathbb{1}_{[0,\tau]}(s) &\left[\phi(s)\partial_s V(s,X(s)) + V(s,X(s))\phi(s)\left(-\gamma(s) + \frac{1}{2}\|\eta(s)\|^2_{HS(U,\mathbb{R})}\right)\right] ds \\
&\quad - \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)V(s,X(s))\eta(s)dW(s) \\
&\quad + \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\partial_x V(s,X(s))(\mu(s)ds + \sigma(s)dW(s))
\end{aligned}$$

$$+\frac{1}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\mathrm{tr}\left(\sigma\sigma^T(s)H_xV(s,X(s))\right)ds$$

$$=\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\left[\partial_sV(s,X(s))-V(s,X(s))\left(\gamma(s)-\frac{1}{2}\|\eta(s)\|^2_{HS(U,\mathbb{R})}\right)+\partial_xV(s,X(s))\mu(s)\right]ds$$

$$+\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\left[-V(s,X(s))\eta(s)+\partial_xV(s,X(s))\sigma(s)\right]dW(s)$$

$$+\frac{1}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi(s)\mathrm{tr}\left(\sigma\sigma^T(s)H_xV(s,X(s))\right)ds \tag{67}$$

Much of the simplification in (67) was obtained by using chain rule.

Lemma 6 (Gronwall-Bellman-Opial Inequality)**.** Let $T\in[0,\infty)$, $p\in(1,\infty)$. Further let $y,\beta:[0,T]\times\Omega\to\mathbb{R}^{\geq 0}$ be functions such that for all $t\in[0,T]$:

$$y^p(t)\leq y_0^p+p\int_0^t y^{p-1}(s)\beta(s)ds \tag{68}$$

Then

$$y(t)\leq y_0+\int_0^t\beta(s)ds \tag{69}$$

Now we are ready to prove the main stochastic Gronwall results.

Proof of Theorem 3. Define the sequence of stopping times

$$\tau_n:=\tau\wedge\inf\left\{s\in[0,T]\;\middle|\;V(s,X(s))+\int_0^s\|\partial_xV(r,X(r))\sigma(r)\|^2_{HS(U,H)}\,dr\geq n\right\} \tag{70}$$

for all $n\in\mathbb{N}^+$ such that

$$\tau=\lim_{n\to\infty}\tau_n$$

Apply Itô's formula to the function $(\varepsilon+V(t,x))^p\phi^p(t)$ for all $\varepsilon\in\mathbb{R}^+$, where $\phi(t)$ is defined in (65), and simplify using the same argument as in Lemma 5. Then

$$(\varepsilon+V(t\wedge\tau,X(t\wedge\tau))^p\phi^p(t\wedge\tau)-(\varepsilon+V(0,X_0))^p=\int_0^t\mathbb{1}_{[0,\tau]}(s)\partial_s\left\{(\varepsilon+V(s,X(s)))^p\phi^p(s)\right\}ds$$

$$+\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)\partial_x\{(\varepsilon+V(s,X(s)))^p\}dX(s)$$

$$+\frac{1}{2}\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)H_x\{(\varepsilon+V(s,X(s)))^p\}d\langle X,X\rangle(s)$$

$$=\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)\left[p(\varepsilon+V(s,X(s)))^{p-1}\partial_sV(s,X(s))-p(\varepsilon+V(s,X(s)))^p\alpha(s)\right]ds$$

$$+\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\partial_xV(s,X(s))\left(\mu(s)ds+\sigma(s)dW(s)\right)$$

$$+\frac{1}{2}\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)\mathrm{tr}\left(\sigma\sigma^T(s)H_x(\varepsilon+V(s,X(s)))^p\right)ds \tag{71}$$

and note that the last term of (71) can be simplified to

$$\frac{1}{2}\int_0^t\mathbb{1}_{[0,\tau]}(s)\phi^p(s)\mathrm{tr}\left(\sigma\sigma^T(s)H_x(\varepsilon+V(s,X(s)))^p\right)ds$$

$$
\begin{aligned}
&= \frac{1}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)\mathrm{tr}\left(\sigma\sigma^T(s)p\partial_x\left[(\varepsilon+V(s,X(s)))^{p-1}\partial_x V(s,X(s))\right]\right)ds\\
&= \frac{p}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)\mathrm{tr}\left(\sigma\sigma^T(s)\left[(p-1)(\varepsilon+V(s,X(s)))^{p-2}\partial_x V(s,X(s))\right.\right.\\
&\qquad\left.\left.+(\varepsilon+V(s,X(s)))^{p-1}H_x V(s,X(s))\right]\right)ds\\
&= \frac{p(p-1)}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)(\varepsilon+V(s,X(s)))^{p-2}\left\|\partial_x V(s,X(s))\sigma(s)\right\|^2_{HS(U,H)}\\
&\qquad+\frac{p}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)(\varepsilon+V(s,X(s)))^{p-1}\mathrm{tr}\left(\sigma\sigma^T(s)H_x V(s,X(s))\right)ds
\end{aligned}
\tag{72}
$$

Substituting (72) back into (71) yields

$$
\begin{aligned}
(71) = &\int_0^t \mathbb{1}_{[0,\tau]}(s)p\phi^p(s)\left[(\varepsilon+V(s,X(s)))^{p-1}\partial_s V(s,X(s))-(\varepsilon+V(s,X(s)))^p\alpha(s)\right]ds\\
&+\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\partial_x V(s,X(s))\left(\mu(s)ds+\sigma(s)dW(s)\right)\\
&+\frac{p(p-1)}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)(\varepsilon+V(s,X(s)))^{p-2}\left\|\partial_x V(s,X(s))\sigma(s)\right\|^2_{HS(U,H)}\\
&+\frac{p}{2}\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)(\varepsilon+V(s,X(s)))^{p-1}\mathrm{tr}\left(\sigma\sigma^T(s)H_x V(s,X(s))\right)ds
\end{aligned}
\tag{73}
$$

Now we can apply Assumption 1 condition 2 to upper bound (73):

$$
\begin{aligned}
&(\varepsilon+V(t\wedge\tau,X(t\wedge\tau))^p\phi^p(t\wedge\tau)-(\varepsilon+V(0,X_0))^p\\
&\quad\le\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\partial_x V(s,X(s))\sigma(s)dW(s)\\
&\qquad+\int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds
\end{aligned}
\tag{74}
$$

Now suppose WLOG that

$$
\mathbb{E}\left[|V(0,X_0)|^p\right]<\infty,\qquad \int_0^T\left\|\mathbb{1}_{[0,\tau]}(s)\beta(s)\phi(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}ds<\infty
$$

Then note that each τ_n for every $n\in\mathbb{N}^+$ is a stopping time, and the integrals on the right side of (74) are integrable. Rewrite (74) with $\tau=\tau_n$ and consider the following $\mathcal{L}^p(\mathbb{P};\mathbb{R})$ norm:

$$
\begin{aligned}
&\|(\varepsilon+V(t\wedge\tau_n,X(t\wedge\tau_n))\phi(t\wedge\tau_n)\|^p_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}:=\mathbb{E}\left[|(\varepsilon+V(t\wedge\tau_n,X(t\wedge\tau_n))^p\phi^p(t\wedge\tau_n)|\right]\\
&\quad\le\mathbb{E}\left[|\varepsilon+V(0,X_0)|^p\right]+\mathbb{E}\left[\int_0^t \mathbb{1}_{[0,\tau_n]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\partial_x V(s,X(s))\sigma(s)dW(s)\right]\\
&\qquad+\mathbb{E}\left[\int_0^t \mathbb{1}_{[0,\tau_n]}(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right]\\
&\quad=\|\varepsilon+V(0,X_0)\|^p_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}+\int_0^t\mathbb{E}\left[\mathbb{1}_{[0,\tau_n]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)\right]ds
\end{aligned}
\tag{75}
$$

since integration with respect to the standard Wiener process has mean zero. Split up the exponential integrating factor as $\phi^p(s)=\phi^{p-1}(s)\phi(s)$ in(75), then use Hölder's inequality (Lemma 4) to get:

$$
(75)\le\|\varepsilon+V(0,X_0)\|^p_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}+\int_0^t\mathbb{E}\left[\mathbb{1}_{[0,\tau_n]}(s)\phi^{p-1}(s)\phi(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)\right]ds
$$

$$\leq \|\varepsilon + V(0, X_0)\|^p_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}$$
$$+ p \int_0^t \|\phi(s\wedge\tau_n)(\varepsilon + V(s\wedge\tau_n, X(s\wedge\tau_n)))\|^{p-1}_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} \left\|\mathbb{1}_{[0,\tau_n]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} ds \tag{76}$$

Apply Gronwall-Bellman-Opial (Lemma 6) to (76) with

$$\gamma(s) := \|\phi(s\wedge\tau_n)(\varepsilon + V(s\wedge\tau_n, X(s\wedge\tau_n)))\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}, \qquad \eta(s) := \left\|\mathbb{1}_{[0,\tau_n]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}$$

then take $t \to T$ to get

$$\|\phi(T\wedge\tau_n)(\varepsilon + V(T\wedge\tau_n, X(T\wedge\tau_n)))\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})}$$
$$\leq \|\varepsilon + V(0, X_0)\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} + \int_0^T \left\|\mathbb{1}_{[0,\tau_n]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} ds \tag{77}$$

Now consider the left side of our desired inequality (64). Take $q_1, q_2 \in \mathbb{R}^+$ such that $\frac{1}{q_1} = \frac{1}{q_2} + \frac{1}{p}$. Then we can apply Hölder's inequality to the $\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})$-norm of $V(\tau, X(\tau))$ to get

$$\|V(\tau, X(\tau))\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} = \left\|\phi(\tau)^{-1}V(\tau, X(\tau))\phi(\tau)\right\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} \leq \left\|\phi(\tau)^{-1}\right\|_{\mathcal{L}^{q_2}(\mathbb{P};\mathbb{R})} \|V(\tau, X(\tau))\phi(\tau)\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} \tag{78}$$

The $\mathcal{L}^p(\mathbb{P};\mathbb{R})$-norm of $V(\tau, X(\tau))\phi(\tau)$ can be simplified using monotone convergence:

$$\begin{aligned}
\|V(\tau, X(\tau))\phi(\tau)\|^p_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} &:= \mathbb{E}\left[|V(\tau, X(\tau))\phi(\tau)|^p\right] \\
&= \mathbb{E}\left[\lim_{n\to\infty} |V(\tau_n, X(\tau_n))\phi(\tau_n)|^p\right] \\
&= \lim_{\varepsilon\to 0}\lim_{n\to\infty} \mathbb{E}\left[|(\varepsilon + V(\tau_n, X(\tau_n)))\phi(\tau_n)|^p\right] \quad \text{by Fatou's Lemma} \\
&\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty} \left(\|\varepsilon + V(0, X_0)\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} + \int_0^T \left\|\mathbb{1}_{[0,\tau_n]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} ds\right)^p \quad \text{by (77)} \\
&= \left(\|V(0, X_0)\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} + \int_0^T \left\|\mathbb{1}_{[0,\tau]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} ds\right)^p \quad \text{by dominated convergence}
\end{aligned} \tag{79}$$

Substituting (79) inside (78) yields:

$$\|V(\tau, X(\tau))\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} \leq \left\|\phi(\tau)^{-1}\right\|_{\mathcal{L}^{q_2}(\mathbb{P};\mathbb{R})} \left(\|V(0, X_0)\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} + \int_0^T \left\|\mathbb{1}_{[0,\tau]}(s)\phi(s)\beta(s)\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} ds\right) \tag{80}$$

which is exactly the desired right side of (64). This concludes the proof of Theorem 3. ∎

The proof of Theorem 4, requires use of the following result from [6].

Lemma 7 (Bound on Supremum and Infimum Processes)**.** Suppose that X is a supermartingale process with $X_0 = 0$ a.s., and denote

$$M^+ := \sup_{t\in[0,T]} \max\{X(t), 0\}, \qquad M^- := \inf_{t\in[0,T]} \min\{X(t), 0\} \tag{81}$$

Then for all $p \in (0, 1)$, the following inequality holds:

$$\left\|M^+\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} \leq c_p \left\|M^-\right\|_{\mathcal{L}^p(\mathbb{P};\mathbb{R})} \tag{82}$$

where the constant $c_p > 0$ is defined as

$$c_p := \left(\left(\frac{1}{p} - 1\right)^p - \int_{\frac{1}{p}-1}^{\infty} \frac{s^{p-1}}{s+1} ds\right)^{\frac{1}{p}} \tag{83}$$

We also require the following generalized triangle inequality for exponents less than 1.

Lemma 8 (Generalized Triangle Inequality)**.** Let $p \in (0,1)$ and $x, y \in \mathbb{R}^+$. Then

$$(x+y)^p \leq x^p + y^p$$

Proof. Recall that if $z \in (0,1)$, then $z < z^p$. Hence,

$$1 = \frac{x}{x+y} + \frac{y}{x+y} \leq \frac{x^p}{(x+y)^p} + \frac{y^p}{(x+y)^p} = \frac{x^p + y^p}{(x+y)^p} \tag{84}$$

Multiplying both sides by $(x+y)^p$ gives the desired result. ■

We are now ready to prove Theorem 4.

Proof of Theorem 4. For any $\varepsilon \in \mathbb{R}^+$, define the martingale process $M : [0,T] \times \Omega \to \mathbb{R}$ to be

$$M^\varepsilon(t) := \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon + V(s,X(s)))^{p-1}\partial_x V(s,X(s))\sigma(s)dW(s) \tag{85}$$

where $\phi(s)$ is defined in (65).

Define $q_1, q_2, q_3 \in \mathbb{R}^+$ such that $q_3 < p$ and $\frac{1}{q_1} = \frac{1}{q_2} + \frac{1}{q_3}$. Further assume WLOG

$$\left\| V(0,X_0) + \int_0^\tau \beta(s)\phi(s)ds \right\|_{\mathcal{L}^{q_3}(\mathbb{P};\mathbb{R})} < \infty$$

Define the sequence of stopping times τ_n as before in (70) of the proof to Theorem 3. We also use the result of the Itô's formula obtained from (74). Now consider, for $q \in (0,1)$:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^p\phi^p(t)|^q\right] \leq \mathbb{E}\left[\sup_{t\in[0,T]} |(\varepsilon + V(t\wedge\tau_n, X(t\wedge\tau_n)))^p\phi^p(t\wedge\tau_n)|^q\right] \\
&\leq \mathbb{E}\left[\sup_{t\in[0,T]}\left|(\varepsilon + V(0,X_0))^p + M^\varepsilon(t\wedge\tau_n) + \int_0^t \mathbb{1}_{[0,\tau_n]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right|^q\right] \\
&\leq \mathbb{E}\left[\sup_{t\in[0,T]}|M^\varepsilon(t\wedge\tau_n)|^q + \sup_{t\in[0,T]}\left|(\varepsilon + V(0,X_0))^p + \int_0^t \mathbb{1}_{[0,\tau_n]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right|^q\right] \\
&\leq \mathbb{E}\left[\sup_{t\in[0,T]}|M^\varepsilon(t\wedge\tau_n)|^q\right] + \mathbb{E}\left[\left|(\varepsilon + V(0,X_0))^p + \int_0^{\tau_n} \phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right|^q\right]
\end{aligned} \tag{86}$$

where the second-to-last inequality follows from Lemma 8 and the property of supremum where $\sup_t(x(t)+y(t)) \leq \sup_t x(t) + \sup_t y(t)$ for nonnegative, real-valued functions x, y.

Note that we can apply Lemma 7 to the first term of (86)

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|M^\varepsilon(t\wedge\tau_n)|^q\right] &\leq \mathbb{E}\left[\left(\sup_{t\in[0,T]}\max\{M^\varepsilon(t\wedge\tau_n),0\}\right)^q\right] \\
&\leq \left(\left(\frac{1-q}{q}\right)^q - \int_{\frac{1-q}{q}}^\infty \frac{s^{q-1}}{s+1}ds\right)\mathbb{E}\left[\left(\inf_{t\in[0,T]}\min\{M^\varepsilon(t\wedge\tau_n),0\}\right)^q\right]
\end{aligned}$$

$$= -\left(\left(\frac{1-q}{q}\right)^q - \int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds\right) \mathbb{E}\left[\left(\sup_{t\in[0,T]} \max\{-M^\varepsilon(t\wedge\tau_n), 0\}\right)^q\right]$$
$$\leq \left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds\right) \mathbb{E}\left[\left(\sup_{t\in[0,T]} \max\{-M^\varepsilon(t\wedge\tau_n), 0\}\right)^q\right] \tag{87}$$

where the second-to-last equality comes from the fact that $\inf(A) = -\sup(-A)$ for any bounded set $A \subset \mathbb{R}$, and the last inequality comes from the fact that $((1-q)/q)^q$ is positive for $q \in (0,1)$. The constant that is obtained in [4] is different:

$$\left(\frac{1}{q}\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q}}{s+1} ds\right) \mathbb{E}\left[\left(\sup_{t\in[0,T]} \max\{-M^\varepsilon(t\wedge\tau_n), 0\}\right)^q\right]$$

instead.

Note that by the nonnegativity of β, σ, V, we can claim that the right side of (74) from the proof of Theorem 3 is nonnegative. Rearranging the resulting expression yields:

$$0 \leq (\varepsilon + V(0,X_0))^p + M^\varepsilon(t\wedge\tau_n) + \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds$$
$$\implies \max\{-M^\varepsilon(t\wedge\tau_n),0\} \leq (\varepsilon + V(0,X_0))^p + \int_0^t \mathbb{1}_{[0,\tau]}(s)\phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds \tag{88}$$

for all $\varepsilon \in \mathbb{R}^+, n \in \mathbb{N}^+$ ad $t \in [0,T]$. Applying (88) to (87), then substituting (87) back into (86) yields:

$$\mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^p\phi^p(t)|^q\right]$$
$$\leq \left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds + 1\right) \mathbb{E}\left[\left|(\varepsilon+V(0,X_0))^p + \int_0^{\tau_n} \phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right|^q\right] \tag{89}$$

Now we can simplify the expected value expression in (89). For all $n \in \mathbb{N}^+, \varepsilon \in \mathbb{R}^+, q \in (0,1)$:

$$\mathbb{E}\left[\left|(\varepsilon+V(0,X_0))^p + \int_0^{\tau_n} \phi^p(s)p(\varepsilon+V(s,X(s)))^{p-1}\beta(s)ds\right|^q\right]$$
$$\leq \mathbb{E}\left[\left(\sup_{s\in[0,\tau_n]} (\varepsilon+V(s,X(s)))^{(p-1)q}\phi^{(p-1)q}(s)\right)\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n}\phi(s)\beta(s)ds\right|^q\right]$$
$$= \mathbb{E}\left[\left(\sup_{s\in[0,\tau_n]} (\varepsilon+V(s,X(s)))^{pq}\phi^{pq}(s)\right)^{\frac{p-1}{p}}\left(\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n}\phi(s)\beta(s)ds\right|^{pq}\right)^{\frac{1}{p}}\right] \tag{90}$$

since we essentially pull out a factor of $(\varepsilon + V(s,X(s)))^{p-1}\phi^{p-1}(s)$, then use the nonnegativity of V to bound it above by the supremum value. There is a mistake in the version of [4]: a missing factor of p next to $\int_0^{\tau_n}\phi(s)\beta(s)ds$.

By Hölder's inequality (Lemma 4),

$$(90) \leq \mathbb{E}\left[\sup_{s\in[0,\tau_n]} (\varepsilon+V(s,X(s)))^{pq}\phi^{pq}(s)\right]^{\frac{p-1}{p}} \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n}\phi(s)\beta(s)ds\right|^{pq}\right]^{\frac{1}{p}} \tag{91}$$

Combining (91) with (89) yields:

$$\left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds + 1\right)^{-1} \mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^p \phi^p(t)|^q\right]$$
$$\leq \mathbb{E}\left[\sup_{s\in[0,\tau_n]} (\varepsilon + V(s,X(s)))^{pq} \phi^{pq}(s)\right]^{\frac{p-1}{p}} \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n} \phi(s)\beta(s)ds\right|^{pq}\right]^{\frac{1}{p}} \tag{92}$$

Recall by the theorem hypothesis that $q_3 < p$. In (92) above, we choose the specific $q \in (0,1)$ such that $q = q_3/p$. Rearranging (92), then taking an exponent of $1/q$ across the entire inequality gives us:

$$\left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds + 1\right)^{-1} \mathbb{E}\left[\sup_{s\in[0,\tau_n]} (\varepsilon + V(s,X(s)))^{pq} \phi^{pq}(s)\right]^{\frac{1-p}{p}} \mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^{pq} \phi^{pq}(t)|\right]$$
$$\leq \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n} \phi(s)\beta(s)ds\right|^{pq}\right]^{\frac{1}{p}}$$
$$\Rightarrow \left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds + 1\right)^{-1} \mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^{pq} \phi^{pq}(t)|\right]^{\frac{1}{p}} \leq \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n} \phi(s)\beta(s)ds\right|^{pq}\right]^{\frac{1}{p}}$$
$$\Rightarrow \left(\int_{\frac{1-q}{q}}^{\infty} \frac{s^{q-1}}{s+1} ds + 1\right)^{-\frac{1}{q}} \mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^{pq} \phi^{pq}(t)|\right]^{\frac{1}{pq}} \leq \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n} \phi(s)\beta(s)ds\right|^{pq}\right]^{\frac{1}{pq}} \tag{93}$$

Now consider the left side of our desired inequality (66). The steps follow in much the same way as in the proof of Theorem 3.

$$\left\|\sup_{t\in[0,\tau]} V(t,X(t))\right\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})} \leq \left\|\phi(\tau)^{-1} \sup_{t\in[0,\tau]} V(t,X(t))\phi(t)\right\|_{\mathcal{L}^{q_1}(\mathbb{P};\mathbb{R})}$$
$$\leq \left\|\phi(\tau)^{-1}\right\|_{\mathcal{L}^{q_2}(\mathbb{P};\mathbb{R})} \left\|\sup_{t\in[0,\tau]} V(t,X(t))\phi(t)\right\|_{\mathcal{L}^{q_3}(\mathbb{P};\mathbb{R})} \quad \text{by Hölder inequality} \tag{94}$$

By condition 2 of Assumption 1,

$$\left\|\sup_{t\in[0,\tau]} V(t,X(t))\phi(t)\right\|_{\mathcal{L}^{q_3}(\mathbb{P};\mathbb{R})}$$
$$\leq \lim_{\varepsilon\to 0} \mathbb{E}\left[\lim_{n\to\infty} \sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^{q_3} \phi^{q_3}(t)|\right]^{\frac{1}{q_3}} \quad \text{by monotone convergence}$$
$$\leq \lim_{\varepsilon\to 0} \lim_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,\tau_n]} |(\varepsilon + V(t,X(t)))^{q_3} \phi^{q_3}(t)|\right]^{\frac{1}{q_3}} \quad \text{by dominated convergence}$$
$$\leq \lim_{\varepsilon\to 0} \lim_{n\to\infty} \left(\int_{\frac{p-q_3}{q_3}}^{\infty} \frac{s^{\frac{q_3}{p}-1}}{s+1} ds + 1\right)^{\frac{p}{q_3}} \mathbb{E}\left[\left|\varepsilon + V(0,X_0) + p\int_0^{\tau_n} \phi(s)\beta(s)ds\right|^{q_3}\right]^{\frac{1}{q_3}}$$
$$= \left(\int_{\frac{p-q_3}{q_3}}^{\infty} \frac{s^{\frac{q_3}{p}-1}}{s+1} ds + 1\right)^{\frac{p}{q_3}} \mathbb{E}\left[\left|V(0,X_0) + p\int_0^{\tau} \phi(s)\beta(s)ds\right|^{q_3}\right]^{\frac{1}{q_3}} \tag{95}$$

where the second-to-last inequality follows from applying the chosen value of $q = q_3/p$ to (93). Substituting (95) back into (94) yields exactly (66), and so we are done with the proof. ■

References

[1] H. Ye, J. Gao, and Y. Ding, "A generalized Gronwall inequality and its application to a fractional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 328, no. 2, pp. 1075–1081, 2007.

[2] Q.-C. Pham, "A variation of Gronwall's lemma," *ArXiv preprint, arXiv:0704.0922v3*, Jan 2009.

[3] D. S. Clark, "Short proof of a discrete Gronwall inequality," *Discrete Applied Mathematics*, vol. 16, no. 3, pp. 279–281, 1987.

[4] A. Hudde, M. Hutzenthaler, and S. Mazzonetto, "A stochastic Gronwall inequality and applications to moments, strong completeness, strong local Lipschitz continuity, and perturbations," *ArXiv preprint, arXiv:1903.08727*, Mar 2019.

[5] M. Scheutzow, "A stochastic Gronwall lemma," *ArXiv preprint, arXiv:1304.5424*, Apr 2013.

[6] R. Bañuelos and A. Osekowski, "Sharp maximal L^p-estimates for martingales," *Illinois Journal of Mathematics*, vol. 58, no. 1, pp. 149 – 165, 2014.