SIAM J. FINANCIAL MATH.
Vol. 8, pp. 1–27

Weak Convergence of Path-Dependent SDEs in Basket Credit Default Swap Pricing with Contagion Risk*

Yao Tung Huang†, Qingshuo Song‡, and Harry Zheng§

Abstract. We investigate computational aspects of basket credit default swap pricing with counterparty credit risk under a multiname contagion model. This model enables us to capture systematic volatility increases in the market triggered by particular bankruptcies. A drawback of this model is its analytical intractability due to a combination of path-dependent coefficients and a path-dependent functional, which furthermore causes potential failure of convergence of numerical approximations under standing assumptions. In this paper, we find sufficient conditions for the desired convergence of functionals associated with approximated solution of certain path-dependent stochastic differential equations.

Key words. path-dependent SDE, weak convergence, correlated first-passage times, basket CDS, contagion risk, counterparty risk

AMS subject classifications. 60F05, 60H30, 60J60, 91G40

DOI. 10.1137/15M1052329

1. Introduction.

It is well known that the first-passage time of a drifting Brownian motion crossing a deterministic level is an inverse Gaussian random variable, as its running maximal process of a drifting Brownian motion can be characterized with the reflection principle and Girsanov theorem. This result has been applied, among other cases, in studying the default time of a firm with a structural framework in which a firm defaults at the first time when its asset value falls below its liability value. The structural model is acclaimed to have a strong economic foundation and is one of the most popular models used in pricing single-name credit derivatives.

It is natural to ask whether one can also characterize the joint distribution of first-passage times of correlated drifting Brownian motions crossing some deterministic levels. When there are two correlated Brownian motions, [10, 18] find the joint distribution of first-passage times, which is an infinite sum of modified Bessel functions of the first kind. Little is known for first-passage times of three or more correlated Brownian motions. This limitation makes it difficult to study the default times of multiple names with a first-passage time structural model and is one of the reasons that intensity-based reduced-form models are used in pricing

*Received by the editors December 11, 2015; accepted for publication (in revised form) October 24, 2016; published electronically January 5, 2017.
http://www.siam.org/journals/sifin/8/M105232.html
Funding: This research supported by the RGC through CityU grant 109613.
†Magnum Research Limited, Entrepreneurship Center, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (don.huang@magnumwm.com).
‡Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong (qingsong@cityu.edu.hk).
§Department of Mathematics, Imperial College, London SW7 2AZ, UK (h.zheng@imperial.ac.uk).

portfolio credit derivatives. After all, intensity models give more analytic tractability but do not provide economic reasons why firms default (see [5, 26] and the references therein for some recent results in this direction).

In this paper we discuss the pricing of a basket credit default swap (CDS) with counterparty (CDS writer) risk. The default times of names in a reference portfolio and that of the counterparty are modeled by a first-passage time structural model and correlated Brownian motions. Furthermore, we include contagion risk in the model, in which the default of a name in the reference portfolio causes a jump increase of volatility of the counterparty, which increases the default probability of the counterparty. The aforementioned model, for instance [6], incorporates both counterparty risk and contagion risk, whereby it can realistically explain the severe difficulties experienced by some seemingly default-remote banks underwriting super senior tranche collateralized debt obligations during the financial crisis of 2007–08.

Since the joint distribution of correlated default times in a structural model is unknown, we compute the price of a basket CDS by the Euler scheme for SDEs and the Monte Carlo simulation. As an effective computational tool, the Euler scheme has been widely adopted in credit risk computing for its simplicity and robustness. Is it possible that a price computed from these "robust" algorithms could actually be overvalued in the trillions-of-dollars market?

As we will illustrate in Example 9, the answer is affirmative in general. Then, is there a set of broad conditions that ensures the computation going to the correct value? In this paper, we aim to reveal the reasons for the possible mispricing and further to provide a rigorous justification on the mispricing scenarios. To illustrate the idea, we take a simplified example which indeed motivates our general setting in section 2.

Example 1. Let V_0 be the firm value process of the CDS writer and $\{V_i : i = 1, \dots, k\}$ those of k basket reference names. We assume that the default time of the firm i is given by $\tau_i = \inf\{t > 0 : V_i(t) \leq 1\}$, $i = 0, 1, \dots, k$. Moreover, we assume the volatility of firm i follows $\bar{\sigma}_i \cdot (1 + \alpha(t))$, where

$$\alpha(t) = \textstyle\sum_{j=1}^{k} I(\tau_j \leq t) \tag{1}$$

is the total default number of reference names by time t. The main interest of this paper is to compute a premium rate $\hat{c}$ per annum in the form of

$$\hat{c} = \frac{\mathbb{E}[f_1(\tau_0, \dots, \tau_k)]}{\mathbb{E}[f_2(\tau_0, \dots, \tau_k)]} \tag{2}$$

for some discontinuous functions f_1 and f_2 with domain $\mathbb{R}^{k+1}$; see also the exact formula (4).

Given that Euler scheme with step size h in the approximation of the underlying firm values and further their default times, denoted by V_i^h and τ_i^h, respectively, one can approximate premium rate $\hat{c}$ by replacing τ_i by τ_i^h in formula (2), denoted by $\hat{c}^h$. Our question is whether the convergence $\hat{c}^h \to \hat{c}$ holds as $h \to 0$.

Regarding the numerical schemes of SDEs, there has been extensive research on the topics of both weak and strong convergences; see [7, 8, 12, 16, 24] and the references therein for excellent expositions. In particular, the book [12] introduces a systematic and rigorous treatment to the numerical approximation of the various types of SDEs. To the best of our knowledge, essentially all the above literature on numerical SDEs studies the convergence at fixed grid points in Markovian settings. Back to Example 1, the volatility is path-dependent and the presence of first-passage times requires information on the whole path.

In this regard, we turn our attention to studying associated mappings from Skorohod path space $\mathbb{D}$ to real numbers. For instance, the first-passage time τ_i can be rewritten by $\tau_i = \pi(V_i)$, where $\pi : \mathbb{D} \to \mathbb{R}$ is a mapping defined by $\pi(x) = \inf\{t > 0 : x(t) \leq 1\}$. A similar idea can be applied to rewrite the premium rate by $\hat{c} = \mathbb{E}[F_1(V)]/\mathbb{E}[F_2(V)]$, where each F_i is a functional on $\mathbb{R}^{k+1}$-valued path space $\mathbb{D}^{k+1}$ with its argument $V = (V_i)_{i=0,\dots,k} \in \mathbb{D}^{k+1}$. To this end, our work can be clearly divided into two steps: the convergence $\hat{c}^h \to \hat{c}$ shall be true by the continuous mapping theorem (CMT) if

(Q1) V^h converges to V in distribution (see section 2.3.3), denoted by $V^h \Rightarrow V$;

(Q2) F_1 and F_2 are continuous almost surely at V with respect to Skorohod topology.

Regarding (Q1), the weak limit theorems for the whole path in a non-Markovian setting can be found in [13], and our approach for the weak convergence $V^h \Rightarrow V$ is also closely related to [13]. However, [13] establishes the convergence based on the continuity assumption of coefficient functions, while the volatility of Example 1 is not continuous as a mapping on a path space, and hence their result cannot be directly applied here. The main reason for the discontinuity is due to the dependence on the number of defaults α; see also the tangency problem of [14]. Nonetheless, α as a function on a path space is almost surely continuous with respect to the probability induced by V; see Example 10 for details. As such, we bravely attempt to show the desired weak convergence under the almost sure continuity assumption of coefficient functions:

(H) As a mapping from path space $\mathbb{D}^{k+1}$ to $\mathbb{R}$, the function α is continuous under Skorohod topology almost surely with respect to $\mathbb{P}V^{-1}$.

In the above, $\mathbb{P}V^{-1}$ refers to the probability measure on $\mathbb{D}^{k+1}$ induced by V; see a further explanation in section 2.3.3. Although (H) is enough for our purpose to cover our motivated example, it is still inappropriate since the unknown solution V shall not be included in the assumption. This leads to Assumption 4, which serves the same role as (H) with the help of Assumption 3.

To this end, it is inevitable to go through the entire procedure and carefully reexamine all the necessary steps in the weak convergence. First, we show the tightness of the discrete Euler processes and deduce the convergence of approximating processes to a limiting process almost surely by the Skorohod representation theorem. Second, we claim the continuity of the limiting process, which plays a crucial role in the proof; see Remark 13. Finally, we complete the proof by showing that the limiting process is the weak solution of the underlying SDE.

Regarding (Q2), provided the completion of (Q1), we shall show the convergence in distribution $f_i(\tau_0^h, \dots, \tau_k^h) \Rightarrow f_i(\tau_0, \dots, \tau_k)$ of (2). Although the form of f_i corresponding to the pricing formula (4) is complicated, it's enough to examine the weak convergence on the following two simple quantities by the CMT:

1. (The convergence in single name risk) One shall verify $I(\tau_i^h > t) \Rightarrow I(\tau_i > t)$ for arbitrary i and t. Applying CMT, it's sufficient to show $\mathbb{P}(\tau_i = t) = 0$ for any t, i.e., all underlying firms have zero probability to get into default at a particular time. This is guaranteed by nondegeneracy Assumption 3.
2. (The convergence in counterparty risk) One shall verify $I(\tau_0^h > \tau_i^h) \Rightarrow I(\tau_0 > \tau_i)$ for each $i \geq 1$. Again by CMT, it's enough to show $\mathbb{P}\{\tau_0 = \tau_i\} = 0$, i.e., there shall be zero probability that two companies default simultaneously. It boils down to two

sufficient conditions in turn: if either (a) a CDS writer is independent to all reference names or (b) a CDS writer is not perfectly correlated to all reference names and the volatilities are all piecewise constants, then the above convergence holds.

To close the introduction, our contribution is summarized as follows. We establish the convergence for the approximation of CDS pricing; see Theorem 7. As illustrated above, the major mathematical difficulty compared to the existing literature stems from the following added features in our model: (a) the contagion risk due to the dependence of the volatility in total number of defaults; (b) single/counterparty default risk in terms of first-passage times. As for any other weak limit results on the path space, our result has to be established by many assumptions, for which we add careful explanations of why they are needed; see section 2.3. As a result, the seemingly cumbersome assumptions can be reduced to a simple condition in Example 1: If $\bar{\sigma}_i > 0$ for all i and V_0 is not perfectly correlated to any of $\{V_i, i = 1, \dots k\}$, then $\hat{c}^h \to \hat{c}$ holds.

The paper is organized as follows. Section 2 presents the problem setup and our main result, which is to cover rather general scenarios than Example 1. Note that some notions used in the introduction are also slightly extended in an obvious way. Section 3 includes the technical proof of the main result. Section 4 includes further discussions related to this work.

2. Main results.

2.1. Problem setting. Let $\hat{T} = T + 1$ for some positive constant T. Denote by $\mathbb{D}^n$ the space of càdlàg functions, i.e., functions that are right continuous with left limits defined on $[0, \hat{T}]$ taking values in $\mathbb{R}^n$. $\mathbb{D}^1$ is abbreviated as $\mathbb{D}$. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, on which a standard $k+1$ dimensional Brownian motion $W = (W_0, W_1, \dots, W_k)^T$ is defined. Here, x^T is the transpose of x. Suppose that $\mu(\cdot,\cdot) : \mathbb{D}^{k+1} \times [0,T] \mapsto \mathbb{R}^{k+1}$ and $\sigma(\cdot,\cdot) : \mathbb{D}^{k+1} \times [0,T] \mapsto \mathbb{R}^{(k+1)\times(k+1)}$ are nonanticipating in the sense that $\mu(x,t) = \mu(x(\cdot \wedge t), t)$ and $\sigma(x,t) = \sigma(x(\cdot \wedge t), t)$ for all $t \geq 0$ and $x \in \mathbb{D}^{k+1}$ (see [13] for details). We consider $k+1$ companies with the firm value processes $V := (V_0, V_1, \dots, V_k)^T$ satisfying

$$dV(t) = \operatorname{diag}(V(t)) \left(\mu(V,t)\, dt + \sigma(V,t)\, dW(t)\right), \quad t \geq 0, \tag{3}$$

where the constant $V(0)$ is the firm value at $t = 0$, $\operatorname{diag}(V(t))$ is a $(k+1)\times(k+1)$ diagonal matrix with diagonal elements $V_i(t)$, $i = 0, 1, \dots, k$, and $\mu = (\mu_0, \mu_1, \dots, \mu_k)^T$ and $\sigma = (\sigma_{ij})_{0 \leq i,j \leq k}$ represent asset appreciation and volatility rates, respectively. The product of σ and its transpose reflects the covariance between the movements in the asset values of firms, thus playing a critical role in determining the dependence structure among the firm values.

Let ϕ stand for μ and σ. We impose the following structure to $\phi = \mu, \sigma$: For any $x \in \mathbb{D}^{k+1}$, $\phi(x,\cdot)$ can be decomposed into a continuous deterministic process $\phi_c(\cdot)$ and a pure jump process $\phi_J(x,\cdot)$ as follows:

$$\phi(x,t) = \phi_c(t) + \phi_J(x,t) = \phi_c(t) + \sum_{i=1}^{\mathcal{N}^\phi(x,t)} J_i^\phi(x), \quad 0 \leq t \leq \hat{T},$$

where functions μ_J (respectively, σ_J) : $\mathbb{D}^{k+1} \times [0,T] \mapsto \mathbb{R}^{k+1}$ (respectively, $\mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$) and $\mathcal{N}^\phi : \mathbb{D}^{k+1} \times [0,T] \mapsto \mathbb{N}$ are measurable and nonanticipating in the sense that

$\phi_J(x,t) = \phi_J(x(\cdot \wedge t), t)$ and $\mathcal{N}^\phi(x,t) = \mathcal{N}^\phi(x(\cdot \wedge t), t)$. The function J_i^μ (respectively, J_i^σ) : $\mathbb{D}^{k+1} \mapsto \mathbb{R}^{k+1}$ (respectively, $\mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$) is measurable for $i = 1, 2, \ldots$. In the above, $\mathcal{N}^\phi(x,t)$ and $J_i^\phi(x)$ represent the number of jumps of ϕ up to time t and the jump size for the ith jump, respectively. Moreover, the processes $t \mapsto \phi(V,t)$ and $t \mapsto \mathcal{N}^\phi(V,t)$ are $\{\mathcal{F}_t\}_{t\geq 0}$-adapted.

Without loss of generality, let V_0 be the firm value process of a CDS writer and V_i, $i = 1, 2, \ldots, k$, be the firm value processes of the companies in the reference portfolio. Each company has an exponential default barrier L_i in the form of

$$L_i(t) = K_i e^{\gamma_i t}, \quad t \geq 0, i = 0, 1, \ldots, k,$$

where γ_i and K_i are nonnegative constants. The default time for company i is defined by

$$\tau_i := \inf\{t > 0 : V_i(t) \leq L_i(t)\} \wedge \hat{T}, \ i = 0, 1, \ldots, k,$$

the first time the firm value falls below the default barrier. We denote by $\{\tau_{(i)} : i = 1, 2, \ldots, k\}$ the order statistics of $\{\tau_i : i = 1, \ldots, k\}$, i.e., $\tau_{(i)}$ is the ith default time among k companies in the reference portfolio.

Note that the definition of the default time τ_i is truncated by $\hat{T} = T + 1$ but not by the maturity T only for its convenience. The advantage is that our work is reduced to càdlàg space from infinite time interval to a finite time interval, while we can keep the probability of $\{\tau_i = T\}$ as zero under a nondegenerate condition. This feature will be used to show the weak convergence involved with $\mathbf{1}(\tau_{(i)} \leq T)$, while preserving the structure of the swap rate defined in (4).

In pricing derivative securities with the structural model, it is normally assumed that the drift coefficient μ of V in (3) is equal to the risk-free interest rate r in a risk-neutral setting. We do not insist that μ equal to r in this paper as the firm value is not a traded asset and cannot be hedged with the no-arbitrage and martingale representation argument. The firm value process V is a measure that may have a close relation to traded assets but is mainly used to define default events. All results still hold if μ is replaced by r. For the sake of simplicity, the risk-free interest rate is assumed to be a positive constant r. The extension to the stochastic interest rate model can be done under the framework of this paper (see Remark 23 for details).

A basket CDS is an insurance product in which the underlying is a portfolio of defaultable companies and the writer (seller) of the ith default CDS promises to pay $1 - \delta_i$ to the buyer of the insurance at the ith default time $\tau_{(i)}$ if that happens before the maturity time T of the contract, in return the buyer of the ith CDS agrees to pay the writer a premium fee at rate $\hat{c}_i$ per annum on each of prespecified dates $\{0 < t_1 < t_2 < \cdots < t_m = T\}$ as long as the ith default has not occurred. (In fact, the triggering time for the basket CDS does not have to be the default time of a company; it can be any predefined event). If the writer defaults before the maturity of the contract or the ith default time, then the CDS contract terminates and there are no further cash flows. The risk neutral swap rate $\hat{c}_i$ is given by

$$\hat{c}_i = \frac{\mathbb{E}\left[e^{-r\tau_{(i)}}(1-\delta_i)\mathbf{1}\left(\tau_{(i)} \leq T\right)\mathbf{1}\left(\tau_0 > \tau_{(i)} \wedge T\right)\right]}{\mathbb{E}\left[\sum_{j=1}^m e^{-rt_j}\Delta t_j \mathbf{1}\left(\tau_{(i)} > t_j\right)\mathbf{1}\left(\tau_0 > t_j\right)\right]}, \tag{4}$$

where $\Delta t_j = t_j - t_{j-1}$, $j = 1, \ldots, m$, $t_0 = 0$, and $\mathbf{1}(\cdot)$ is the indicator function which equals 1 if an event occurs and 0 otherwise. The values of τ_0 and $\tau_{(i)}$ are dependent on the realized path of the process V. The evaluation of $\hat{c}_i$ in (4) involves the expectations of path-dependent functionals, which may naturally be computed with the Monte Carlo and the Euler approximation method.

Let the time interval $[0, T]$ be partitioned into N equally spaced subintervals with grid points $t_n^h = nh$, $n = 0, \ldots, N$, $h = T/N$, and let V^h, valued in $\mathbb{R}^{k+1}$, be the Euler approximating process for V, defined recursively by

$$V_{n+1}^h := V_n^h + \mathrm{diag}\left(V_n^h\right)\left(\mu_n^h h + \sigma_n^h(W(t_{n+1}^h) - W(t_n^h))\right), \quad n = 0, \ldots, N-1,$$

where $\mu_n^h := \mu(V^h, nh)$ and $\sigma_n^h := \sigma(V^h, nh)$. In the rest of the paper, for ease of the notational complexity without ambiguity, we retain the notation (V^h, μ^h, σ^h) to denote the piecewise constant interpolation of sequences $\{(V_n^h, \mu_n^h, \sigma_n^h) : n = 0, 1, \ldots N\}$, i.e.,

$$V^h(t) = V_n^h, \quad \mu^h(t) = \mu_n^h, \quad \sigma^h(t) = \sigma_n^h, \quad t \in [nh, (n+1)h). \tag{5}$$

Random variables V_n^h are $\mathcal{F}_n^h$-measurable, where $\mathcal{F}_n^h := \mathcal{F}_{t_n^h}$ is the information available at time t_n^h. Since W is a Brownian motion, without changing their distributions, we may generate $\{V_n^h : n = 1, 2, \ldots, N\}$ by the recursive formula

$$V_{n+1}^h = V_n^h + \mathrm{diag}\left(V_n^h\right)\mu_n^h h + \mathrm{diag}\left(V_n^h\right)\sigma_n^h\sqrt{h}Z_{n+1}, \tag{6}$$

where Z_n, $n = 1, \ldots, N$, are independent $k+1$ dimensional standard normal variables and Z_l are independent of the filtration $\mathcal{F}_n^h$ for $l > n$ and $n = 1, \ldots, N-1$.

Corresponding to the Euler approximating process V^h, the annual swap rate $\hat{c}_i^h$ has the same form as that in (4) except τ_0, τ_i, and $\tau_{(i)}$ are replaced by τ_0^h, τ_i^h, and $\tau_{(i)}^h$, respectively. In this paper we explore under what conditions we have

$$\lim_{h \to 0} \hat{c}_i^h = \hat{c}_i. \tag{7}$$

2.2. The main results. In this part, we present the main result after several assumptions.

Assumption 2. *With some positive constant K for all $x \in \mathbb{D}^{k+1}$ and $0 \le t_1, t_2 \le \hat{T}$, for $\phi = b, \sigma$*

$$\begin{aligned} |\phi_c(t_1) - \phi_c(t_2)| &\le K|t_1 - t_2|^{1/2}, \\ \left|\mathcal{N}^\phi(x, T)\right| &\le K, \\ \left|J_i^\phi(x)\right| &\le K \forall i. \end{aligned}$$

Assumption 3. *$\sigma(x,t)$ satisfies the uniform nondegeneracy condition, i.e., $\sigma(x,t)\sigma(x,t)^T \ge \lambda I$ for all $x \in \mathbb{D}^{k+1}$ and $t \in [0, \hat{T}]$ for some $\lambda > 0$.*

We next define $\pi : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}$ as

$$\pi\left(x, l\right) := \inf\left\{t > 0 : x\left(t\right) \leq l\left(t\right)\right\} \wedge \hat{T}, \tag{8}$$

which is the first time of the càdlàg function x hitting the barrier l. Let $\mathbb{C}^n$ be the collection of continuous functions defined on $[0, \hat{T}]$ taking values in $\mathbb{R}^n$. $\mathbb{C}^1$ is abbreviated as $\mathbb{C}$. For $i = 0, 1, \ldots, k$, define two disjoint subsets of the space $\mathbb{C}$ as

$$C_1^i = \left\{x \in \mathbb{C} : \pi\left(x, L_i\right) < T \text{ and } \inf\left\{t > \pi\left(x, L_i\right) : x\left(t\right) < L_i\left(t\right)\right\} = \pi\left(x, L_i\right)\right\}$$

and

$$C_2^i = \left\{x \in \mathbb{C} : \pi\left(x, L_i\right) \geq T\right\}.$$

Assumption 4. *The mappings $x \mapsto \mu(x, \cdot)$ and $x \mapsto \sigma(x, \cdot)$ are continuous under Skorohod topology at the set $\{x \in \mathbb{C}^{k+1} : x_i \in C_1^i \cup C_2^i, i = 0, 1, \ldots, k\}$.*

Assumption 5. *For all $t \geq 0$, $\left(\sigma\left(t\right)\right)_{0i} = \left(\sigma\left(t\right)\right)_{i0} = 0$, a.s. for $1 \leq i \leq k$.*

Assumption 6. *σ is piecewise constant almost surely, i.e., for strictly increasing stopping time sequence $\{\theta_0, \theta_1, \ldots\}$ such that $\theta_0 = 0$ and $\lim_{n\to\infty} \theta_n = \hat{T}$, the process σ is in the following form:*

$$\sigma\left(t\right) = \sum_{i=1}^{\infty} \sigma_i \mathbf{1}\left(\theta_{i-1} \leq t < \theta_i\right) \quad a.s.,$$

where σ_i is a nonsingular $(k+1) \times (k+1)$ matrix, measurable with respect to $\mathcal{F}_{\theta_{i-1}}$, $i = 1, 2, \ldots$.

We now state the main result of the paper, and the explanations are immediately followed by section 2.3. A short remark on convergence in distribution is also included in section 2.3.3.

Theorem 7. *Let V be the $k+1$ dimensional $\mathcal{F}_t$-adapted continuous process of the form (3) and V^h be Euler approximating process for V given by (5). If Assumptions 2, 3, and 4 hold, then V^h converges to V in distribution. If, in addition, either Assumption 5 or 6 holds, then $\lim_{h\to 0} \hat{c}_i^h = \hat{c}_i$ for all $i = 1, 2, \ldots, k$.*

2.3. Discussions of assumptions with its application to Example 1. In this part, we discuss the above assumptions combined with Example 1 to illustrate the main result.

2.3.1. Discussion of Assumption 2. Assumption 2 is imposed to guarantee the existence and uniqueness of a solution V to (3). As one shall further note from Assumption 2, $\phi(x, \cdot), \phi_c(\cdot)$ and $\phi_J(x, \cdot)$ are all bounded by some constant K. (In this paper, K is a generic constant whose value may change at each line.) We may relax the Holder-1/2 continuity of ϕ_c by a condition $|\phi_c(t_1) - \phi_c(t_2)| \leq g(|t_1 - t_2|)$, where g is a bounded function satisfying $g(0) = 0$ and $g(s)$ tends to 0 as s tends to 0.

2.3.2. Skorohod space. To discuss other assumptions, we shall first state the Skorohod metric for $\mathbb{D}^{k+1}$ space and related notions, which are adopted by [1]. Define a uniform metric on $\mathbb{D}^{k+1}$ by

$$\|x - y\| = \sup_{t\in[0,T]} \left|x\left(t\right) - y\left(t\right)\right| \forall x, y \in \mathbb{D}^{k+1}. \tag{9}$$

Let Λ denote the class of strictly increasing, continuous mappings of $[0, \hat{T}]$ onto itself. Then the function space $\mathbb{D}^{k+1}$ is equipped with the Skorohod topology with the metric

$$d(x, y) = \inf_{\lambda \in \Lambda} \{\|\lambda - I\| \vee \|x \circ \lambda - y\|\} \, \forall x, y \in \mathbb{D}^{k+1}. \tag{10}$$

Since $d(x, y) \leq \|x - y\|$ the convergence in Skorohod topology does not imply the convergence in uniform topology. However, if the limit is in $\mathbb{C}^{k+1}$, then they are equivalent.

Proposition 8 (see [1, p. 124]). *Elements x_n of $\mathbb{D}^{k+1}$ converge to a limit x in the Skorohod topology if and only if there exist functions λ_n in Λ such that $\lim_n x_n(\lambda_n t) = x(t)$ uniformly in t and $\lim_n \lambda_n t = t$ uniformly in t. Moreover, if x is in $\mathbb{C}^{k+1}$, then Skorohod convergence implies uniform convergence.*

2.3.3. Convergence in distribution. The notion of the *convergence in distribution* (see [1]) or *weak convergence* (see [14]) may be defined in various ways in the literature. Since it plays an important role in Theorem 7 and throughout the paper, we give a short remark for its clarification. The random elements of our interests V^h and V of Theorem 7 are the maps from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to a Skorohod metric space $(\mathbb{D}^{k+1}, d)$ equipped with Skorohod metric d defined in (10), and we often write it as

$$V^h, V : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto (\mathbb{D}^{k+1}, d)$$

or $V^h, V : \Omega \mapsto \mathbb{D}^{k+1}$ in short if the context has no ambiguity for $\mathcal{F}, \mathbb{P}$, and d. The *distribution* of the random element V refers to $\mathbb{P}V^{-1}$, which is indeed the probability measure on the Borel σ-algebra $\mathcal{B}(\mathbb{D}^{k+1})$, i.e.,

$$\mathbb{P}V^{-1}(A) = \mathbb{P}\{\omega : V(\omega) \in A\} \; \forall A \in \mathcal{B}(\mathbb{D}^{k+1}).$$

$\mathbb{P}V^{-1}$ is sometimes called the law of V, or the probability measure induced by V, or the push forward measure. Similarly, $\mathbb{P}(V^h)^{-1}$ is the distribution of V^h.

We say, as $h \to 0$, V^h converges to V in distribution with respect to Skorohod metric d if the measure $\mathbb{P}(V^h)^{-1}$ weakly converges to $\mathbb{P}V^{-1}$, denoted by $\mathbb{P}(V^h)^{-1} \Rightarrow \mathbb{P}V^{-1}$. It sometimes is said V^h converges to V weakly with respect to the Skorohod metric. By the portmanteau theorem, one can equivalently define the convergence in distribution in any of five different ways provided on p. 26 of [1]. For instance, a common definition adopted in many references is that V^h is said to be convergent to V in distribution with respect to Skorohod metric d if $\mathbb{E}[f(V^h)] \to \mathbb{E}[f(V)]$ for all $f \in C_b(\mathbb{D}^{k+1})$, the space of all bounded continuous (w.r.t. metric d) functions on $\mathbb{D}^{k+1}$.

One may already note that the convergence in distribution relies on the topology of $\mathbb{D}^{k+1}$. Indeed, $C_b(\mathbb{D}^{k+1})$ under uniform topology induced by $\|\cdot\|$ of (9) is a bigger space than $C_b(\mathbb{D}^{k+1})$ induced by Skorohod metric $d(\cdot, \cdot)$ of (10). This immediately yields by the definition that the convergence in distribution with respect to the uniform topology implies convergence in distribution with respect to Skorohod topology. In the rest of the paper, unless it is specified, the space $\mathbb{D}^{k+1}$ is equipped with Skorohod metric d by default, and convergence in distribution means by default the convergence in distribution with respect to Skorhod topology.

2.3.4. Discussion on Assumption 3. Recall the hitting time operator π of (8). This notion enables us to treat the default time τ_i as a function on a random process, i.e., $\tau_i = \pi(V_i, L_i)$. However, one cannot assume $\pi(\cdot, L_i)$ is continuous in $\mathbb{D}$ in general from the following example.

Example 9. $\pi(x, 0)$ is not upper semicontinuous at $x \in \mathbb{C}$ given by

$$x(t) = |t - 1/2|,$$

since $\lim_n \pi(x_n, 0) = \hat{T} > 1/2 = \pi(x, 0)$, where $x_n = x + 1/n$. See Figure 1 for an illustration.

One can also adapt the above idea to illustrate the potential issue arising from the numerical computation of the credit risk model. Let's assume that the firm value V_1 follows the deterministic curve $x + 1$, i.e., $V_1(t) = x(t) + 1 = 1 + |t - 1/2|$, and a tradable derivative price is given by $\hat{c} = \mathbb{E}[\tau_1]$, where $\tau_1 = \inf\{t : V_1(t) \le 1\} \wedge \hat{T}$. One can easily see that an approximation $\hat{c}^h$ by the usual Euler scheme does not guarantee the convergence $\hat{c}^h \to c^h$. Obviously, it is not a realistic example due to its zero volatility. Then, can we avoid mispricing from the computation by assuming nonzero volatility of the firm value?

The nondegenerate condition Assumption 3 is very important throughout the paper. It not only implies any two of firms V_i and V_j are not perfectly correlated but also makes the barriers regular to the underlying diffusion. As an immediate consequence, we have $x \mapsto \pi(x, L_i)$ continuous $\mathbb{P}V_i^{-1}$-almost surely for all i's.

Another use of Assumption 3 is for the total number of defaults α in Example 1. One can rewrite this α in terms of π by

$$\alpha(t) = \sum_{j=1}^{k} I(\pi(V_i, L_i) \le t) := \hat{\alpha}(V) \tag{11}$$

for some mapping $\hat{\alpha}$ defined on $\mathbb{D}^{k+1}$. Example 9 implies that the functional $\hat{\alpha}$ above is not continuous with respect to Skorohod topology, and hence it violates the sufficient condition for the weak convergence given in Condition C5.1 of [13]. However, Example 10 below can verify the almost sure continuity under Assumption 3, which is enough for our purpose.

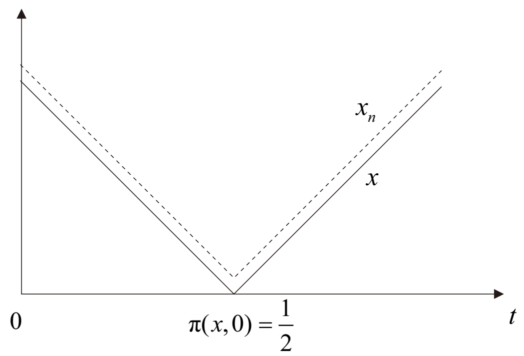

Figure 1. *Illustration of Example 9.*

Example 10. The number of defaults α defined in Example 1 can be rewritten by (11). Example 9 implies that $\hat{\alpha}$ is not continuous in general. However, under Assumption 3, $\hat{\alpha}$ is continuous under Skorohod topology almost surely with respect to $\mathbb{P}V^{-1}$. Indeed, this follows from the following two facts:

1. By the Blumenthal 0-1 law (see [3]), $\pi(\cdot, L_i)$ is continuous with respect to the Skorohod metric almost surely in $\mathbb{P}V_i^{-1}$ for all $i = 0, \dots, k$, i.e.,

$$\mathbb{P}(V_i \in \{x \in \mathbb{D} : \pi(\cdot, L_i) \text{ is continuous at } x\}) = 1.$$

2. Moreover, $I(\pi(\cdot, L_j) \le t)$ is also continuous with respect to the Skorohod metric almost surely in $\mathbb{P}V_j^{-1}$, the probability induced by V_j.

2.3.5. Discussion of Assumption 4. Condition C5.1 in [13] requires the mapping $x \mapsto \sigma(x, \cdot)$ is continuous under Skorohod topology in $\mathbb{D}^{k+1}$, while our volatility violates this condition as of Example 10. In contrast, Assumption 4 may be regarded as a continuity requirement of coefficient functions on a smaller space $\{x \in \mathbb{C}^{k+1} : x_i \in C_1^i \cup C_2^i, i = 0, 1, \dots, k\}$; Figure 2 provides some examples to illustrate the concept of C_1^i and C_2^i. The paths in the union set of C_1^i and C_2^i are regular with respect to the boundary L_i in the sense that once the path touches the boundary L_i the path pushes through the boundary. Continued from Example 10, due to the fact $\mathbb{P}(V_i \in C_1^i \cup C_2^i) = 1$ under Assumption 3, α also satisfies Assumption 4. We refer to section 2.2 of [23] for more detailed descriptions.

2.3.6. Discussion of Assumption 5 or 6. Assumption 5 or 6, together with Assumption 3, ensures the indicator function $\mathbf{1}\left(\tau_0 > \tau_{(i)} \wedge T\right)$ is continuous $\mathbb{P}V^{-1}$-almost surely. Suppose the CDS writer is default free before maturity T; then the indicator function $\mathbf{1}\left(\tau_0 > \tau_{(i)} \wedge T\right)$ is a constant one a.s. $\mathbb{P}V^{-1}$. For such a case, Assumption 5 or 6 is not required in Theorem 7. From the financial perspective, Assumption 5 implies that the CDS writer and the reference names are independent of each other from exogenous factors (W_0 and W_i are independent). This is realistic in practice since an important criteria adopted by practitioners for choosing an

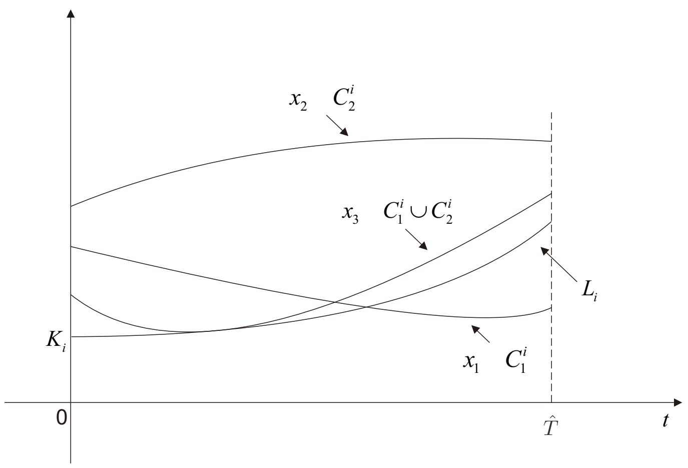

Figure 2. *Illustration of disjoint subsets C_1^i and C_2^i. x_1 and x_2 belong to C_1^i and C_2^i, respectively. x_3 doesn't belong to C_1^i or C_2^i.*

appropriate CDS writer is that the CDS writer has little correlation with the reference names. On the other hand, if the CDS writer has nonzero correlation with reference names (but not perfectly correlated due to Assumption 3), then one can still have almost sure continuity of $\mathbf{1}\left(\tau_0 > \tau_{(i)} \wedge T\right)$ by requiring the piecewise constant form of σ, which is indeed the case in reality. Note that the volatility is calibrated in practice from time to time, not continuously.

2.4. Other operators related to the CDS pricing formula. For later use, we also introduce other related notions here. For any $n \in \mathbb{N}$, define $\mathcal{S}^n : \mathbb{R}^n \times \{1, 2, \ldots, n\} \mapsto \mathbb{R}$ as

$$\mathcal{S}^n (x, j) := \text{the } j\text{th smallest value among } \{x_i\}_{i=1}^n .$$

Finally, define $F_i : \mathbb{D}^{k+1} \mapsto \mathbb{R}$, $i = 1, 2$, as

$$F_1 (x) := e^{-r\mathcal{S}^k\left(\{\pi(x_n, L_n)\}_{n=1}^k, i\right)} (1 - \delta_i) \mathbf{1} \left(\mathcal{S}^k \left(\{\pi (x_n, L_n)\}_{n=1}^k , i\right) \leq T\right)$$
$$\mathbf{1} \left(\pi (x_0, L_0) > \mathcal{S}^k \left(\{\pi (x_n, L_n)\}_{n=1}^k , i\right) \wedge T\right)$$

and

$$F_2 (x) := \sum_{j=1}^m e^{-rt_j} \Delta t_j \mathbf{1} \left(\mathcal{S}^k \left(\{\pi (x_n, L_n)\}_{n=1}^k , i\right) > t_j\right) \mathbf{1} (\pi (x_0, L_0) > t_j) .$$

With the help of mappings π, $\mathcal{S}^k$, and F_i, $i = 1, 2$, the default time τ_i of company i, the ith default time $\tau_{(i)}$ of the reference portfolio, and the swap rates $\hat{c}_i$ and $\hat{c}_i^h$ can be expressed respectively as follows:

$$\begin{aligned}
\tau_i &= \pi(V_i, L_i), \quad i = 0, 1, \ldots, k,\\
\tau_{(i)} &= S^k(\{\pi(V_j, L_j)\}_{j=1}^k, i), \quad i = 1, 2, \ldots, k,\\
\hat{c}_i &= \frac{\mathbb{E}[F_1(V)]}{\mathbb{E}[F_2(V)]} \quad \text{and} \quad \hat{c}_i^h = \frac{\mathbb{E}[F_1(V^h)]}{\mathbb{E}[F_2(V^h)]}.
\end{aligned}$$

The default time for company i is illustrated in Figure 3.

3. Proof of Theorem 7.

In this section we prove Theorem 7 through a number of lemmas.

3.1. Preliminary estimates. We discuss some properties of V^h here, which play a crucial role in the subsequent parts. Let $z^h(t) := [\frac{t}{h}]h$. The Euler scheme can be rewritten in terms of interpolated process in the following way:

$$V^h(t) = V^h_{[\frac{t}{h}]} = V_0 + \int_0^{z^h(t)} \operatorname{diag}(V^h(s))\mu^h(s)ds + \int_0^{z^h(t)} \operatorname{diag}(V^h(s))\sigma^h(s)dW(s).$$

For convenience, we also denote $\Delta V_n^h := V_{n+1}^h - V_n^h$ and $\Delta M_n^h := \Delta V_n^h - E\left[\Delta V_n^h | \mathcal{F}_n^h\right]$.

Lemma 11. *If Assumption 2 holds, then V^h satisfies, for $p \geq 2$ and $h > 0$,*

$$\mathbb{E}\left[\sup_{0\leq t\leq \hat{T}} \left|V^h(t)\right|^p\right] \leq K, \tag{12}$$

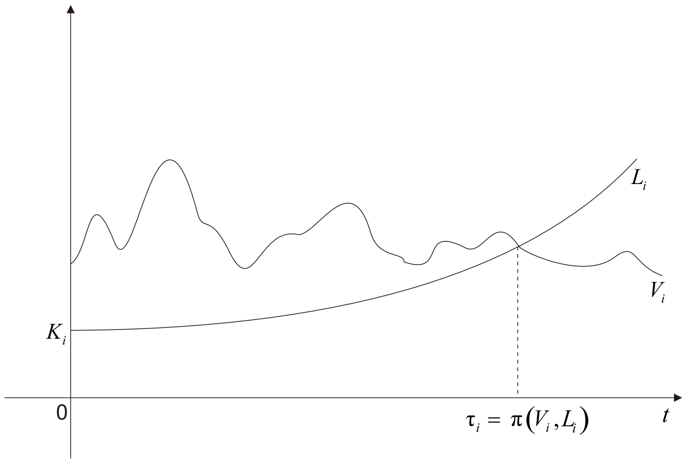

Figure 3. *Illustration of the default time for company i, $\tau_i = \pi(V_i, L_i)$.*

$$\mathbb{E}\left[\sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\sum_{n=[\frac{t_1}{h}]}^{m}\Delta V_n^h\right|^p\right] \le K\left[z^h(t_2)-z^h(t_1)\right] < K(t_2-t_1+h)\quad \forall t_1<t_2, \tag{13}$$

$$\mathbb{E}\left[\sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\sum_{n=[\frac{t_1}{h}]}^{m}\Delta V_n^h\right|^p\middle|\mathcal{F}_{t_1}\right] \le K\mathbb{E}\left[\int_{z^h(t_1)}^{z^h(t_2)}|V^h(s)|^p ds\middle|\mathcal{F}_{t_1}\right]\quad \forall t_1<t_2, \tag{14}$$

where $[x]$ is the largest integer not greater than x.

Proof. For any $t\in[h,\hat{T}]$, $h>0$ and $p\ge 2$, using the Burkholder inequality and Assumption 2 we estimate the pth moment of $V^h(t)$ as follows:

$$\begin{aligned}
\mathbb{E}|V^h(t)|^p &\le \mathbb{E}\left(|V_0| + \left|\int_0^{z^h(t)}\mathrm{diag}(V^h(s))\mu^h(s)ds\right| + \left|\int_0^{z^h(t)}\mathrm{diag}(V^h(s))\sigma^h(s)dW(s)\right|\right)^p\\
&\le K\mathbb{E}\left(|V_0|^p + \left|\int_0^{z^h(t)}\mathrm{diag}(V^h(s))\mu^h(s)ds\right|^p + \left|\int_0^{z^h(t)}\mathrm{diag}(V^h(s))\sigma^h(s)dW(s)\right|^p\right)\\
&\le K + \mathbb{E}\int_0^{z^h(t)}|\mathrm{diag}(V^h(s))\mu^h(s)|^p ds\\
&\quad + K\mathbb{E}\left[\int_0^{z^h(t)}|\mathrm{diag}(V^h(s))\sigma^h(s)\sigma^h(s)^T\mathrm{diag}(V^h(s))|ds\right]^{p/2}\\
&\le K + K\int_0^t \mathbb{E}|V^h(s)|^p ds.
\end{aligned}$$

From above inequality and Gronwall's inequality, we have

$$\mathbb{E}|V^h(t)|^p \le Ke^K. \tag{15}$$

For the case of $t < h$, (15) still holds by noting $V^h(t) = V^h(0)$ when $0 \le t < h$. For any $0 \le t_1 < t_2 \le \hat{T}$, we have

$$\begin{aligned}
&\sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\sum_{n=[\frac{t_1}{h}]}^{m}\Delta V_n^h\right|^p \\
&= \sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\int_{z^h(t_1)}^{(m+1)h}\operatorname{diag}(V^h(s))\mu^h(s)ds + \int_{z^h(t_1)}^{(m+1)h}\operatorname{diag}(V^h(s))\sigma^h(s)dW(s)\right|^p \\
&\le K \sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\int_{z^h(t_1)}^{(m+1)h}\operatorname{diag}(V^h(s))\mu^h(s)ds\right|^p \\
&\quad + K \sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\int_{z^h(t_1)}^{(m+1)h}\operatorname{diag}(V^h(s))\sigma^h(s)dW(s)\right|^p \\
&\le K\int_{z^h(t_1)}^{z^h(t_2)}\left|\operatorname{diag}(V^h(s))\mu^h(s)\right|^p ds + K \sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\int_{z^h(t_1)}^{(m+1)h}\operatorname{diag}(V^h(s))\sigma^h(s)dW(s)\right|^p.
\end{aligned}$$

Since the integral of the second term is a martingale, applying (15), Assumption 2, Burkholder's inequality, and the Holder inequality gives us

$$\begin{aligned}
\mathbb{E}\sup_{[\frac{t_1}{h}]\le m\le[\frac{t_2}{h}]-1}\left|\sum_{n=[\frac{t_1}{h}]}^{m}\Delta V_n^h\right|^p &\le K\mathbb{E}\int_{z^h(t_1)}^{z^h(t_2)}\left|\operatorname{diag}(V^h(s))\right|^p ds \\
&\quad + K\mathbb{E}\left[\int_{z^h(t_1)}^{z^h(t_2)}\left|\operatorname{diag}(V^h(s))\sigma^h(s)\right|^2 ds\right]^{p/2} \\
&\le K\mathbb{E}\int_{z^h(t_1)}^{z^h(t_2)}\left|\operatorname{diag}(V^h(s))\right|^p ds \qquad (16)\\
&\le K(t_2 - t_1 + h).
\end{aligned}$$

Setting $t_1 = 0$ and $t_2 = \hat{T}$, we obtain

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le t\le\hat{T}}\left|V^h(t)\right|^p\right] &\le K|V_0|^p + K\mathbb{E}\left[\sup_{0\le m\le[\frac{\hat{T}}{h}]-1}\left|\sum_{n=0}^{m}\Delta V_n^h\right|^p\right] \\
&\le K.
\end{aligned}$$

A review of the proof shows that the inequality (16) still holds when the expectation is replaced by the conditional expectation. Hence we have (14). ■

Lemma 12. *If Assumption 2 holds, then $j(V^h) \Rightarrow 0$ as $h \to 0$, where $j(x) := \sup_{0<t\le T}|x(t) - x(t^-)|$ for $x \in \mathbb{D}^{k+1}$.*

Proof. Since boundedness of μ^h and σ^h implies that

$$|\Delta V_n^h| \le K|V_n^h|(h + \sqrt{h}|Z_{n+1}|),$$

we have

$$j(V^h) = \sup_n |\Delta V_n^h| \le \underbrace{Kh \sup_n |V_n^h|}_{\mathrm{I}^h} + \underbrace{K\sqrt{h} \sup_n |V_n^h| \sup_n |Z_{n+1}|}_{\mathrm{II}^h},$$

where in the above inequality the first and second terms are denoted by I^h and II^h, respectively. The convergence of $E[\mathrm{I}^h]$ to zero as $h \to 0$ can be obtained from

$$\mathbb{E}[\mathrm{I}^h] \le Kh\mathbb{E}[\sup_n |V_n^h|] \le Kh(\mathbb{E}[\sup_n |V_n^h|^2])^{1/2} \le Kh.$$

The last inequality in the above is due to (12). On the other hand,

$$\begin{aligned}
\mathbb{E}[\mathrm{II}^h] &\le K\sqrt{h}\left(\mathbb{E}\sup_n |V_n^h|^{\frac{4}{3}}\right)^{\frac{3}{4}} \left(\mathbb{E}\sup_n |Z_{n+1}|^4\right)^{\frac{1}{4}} \\
&= Kh^{\frac{1}{4}}\left(\mathbb{E}\sup_n |V_n^h|^{\frac{4}{3}}\right)^{\frac{3}{4}} \left(h\mathbb{E}\sup_n |Z_{n+1}|^4\right)^{\frac{1}{4}} \\
&\le Kh^{\frac{1}{4}}\left(\mathbb{E}\sup_n |V_n^h|^{\frac{4}{3}}\right)^{\frac{3}{4}} \left(\frac{\mathbb{E}\sum_{n=1}^N |Z_n|^4}{N}\right)^{\frac{1}{4}} \\
&= Kh^{\frac{1}{4}}\left(\mathbb{E}\sup_n |V_n^h|^{\frac{4}{3}}\right)^{\frac{3}{4}} \left(\mathbb{E}|Z_1|^4\right)^{\frac{1}{4}} \\
&\to 0 \quad \text{as } h \to 0.
\end{aligned}$$

The conclusion follows from that $j(V^h)$ converges to 0 in L^1 as $h \to 0$. ■

Remark 13. Lemma 12 is the key result that enables us to show that the weak limit process $\bar{V}$ of $\{V^h\}$ is continuous. The Skorohod representation theorem allows us to treat the limit $V^h \Rightarrow \bar{V}$ in an almost sure sense, that is,

$$\lim_h d(V^h, \bar{V}) = 0 \text{ a.s.} \tag{17}$$

However, it does not imply an almost sure limit with uniform topology, i.e.,

$$\lim_h \|V^h - \bar{V}\| = 0 \text{ a.s.}, \tag{18}$$

may not be true. In other words, (17) does not imply

$$\lim_h f(V^h(t)) = f(\bar{V}(t)) \ \forall t$$

even for a bounded continuous function f, which is useful in characterizing properties of $\bar{V}$. Indeed, one shall prove continuity of $\bar{V}$ (i.e., $\mathbb{P}\{\bar{V} \in \mathbb{C}^{k+1}\} = 1$) in advance to make use of (18) from Proposition 8. We note that the proof on the continuity of $\bar{V}$ is missing in [22] and the related references therein.

3.2. Weak convergence of approximating solutions. This part shows that the approximating processes V^h converge to V in distribution. A general approach to this goal is first to prove tightness of $\{\mathbb{P}(V^h)^{-1}\}$ for extracting a weak limit from any subsequence, then to apply the Skorohod representation theorem for passing the limit almost surely, and finally to characterize the limiting process as the solution of the underlying SDE, provided there exists a unique weak solution. The main result of this subsection is Theorem 17.

Proposition 14. *If Assumption 2 holds, then there exists a unique weak solution to SDE (3).*

Proof. For the sake of a simple presentation, we assume no jump for μ. Taking logarithm, it is sufficient to consider the following SDE:

$$dV\left(t\right)=\mu\left(V,t\right)\,dt+\sigma\left(V,t\right)\,dW\left(t\right),\quad t\geq0. \tag{19}$$

By definition, V can be constructed uniquely by the following steps:

1. Let $V_1(t)=V_0+\int_0^t\mu(s)ds+\sigma_c(s)dW(s)$ for $t>0$.
2. Let $\tau_1=\inf\{t>0:\mathcal{N}^\sigma(V_1,t)=1\}\wedge\hat{T}$.
3. Let $V_2(t)=V_1(t)$ for $t\leq\tau_1$; otherwise,
$$V_2(t)=V_1(\tau_1)+\int_{\tau_1}^t\mu(s)ds+(\sigma_c(s)+J_1^\sigma(V_1))dW(s)$$
for $t>\tau_1$.
4. Let $\tau_2=\inf\{t>\tau_1:\mathcal{N}^\sigma(V_2,t)=2\}\wedge\hat{T}$.
5. Repeat above steps to construct V_i, τ_i, $i=1,2,\ldots$, until $\tau_i=\hat{T}$.

According to [25, Theorem 1.6.3], V_1 has a unique strong solution from 0 to $\hat{T}$; hence τ_1 is well defined. Since τ_1 is no greater than $\hat{T}$, V_1 has a unique solution from 0 to τ_1. Yong and Zhou [25, Theorem 1.6.3] also imply that $\mathbb{E}|V_1(\tau_1)|^2\leq K$, so V_2 has a unique strong solution from τ_1 to T with initial value $V_1(\tau_1)$ at time τ_1. Therefore, V_2 has a unique strong solution from 0 to $\hat{T}$. Since the number of jumps of σ is finite due to Assumption 2, we can proceed with the above procedure inductively until $\tau_n=\hat{T}$, where n is some finite number. Then V_n is the unique strong solution of (19) from 0 to $\hat{T}$. Hence, the uniqueness of the weak solution is ensured according to [20, Theorem 9.1.7]. ■

Lemma 15. *If Assumption 2 holds, then $\{\mathbb{P}(V^h)^{-1}:h>0\}$ is tight.*

Proof. According to [4, Theorem 3.8.6] it is sufficient to verify condition (a) in [4, Theorem 3.7.2] and condition (b) in [4, Theorem 3.8.6] in order to show $\{\mathbb{P}(V^h)^{-1}:h>0\}$ is tight. To verify condition (a) in [4, Theorem 3.7.2] we need to verify that for every ϵ and rational $t\geq0$, there exists a compact set $\Gamma_{\epsilon,t}\subset\mathbb{R}^{k+1}$ such that

$$\inf_h\mathbb{P}\{V^h(t)\in\Gamma^\epsilon_{\epsilon,t}\}\geq1-\epsilon,$$

where $\Gamma^\epsilon_{\epsilon,t}:=\{x\in\mathbb{R}^{k+1}:\inf_{y\in\Gamma_{\epsilon,t}}|x-y|<\epsilon\}$. It is worth noting that $\Gamma^\epsilon_{\epsilon,t}\supset\Gamma_{\epsilon,t}$. Denote by K_0 the bound of $\mathbb{E}|V^h(t)|^4$ from (12). For every ϵ and $t\geq0$, by setting $\delta=(\frac{K_0}{\epsilon})^{\frac14}$ and

$\Gamma_{\epsilon,t} = \{x \in \mathbb{R}^{k+1} : |x| \le \delta\}$, we have

$$\begin{aligned}\inf_{h>0} \mathbb{P}(V^h(t) \in \Gamma^{\epsilon}_{\epsilon,t}) &= 1 - \sup_{h>0} \mathbb{P}(V^h(t) \notin \Gamma^{\epsilon}_{\epsilon,t}) \\ &\ge 1 - \sup_{h>0} \mathbb{P}(V^h(t) \notin \Gamma_{\epsilon,t}) \\ &= 1 - \sup_{h>0} \mathbb{P}(|V^h(t)| > \delta) \\ &\ge 1 - \sup_{h>0} \frac{\mathbb{E}|V^h(t)|^4}{\delta^4} = 1 - \epsilon.\end{aligned}$$

To verify condition (b) in [4, Theorem 3.8.6], we need to find some positive β and a family $\{\gamma^h(\delta) : 0 < \delta < 1, \text{ all } h\}$ of nonnegative random variables satisfying

$$\mathbb{E}[|V^h(t+u) - V^h(t)|^\beta \wedge 1|\mathcal{F}_t](|V^h(t) - V^h(t-\nu)|^\beta \wedge 1) \le \mathbb{E}[\gamma^h(\delta)|\mathcal{F}_t] \tag{20}$$

for $0 \le t \le T$, $0 \le u \le \delta$, and $0 \le v \le \delta \wedge t$; in addition,

$$\lim_{\delta \to 0} \sup_h \mathbb{E}[\gamma^h(s)] = 0 \tag{21}$$

and

$$\lim_{\delta \to 0} \sup_h \mathbb{E}[|V^h(\delta) - V^h(0)|^\beta \wedge 1] = 0. \tag{22}$$

We claim that $\beta = 4$ and a family of nonnegative random variables $\{\sup_{0 \le s \le T} |V^h(s)|^4[\delta + 2(h \wedge 2\delta)] : \delta > 0, h > 0\}$ satisfy (20), (21), and (22). For (20), since either $|V^h(t+u) - V^h(t)|^4$ or $|V^h(t) - V^h(t-\nu)|^4$ is zero when $h > 2\delta$ due to piecewise constant form of V^h, we then have

$$\begin{aligned}&\mathbb{E}[|V^h(t+u) - V^h(t)|^4 \wedge 1|\mathcal{F}_t](|V^h(t) - V^h(t-\nu)|^4 \wedge 1) \\ &= \mathbb{E}[|V^h(t+u) - V^h(t)|^4 \wedge 1|\mathcal{F}_t](|V^h(t) - V^h(t-\nu)|^4 \wedge 1)\mathbf{1}(h \le 2\delta) \\ &\le \mathbb{E}\left[\sup_{[\frac{t}{h}] \le m \le [\frac{t+\delta+h}{h}]-1} \left|\sum_{n=[\frac{t}{h}]}^{m} \Delta V^h_n\right|^4 \big|\mathcal{F}_t\right] \mathbf{1}(h \le 2\delta) \\ &\le \mathbb{E}\left[\int_{z^h(t)}^{z^h(t+\delta+h)} |V^h(s)|^4 ds \big|\mathcal{F}_t\right] \mathbf{1}(h \le 2\delta) \\ &\le \mathbb{E}\left[\sup_{0 \le s \le T} |V^h(s)|^4[\delta + 2(h \wedge 2\delta)] \big|\mathcal{F}_t\right].\end{aligned}$$

Equation (21) follows from (12) and the Cauchy–Schwartz inequality that

$$\lim_{\delta \to 0} \sup_h \mathbb{E}\left[\sup_{0 \le s \le T} |V^h(s)|^4[\delta + 2(h \wedge 2\delta)]\right] \le \lim_{\delta \to 0} K\delta.$$

Equation (22) is shown as follows:

$$\begin{aligned}\lim_{\delta\to 0}\sup_{h>0}\mathbb{E}[|V^h(\delta)-V^h(0)|^4] &= \lim_{\delta\to 0}\sup_{0<h\leq\delta}\mathbb{E}[|V^h(\delta)-V^h(0)|^4] \\ &\leq \lim_{\delta\to 0}\sup_{0<h\leq\delta}\mathbb{E}\left[\sup_{0\leq m\leq[\frac{\delta+h}{h}]-1}\left|\sum_{n=0}^{m}\Delta V_n^h\right|^4\right] \\ &\leq \lim_{\delta\to 0}\sup_{0<h\leq\delta}K(\delta+h+h)=0. \quad \text{(by (13))}.\end{aligned}$$

So, $\{\mathbb{P}(V^h)^{-1}: h>0\}$ is tight. ■

The next result is needed in the proof of Theorem 17.

Lemma 16 (Rosenthal's inequality [21]). *Let $p\geq 2$ and $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent random variables such that, for any $n\in\mathbb{N}$ and any $i\in\{1,2,\ldots,n\}$, $\mathbb{E}(X_i)=0$ and $\mathbb{E}(|X_i|^p)<\infty$. Then we have*

$$\mathbb{E}\left(\left|\sum_{i=1}^{n}X_i\right|^p\right)\leq c_p\max\left(\sum_{i=1}^{n}\mathbb{E}|X_i|^p,\left(\sum_{i=1}^{n}\mathbb{E}(X_i^2)\right)^{p/2}\right),$$

where c_p is some constant depending on p.

We now state the main theorem of this subsection on weak convergence of the Euler scheme.

Theorem 17. *Let V^h be the Euler approximating process for V. If Assumptions 2, 3, and 4 hold, then V^h converges weakly to V as $h\to 0$.*

Proof. Since $\{\mathbb{P}(V^h)^{-1}: h>0\}$ is tight, for an arbitrary infinite sequence, there exists a subsequence that has a weak limit. We denote this subsequence again by $\{V^h\}$ and its limit by $\overline{V}$. Due to the uniqueness of the weak solution (see Proposition 14), it suffices to show that $\overline{V}$ is the weak solution of (3). Tightness of $\{\mathbb{P}(V^h)^{-1}: h>0\}$ implies $\overline{V}$ is in $\mathbb{D}^{k+1}$. Moreover, since $j(V^h)\Rightarrow 0$ from Lemma 12, $\overline{V}$ is a continuous process due to [1, Theorem 13.4]. Using Skorohod representation (see [14, Theorem 9.17]) we can find $\widetilde{V}^h$ and $\hat{V}$ in the same probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},\{\widetilde{\mathcal{F}}_t\},\widetilde{\mathbb{P}})$ such that their distributions are the same as those of V^h and $\overline{V}$, respectively, and $\widetilde{V}^h$ converges to $\hat{V}$ almost surely. According to Remark 13 and Proposition 8, $\widetilde{V}^h$ converges to $\hat{V}$ a.s. under both Skorohod topology and uniform topology. Since σ^h is uniformly bounded from below by a positive constant due to Assumption 3, $\hat{V}_i, i=0,1,\ldots,k$, is regular with respect to L_i (see [23, Proposition A.1]). Hence $\phi(\widetilde{V}^h,\cdot)\to\phi(\hat{V},\cdot)$ a.s. for $\phi=\mu,\sigma$ in Skorohod topology as $h\to 0$ by Assumption 4. Denote by $D(\hat{V},\phi)$ the set $\{t\in[0,T]:\phi(\hat{V},\cdot)\text{ is discontinuous at }t\}$. Since the number of discontinuities of $\phi(\hat{V},t)$ is bounded almost surely, we can write $[0,T]\setminus D(\hat{V},\phi)$ as the finite union of disjoint intervals I_i, $i=1,\ldots,K$. Since $\phi(\hat{V},\cdot)$ is Holder-1/2 continuous at each I_i, $\phi(\hat{V},\cdot)$ is uniformly continuous at each I_i. Therefore as $h\to 0$,

$$\|\phi(\widetilde{V}^h,\cdot)-\phi(\hat{V},\cdot)\|_{I_i}\to 0 \quad \text{a.s., } 1\leq i\leq K.$$

For $t \in [0,T]$ denote by

$$M(V,t) := V(t) - V(0) - \int_0^t \mathrm{diag}(V(s))\, \mu(V,s)\, ds.$$

Noting that $D(\hat{V},\phi)$ is finitely many by Assumption 2 and $\hat{V}(0) = V^h(0) = V_0$, we have

$$\begin{aligned}
&\lim_{h\to 0} |M(\widetilde{V}^h, t) - M(\hat{V}, t)| \\
&= \lim_{h\to 0} |\widetilde{V}^h(t) - \hat{V}(t)| + \lim_{h\to 0} \left| \sum_{i=1}^K \int_{I_i \cap [0,t]} \mathrm{diag}(\widetilde{V}^h(s))\mu(\widetilde{V}^h, s) - \mathrm{diag}(\hat{V}(s))\mu(\hat{V}, s) ds \right| \\
&\le \lim_{h\to 0} \left| \int_{[0,t]} \mathrm{diag}(\widetilde{V}^h(s) - \hat{V}(s))\mu(\widetilde{V}^h, s) ds \right| \\
&\quad + \lim_{h\to 0} \left| \sum_{i=1}^K \int_{I_i \cap [0,t]} \mathrm{diag}(\hat{V}(s))[\mu(\widetilde{V}^h, s) - \mu(\hat{V}, s)] ds \right| \\
&\le K \lim_{h\to 0} \| \widetilde{V}^h - \hat{V} \|_{[0,t]} + \lim_{h\to 0} \sum_{i=1}^K \| \mu(\widetilde{V}^h, \cdot) - \mu(\hat{V}, \cdot) \|_{I_i \cap [0,t]} \int_{I_i \cap [0,t]} \mathrm{diag}(\hat{V}(s)) ds \\
&= 0.
\end{aligned}$$

Moreover, noting that $\widetilde{V}^h$ and V^h have the same law and $\{M(\widetilde{V}^h, t) : h > 0\}$ is uniformly integrable due to (12), we have

$$\mathbb{E}[M(\overline{V}, t)] = \mathbb{E}[M(\hat{V}, t)] = \lim_{h\to 0} \mathbb{E}[M(\widetilde{V}^h, t)] = \lim_{h\to 0} \mathbb{E}[M(V^h, t)],$$

which shows that

$$\begin{aligned}
\mathbb{E}[M(\overline{V}, t)] &= \lim_{h\to 0} \mathbb{E}\left[\sum_{n=1}^{[\frac{t}{h}]} \left(\left(V_n^h - V_{n-1}^h \right) - \int_{(n-1)h}^{nh} \mathrm{diag}\left(V_{n-1}^h \right) \mu(V^h, s)\, ds \right) \right] \\
&\quad - \lim_{h\to 0} \mathbb{E}\left[\int_{z^h(t)}^t \mathrm{diag}\left(V_{[\frac{t}{h}]}^h \right) \mu(V^h, s)\, ds \right].
\end{aligned}$$

For the last term above, since μ is bounded and $V_{[\frac{t}{h}]}^h = V^h(t)$, we have

$$\lim_{h\to 0} \mathbb{E}\left[\left| \int_{z^h(t)}^t \mathrm{diag}\left(V_{[\frac{t}{h}]}^h \right) \mu(V^h, s)\, ds \right| \right] \le K \lim_{h\to 0} h\mathbb{E}\left[\left| V^h(t) \right| \right] = 0.$$

Denote by $\mathcal{N}_n^{\mu,h} := \mathcal{N}^\mu(V^h, nh)$. Using the Euler recursive formula (6) and Assumption 2, we have

$$\begin{aligned}
&\left| \mathbb{E}\left[\sum_{n=1}^{[\frac{t}{h}]} \left(\left(V_n^h - V_{n-1}^h \right) - \int_{(n-1)h}^{nh} \mathrm{diag}\left(V_{n-1}^h \right) \mu(V^h, s)\, ds \right) \right] \right| \\
&= \left| \mathbb{E}\left[\sum_{n=1}^{[\frac{t}{h}]} \left(\int_{(n-1)h}^{nh} \mathrm{diag}\left(V_{n-1}^h \right) (\mu_{n-1}^h - \mu(V^h, s))\, ds \right) \right] \right|
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{n=1}^{[\frac{t}{h}]} h\mathbb{E}\left[\left|V_{n-1}^h\right|\left(Kh^{1/2}+\sum_{k=\mathcal{N}_{n-1}^{\mu,h}}^{\mathcal{N}_n^{\mu,h}}\left|J_k^\mu(V^h)\right|\right)\right]\\
&\leq Kh^{3/2}\sum_{n=1}^{[\frac{t}{h}]}\mathbb{E}\left|V_{n-1}^h\right| + Kh\sum_{n=1}^{[\frac{t}{h}]}\mathbb{E}\left[\left|V_{n-1}^h\right|\ [\mathcal{N}_n^{\mu,h}-\mathcal{N}_{n-1}^{\mu,h}]\right].\\
&\leq Kh^{3/2}\sum_{n=1}^{[\frac{t}{h}]}\mathbb{E}\left[\sup_{0\leq t\leq T}\left|V^h(t)\right|\right] + Kh\mathbb{E}\left[\sup_{0\leq t\leq T}\left|V^h(t)\right|\sum_{n=1}^{[\frac{t}{h}]}[\mathcal{N}_n^{\mu,h}-\mathcal{N}_{n-1}^{\mu,h}]\right]\\
&\leq Kh^{1/2}\mathbb{E}\left[\sup_{0\leq t\leq T}\left|V^h(t)\right|\right] + Kh\mathbb{E}\left[\sup_{0\leq t\leq T}\left|V^h(t)\right|\right],
\end{aligned}
$$

which tends to 0 as $h\to 0$ from (12). Here we have used the fact that $\mathcal{N}^\mu(V^h,\cdot)$ is uniformly bounded given Assumption 2 in the last inequality. Therefore, $\mathbb{E}\left[M\left(\overline{V},t\right)\right]=0$. Using the same procedure one can show $\mathbb{E}\left[M\left(\overline{V},t\right)|\mathcal{F}_s\right]=M\left(\overline{V},s\right)$ for any $0\leq s\leq t$; thus $M(\overline{V},\cdot)$ is a martingale.

Denote the lth component of the vector process X by X_l, and let $\langle X_l, X_q\rangle(t)$ be the cross-variation of two real processes X_l and X_q up to time t. For $t\in[0,T]$ denote

$$
Q_{lq}(V,t) := \langle V_l, V_q\rangle(t) - \int_0^t \Big(\mathrm{diag}\left(V\left(s\right)\right)\sigma\left(V,s\right)\sigma\left(V,s\right)^T\mathrm{diag}\left(V\left(s\right)\right)\Big)_{lq} ds.
$$

Again, due to the uniform topology of the convergence of $\widetilde{V}^h$ to $\hat{V}$, boundedness of σ^h and (5), and the uniformly integrability of $\{Q_{lq}(\widetilde{V}^h,t): h>0\}$ from (12), we have

$$
\mathbb{E}\big|Q_{lq}(\overline{V},t)\big| = \mathbb{E}\big|Q_{lq}(\hat{V},t)\big| = \lim_{h\to 0}\mathbb{E}\big|Q_{lq}(\widetilde{V}^h,t)\big| = \lim_{h\to 0}\mathbb{E}\big|Q_{lq}(V^h,t)\big|.
$$

Therefore

$$
\begin{aligned}
\mathbb{E}\big|Q_{lq}(\overline{V},t)\big| &= \lim_{h\to 0}\mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]}\left(V_{l,n}^h-V_{l,n-1}^h\right)\left(V_{q,n}^h-V_{q,n-1}^h\right)\right.\\
&\qquad\left.-\int_0^t\left(\mathrm{diag}\left(V^h\left(s\right)\right)\sigma(V^h,s)\sigma(V^h,s)^T\mathrm{diag}\left(V^h\left(s\right)\right)\right)_{lq}ds\right|\\
&\leq \lim_{h\to 0}\mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]}\left[(V_{l,n}^h-V_{l,n-1}^h)(V_{q,n}^h-V_{q,n-1}^h)-V_{l,n-1}^hV_{q,n-1}^h(\sigma_{n-1}^h(\sigma_{n-1}^h)^T)_{lq}h\right]\right|\\
&\quad+\lim_{h\to 0}\mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]}V_{l,n-1}^hV_{q,n-1}^h\int_{(n-1)h}^{nh}(\sigma(V^h,s)\sigma(V^h,s)^T-\sigma_{n-1}^h(\sigma_{n-1}^h)^T)_{lq}ds\right|\\
&\quad+\lim_{h\to 0}\mathbb{E}\left|\int_{z^h(t)}^t\left(\mathrm{diag}\left(V_{[\frac{t}{h}]}^h\right)\sigma(V^h,s)\sigma(V^h,s)^T\mathrm{diag}\left(V_{[\frac{t}{h}]}^h\right)\right)_{lq}ds\right|.
\end{aligned}
$$

The second and third limit terms in (5) are bounded by

$$\lim_{h\to 0}\left(\sum_{n=1}^{[\frac{t}{h}]} Kh^{\frac{3}{2}}\mathbb{E}\left[V_{l,n}^h V_{q,n}^h\right] + Kh\sum_{n=1}^{[\frac{t}{h}]}\mathbb{E}\left[V_{l,n-1}^h V_{q,n-1}^h\left[\mathcal{N}_n^{\sigma,h}-\mathcal{N}_{n-1}^{\sigma,h}\right]\right]\right) + \lim_{h\to 0}\mathbb{E}\left|V_{l,[\frac{t}{h}]}^h V_{q,[\frac{t}{h}]}^h\right|,$$

which is zero from a similar argument for showing (5) goes to zero as $h\to 0$ and (12). For the first limit term in (5), using Assumption 2 and the Euler recursive formula (6), we have

$$\begin{aligned}
&\mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]}\left[(V_{l,n}^h - V_{l,n-1}^h)(V_{q,n}^h - V_{q,n-1}^h) - V_{l,n-1}^h V_{q,n-1}^h(\sigma_{n-1}^h(\sigma_{n-1}^h)^T)_{lq}h\right]\right| \\
&\le \mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]} V_{l,n-1}^h V_{q,n-1}^h \mu_{l,n-1}^h \mu_{q,n-1}^h h^2\right| \\
&\quad + \mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]}\left\{h^{\frac{3}{2}} V_{l,n-1}^h V_{q,n-1}^h\left[\mu_{l,n-1}^h\sum_{i=1}^{k+1}(\sigma_{n-1}^h)_{qi}(Z_n)_i + \mu_{q,n-1}^h\sum_{i=1}^{k+1}(\sigma_{n-1}^h)_{li}(Z_n)_i\right]\right\}\right| \\
&\quad + \mathbb{E}\left|\sum_{n=1}^{[\frac{t}{h}]} hV_{l,n-1}^h V_{q,n-1}^h\left\{\sum_{i=1}^{k+1}(\sigma_{n-1}^h)_{li}(\sigma_{n-1}^h)_{qi}[(Z_n)_i^2 - 1]\right.\right. \\
&\quad \left.\left. + \sum_{1\le i\ne j\le k+1}(\sigma_{n-1}^h)_{li}(\sigma_{n-1}^h)_{qj}[(Z_n)_i(Z_n)_j]\right\}\right| \\
&\le Kh\mathbb{E}\left[\sup_{0\le s\le T}|V^h(s)|^2\right] + \mathbb{E}\left[K\sup_{0\le s\le T}|V^h(s)|^2 h^{\frac{3}{2}}\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i\right|\right] \\
&\quad + \mathbb{E}\left[K\sup_{0\le s\le T}|V^h(s)|^2 h\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}[(Z_n)_i^2 - 1]\right|\right] \\
&\quad + \mathbb{E}\left[K\sup_{0\le s\le T}|V^h(s)|^2 h\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{1\le i\ne j\le k+1}[(Z_n)_i(Z_n)_j]\right|\right].
\end{aligned}$$

The first term of (5) clearly converges to zero. We only show that the second term converges to zero. The convergence of the other two terms to zero can be shown similarly. Thanks to [3, Theorem 2.5.7] $\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i$ is at the order of $(\frac{1}{h})^{1/2}(\log\frac{1}{h})^{1/2+\epsilon}, \epsilon > 0$. Hence

$$\lim_{h\to 0} h^{\frac{3}{2}}\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i\right| = 0 \quad \text{a.s.}$$

Therefore once we can justify the exchange of limit and expectation, then the convergence of the expectation to zero is shown. To this end, using Lemma 16 and Young's inequality, we show uniform integrability of $\{\sup_{0\le s\le T}|V^h(s)|^2h^{\frac{3}{2}}\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i\right|, h>0\}$ as

$$\begin{aligned}\mathbb{E}\left[K\sup_{0\le s\le T}|V^h(s)|^2h^{\frac{3}{2}}\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i\right|\right]^2 &\le \mathbb{E}\left[K\sup_{0\le s\le T}|V^h(s)|^8\right]+K\mathbb{E}\left[h^6\left|\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}(Z_n)_i\right|^4\right]\\ &\le K+h^6\max\left(\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}\mathbb{E}|(Z_n)_i|^4,\left(\sum_{n=1}^{[\frac{t}{h}]}\sum_{i=1}^{k+1}\mathbb{E}(Z_n)_i^2\right)^2\right)\\ &\le K+K\max(h^5,h^4),\end{aligned}$$

which is finite for all small h, where we have used Lemma 16 in the second to last inequality. Hence we show (5) converges to zero as $h\to 0$. Therefore

$$\left\langle \overline{V}_l,\overline{V}_q\right\rangle(t)=\int_0^t\left(\operatorname{diag}\left(\overline{V}(s)\right)\sigma\left(\overline{V},s\right)\sigma\left(\overline{V},s\right)^T\operatorname{diag}\left(\overline{V}(s)\right)\right)_{lq}ds\quad\text{a.s.}$$

Thanks to [9, Theorem 7.1], there exists a $k+1$ dimensional Brownian motion B such that $M\left(\overline{V},t\right)=\int_0^t\operatorname{diag}\left(\overline{V}(s)\right)\sigma\left(\overline{V},s\right)dB(s)$. Therefore, $\overline{V}$ is the weak solution of (3). ■

3.3. Properties of first-passage times.

Lemma 18. *For any $n\in\mathbb{N}$ and $j\in\{1,2,\ldots,n\}$, the mapping $x\mapsto\mathcal{S}^n(x,j)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is continuous.*

Proof. The case of $n=1$ is trivial. For $n\ge 2$ and $x=(x_1,\ldots,x_n)^T\in\mathbb{R}^n$, functions

$$\mathcal{S}^n(x,1)=\min(x_1,\ldots,x_n)\ \text{ and }\ \mathcal{S}^n(x,n)=\max(x_1,\ldots,x_n)$$

are obviously continuous. For $j\in\{2,\ldots,n-1\}$ (a nonempty set only if $n\ge 3$), we may decompose x into $\hat{x}=(x_1,\ldots,x_{n-1})^T$ and x_n and express $\mathcal{S}^n(x,j)$ as

$$\mathcal{S}^n(x,j)=\min\left\{\max\left(\mathcal{S}^{n-1}(\hat{x},j-1),x_n\right),\mathcal{S}^{n-1}(\hat{x},j)\right\}.$$

Given that $\mathcal{S}^{n-1}(\cdot,j)$ is continuous, and $\mathcal{S}^n(\cdot,1)$ and $\mathcal{S}^n(\cdot,n)$ are continuous, we have that $\mathcal{S}^n(\cdot,j)$ is continuous for any $j\in\{1,2,\ldots,n\}$. ■

Lemma 19. *Given that Assumptions 2 and 3 hold, we have*

$$\mathbb{P}(\tau_0=t_j)=0\ \text{and}\ \mathbb{P}\left(\tau_{(i)}=t_j\right)=0$$

for any $i=1,2,\ldots,k$ and $j=1,2,\ldots,m$.

Proof. Note that firm values V_i, $i=0,1,\ldots,k$, satisfy (3). For any $j=1,2,\ldots,m$, we show $\mathbb{P}(\tau_0=t_j)=0$ as follows:

$$\begin{aligned}\mathbb{P}(\tau_0=t_j)&\le\mathbb{P}(V_0(t_j)=L_0(t_j))\\ &=\mathbb{P}\left(V_0(0)\exp\left(\int_0^{t_j}\left(\mu_0(s)-\frac{b_0^2(s)}{2}\right)ds+\int_0^{t_j}b_0(s)\,dB_0(s)\right)=K_0\exp(\gamma_0t_j)\right)\\ &=\mathbb{P}\left(\int_0^{t_j}\left(\mu_0(s)-\frac{b_0^2(s)}{2}\right)ds+\int_0^{t_j}b_0(s)\,dB_0(s)=\ln\frac{K_0}{V_0(0)}+\gamma_0t_j\right),\end{aligned}$$

where $b_0^2(s) = (1,0,\dots,0)(\sigma(s)\sigma(s)^T)(1,0,\dots,0)^T$. Due to the uniform nondegeneracy Assumption 3 and boundedness of σ, there exist some positive constants λ and Λ such that

$$0 < \lambda < b_0(t) < \Lambda \quad \forall 0 \le t \le T \text{ a.s.}$$

Thanks to the property discussed in [15, Appendix A.5], we know the term in (7) is zero. Similarly, $\mathbb{P}(\tau_i = t_j) = 0$ for $i = 1,2,\dots,k$. Finally, $\mathbb{P}\left(\tau_{(i)} = t_j\right) \le \sum_{i=1}^k \mathbb{P}(\tau_i = t_j) = 0$. ■

Remark 20. In the proof of Lemma 19, we need to show a one-dimensional continuous semimartingale never hits a given point within finite time almost surely; more specifically, see (7). One may wonder if uniform nondegeneracy of σ can be relaxed to positive definite to show this result. The answer is negative. The counterexample for the one-dimensional case is provided in [17].

Lemma 21. *Suppose Assumption 3 is true. Given that Assumption 5 or 6 holds, we have*

$$\mathbb{P}\left(\tau_0 = \tau_{(i)} \wedge T\right) = 0$$

for any $i = 1,2,\dots,k$.

Proof. Since $\left\{\tau_0 = \tau_{(i)} \wedge T\right\} \subseteq \cup_{i=1}^k \{\tau_0 = \tau_i \wedge T\}$, we have

$$\mathbb{P}\left(\tau_0 = \tau_{(i)} \wedge T\right) \le \sum_{i=1}^k \mathbb{P}(\tau_0 = \tau_i \wedge T).$$

Observe that for any $i = 1,2,\dots,k$,

$$\begin{aligned}
&\mathbb{P}(\tau_0 = \tau_i \wedge T)\\
&\quad \le \mathbb{P}(\tau_0 = \tau_i \text{ and } \tau_i \le T) + \mathbb{P}(\tau_0 = T \text{ and } \tau_i > T)\\
&\quad \le \mathbb{P}(\text{2-d random process } \{(V_0(t), V_i(t))\}_t \text{ hits the curve } (L_0(t), L_i(t)), 0 \le t \le T)\\
&\quad = \mathbb{P}(\text{process } \{(V_0(t) - L_0(t), V_i(t) - L_i(t))\}_t \text{ hits } (0,0) \text{ at some time } t \text{ in } [0,T]).
\end{aligned}$$

In the above, we use the fact $\mathbb{P}(\tau_0 = T \text{ and } \tau_i > T) = 0$. Similar to the derivation of (7), the event $\{V_i(t) - L_i(t) = 0\}$ is equivalent to

$$\left\{\int_0^t b_i(s)\, dB_i(s) - \ln\frac{K_i}{V_i(0)} - \gamma_i t + \int_0^t \left(\mu_i(s) - \frac{b_i^2(s)}{2}\right) ds = 0\right\}.$$

Define the processes $\widetilde{V}_0$ and $\widetilde{V}_i$ to be

$$\widetilde{V}_j(t) := \int_0^t b_j(s)\, dB_j(s) - \ln\frac{K_j}{V_j(0)} - \gamma_j t + \int_0^t \left(\mu_j(s) - \frac{b_j^2(s)}{2}\right) ds, \quad j = 0, i.$$

The differential form of the SDE for $\widetilde{V} = \left(\widetilde{V}_0, \widetilde{V}_i\right)^T$ is

$$d\widetilde{V}(t) = \hat{\mu}_i(t)\, dt + \hat{b}_i(t)\, \hat{\rho}_i(t)\, dB'(t),$$

where

$$\hat{\mu}_i(t) = \begin{pmatrix} \mu_0(t) - \frac{b_0^2(t)}{2} - \gamma_0 \\ \mu_i(t) - \frac{b_i^2(t)}{2} - \gamma_i \end{pmatrix}, \hat{b}_i(t) = \begin{pmatrix} b_0(t) & 0 \\ 0 & b_i(t) \end{pmatrix}, \hat{\rho}_i(t) = \begin{pmatrix} 1 & 0 \\ \rho_i(t) & \sqrt{1-\rho_i^2(t)} \end{pmatrix},$$

and $B' = (B_0', B_i')^T$ is a standard two-dimensional Brownian motion, ρ_i is an adapted process satisfying

$$\hat{b}_i(t)\,\hat{\rho}_i(t)\,\hat{\rho}_i(t)^T\,\hat{b}_i(t) = \begin{pmatrix} \sigma_i(t) \\ \sigma_j(t) \end{pmatrix} \begin{pmatrix} \sigma_i(t) \\ \sigma_j(t) \end{pmatrix}^T,$$

and $\sigma_i(t)$ and $\sigma_j(t)$ are the ith and jth row vectors of $\sigma(t)$. Rewriting (8) in terms of the process $\widetilde{V}$, our target is to show that if Assumption 5 or 6 is true, then

$$\mathbb{P}\left(\widetilde{V}(t) = (0,0) \quad \text{for some } 0 < t \leq T\right) = 0. \tag{23}$$

Using the Girsanov theorem, we may assume the drift term is zero:

$$d\widetilde{V}(t) = \hat{b}_i(t)\,\hat{\rho}_i(t)\;dB'(t) := \hat{\sigma}(t)\;dB'(t).$$

We first prove (23) is true under Assumption 5. We have that $\hat{\rho}_i(t)$ is an identity matrix for all $t \geq 0$ almost surely. Since we are investigating the first-passage time before T, without loss of generality, we assume

$$b_0(t) \equiv 1 \text{ and } b_i(t) \equiv 1 \quad \forall t > T \text{ a.s.}$$

Then we have

$$\left\langle \widetilde{V}_0, \widetilde{V}_0 \right\rangle(\infty) = \infty, \left\langle \widetilde{V}_i, \widetilde{V}_i \right\rangle(\infty) = \infty \text{ and } \left\langle \widetilde{V}_0, \widetilde{V}_i \right\rangle(t) = \ 0\ \forall t > 0 \text{ a.s.}$$

Define $\hat{T}_0(t) := \inf\{s : \langle \widetilde{V}_0, \widetilde{V}_0 \rangle(s) > t\}$ and $\hat{T}_i(t) := \inf\{s : \langle \widetilde{V}_i, \widetilde{V}_i \rangle(s) > t\}$. By the time change results for multidimensional continuous local martingales in [20, Theorem 5.1.9], we have

$$\hat{B}(s) := \left(\widetilde{V}_0\left(\hat{T}_0(s)\right), \widetilde{V}_i\left(\hat{T}_i(s)\right)\right)^T, \quad s \geq 0,$$

is a standard two-dimensional Brownian motion under $\mathbb{P}$ with initial position $(\widetilde{V}_0(0), \widetilde{V}_i(0))$. Since $\langle \widetilde{V_0}, \widetilde{V_0} \rangle(\cdot)$ and $\langle \widetilde{V}_i, \widetilde{V}_i \rangle(\cdot)$ are both continuous and strictly increasing, the inverse maps $\hat{T}_0^{-1}$ and $\hat{T}_i^{-1}$ exist and are both strictly increasing. It suffices to show

$$\mathbb{P}\left(\hat{B}(t) \text{ hits origin at some } t \text{ in } \left(0, \max\left(\hat{T}_0^{-1}(T), \hat{T}_i^{-1}(T)\right)\right]\right) = 0.$$

This is true due to $\max(\hat{T}_0^{-1}(T), \hat{T}_i^{-1}(T)) < \infty$ a.s., and nonattainability of the origin by the two-dimensional Brownian path shown in [11, Proposition 3.3.22].

We next prove (23) is true under Assumption 6. We have that $\hat{\sigma}$ is a piecewise constant process almost surely. In other words, for strictly increasing stopping time sequence $\{\theta_0, \theta_1, \dots\}$ such that $\theta_0 = 0$ and $\lim_{n\to\infty} \theta_n = T + 1$, the process $\hat{\sigma}$ is in the form of

$$\hat{\sigma}(t) = \sum_{i=1}^{\infty} \hat{\sigma}_i \mathbf{1}(\theta_{i-1} \leq t < \theta_i) \quad \text{a.s.},$$

where $\hat{\sigma}_i$ is a nonsingular 2×2 dimensional matrix, measurable with respect to $\mathcal{F}_{\theta_{i-1}}$. Let p be any given point in $\mathbb{R}^2$. Partitioning the event in (23), we have

$$
\begin{aligned}
&\mathbb{P}\left(\widetilde{V}(t) = p \quad \text{for some } 0 < t \leq T\right) \\
&\leq \sum_{i=1}^{\infty} \mathbb{P}\left(\widetilde{V}(t) = p \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1\right) \\
&= \sum_{i=1}^{\infty} \int_{\mathbb{R}^2} \mathbb{P}\left(\widetilde{V}(t) = p \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1 \middle| \widetilde{V}(\theta_{i-1}) = x\right) \mathbb{P}\left(\widetilde{V}(\theta_{i-1}) \in dx\right).
\end{aligned}
$$

But the integrand in the above integral is seen to be zero from the following derivation:

$$
\begin{aligned}
&\mathbb{P}\left(\widetilde{V}(t) = p \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1 \middle| \widetilde{V}(\theta_{i-1}) = x\right) \\
&= \mathbb{P}\left(\int_{\theta_{i-1}}^{t} \hat{\sigma}_i \, dW(s) = p - \widetilde{V}(\theta_{i-1}) \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1 \middle| \widetilde{V}(\theta_{i-1}) = x\right) \\
&\leq \sup_{C} \mathbb{P}\left(\int_{\theta_{i-1}}^{t} C \, dW(s) = p - x \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1 \middle| \widetilde{V}(\theta_{i-1}) = x\right) \\
&= \sup_{C} \mathbb{P}\left(W(t) - W(\theta_{i-1}) = C^{-1}(p-x) \quad \text{for some } \theta_{i-1} \leq t < \theta_i < T+1 \middle| \widetilde{V}(\theta_{i-1}) = x\right) \\
&\leq \sup_{C} \mathbb{P}\left(\hat{W}(t) = C^{-1}(p-x) \quad \text{for some } 0 < t \leq T+1\right) \\
&= 0.
\end{aligned}
$$

Here the supremum is taken among any constant 2×2 nonsingular matrix, since $\hat{\sigma}_i$ is measurable with respect to $\mathcal{F}_{\theta_{i-1}}$ and constant from θ_{i-1} to θ_i. Also, $\hat{W}$ denotes a two-dimensional Brownian motion starting from zero. The last inequality uses the Markovian property of Brownian motion and $\theta_i - \theta_{i-1}$ is no greater than $T+1$ almost surely. Hence for each i the integral is zero, and so is the countable summation. ∎

Assumption 5 or 6 is used to prove (23). Although it is intuitive to think the two-dimensional continuous nondegenerate local martingale (without Assumption 5 or 6) should not hit a given point within finite time, it seems difficult to prove this rigorously. Since this unsolved question is interesting on its own, we describe it in detail below.

Remark 22 (an open question). Let W be a two-dimensional standard Brownian motion with respect to a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0})$. Define a continuous local martingale Y by

$$
Y(t) = Y(0) + \int_0^t \sigma(s) dW(s), \quad t \geq 0,
$$

with $Y(0) \neq 0$. Assume that σ is an adapted 2×2 matrix process satisfying

$$
\lambda \|\xi\|^2 \leq \xi^T \sigma(t) \sigma(t)^T \xi \leq \Lambda \|\xi\|^2, \text{ for all } (\xi, t) \in \mathbb{R}^2 \times (0, T) \text{ with some } \lambda, \Lambda > 0 \text{ a.s.}
$$

The question we have is whether the equality

$$\mathbb{P}\{Y(t) = (0,0) \ \text{ for some } t \in [0,T]\} = 0$$

is true. We leave this to our future work.

3.4. Completion of the proof. Now we are ready to complete the proof of Theorem 7. Let C_{F_1} and C_{F_2} be the sets of the continuities of functions F_1 and F_2, respectively. Due to [22, Lemma 4], with fixed boundary L_i, the mapping $\pi(\cdot, L_i)$ is continuous at each $x \in C_1^i \cup C_2^i$. Notice that F_1 and F_2 are composition of indicator functions, $\mathcal{S}^k(\cdot, j)$ and $\pi(\cdot, L_i)$, and $\mathcal{S}^k(\cdot, j)$ is continuous by Lemma 18. Hence we have

$$\begin{aligned}\widetilde{C}_{F_1} := \{&x \in \mathbb{C}^{k+1}[0,T] : x_i \in C_1^i \cup C_2^i, i = 0, \ldots, k, \text{ and } \mathcal{S}^k(\{\pi(x_n, Ln\}_{n=1}^k, i) \neq T, \\ &\text{and } \pi(x_0, L_0) \neq \mathcal{S}^k(\{\pi(x_n, L_n)\}_{n=1}^k, i) \wedge T\} \subset C_{F_1}\end{aligned}$$

and

$$\begin{aligned}\widetilde{C}_{F_2} := \{&x \in \mathbb{C}^{k+1}[0,T] : x_i \in C_1^i \cup C_2^i, i = 0, \ldots, k, \text{ and} \\ &\mathcal{S}^k(\{\pi(x_n, Ln\}_{n=1}^k, i) \neq t_j, \text{ and } \pi(x_0, L_0) \neq t_j\} \subset C_{F_2}.\end{aligned}$$

To apply the mapping theorem (see [1, Theorem 2.7]) we need to show $\mathbb{P}(\omega : V(\omega) \in C_{F_1}) = 1$ and $\mathbb{P}(\omega : V(\omega) \in C_{F_2}) = 1$. It can be seen by

$$\begin{aligned}&\mathbb{P}(\omega : V(\omega) \in C_{F_1}) \\ &\geq \mathbb{P}(\omega : V(\omega) \in \widetilde{C}_{F_1}) \\ &= \mathbb{P}(V_i \in C_1^i \cup C_2^i, i = 0, \ldots, k, \text{ and } \mathcal{S}^k(\{\pi(V_n, Ln\}_{n=1}^k, i) \neq T, \\ &\quad \text{and } \pi(V_0, L_0) \neq \mathcal{S}^k(\{\pi(V_n, L_n)\}_{n=1}^k, i) \wedge T) \\ &= \mathbb{P}(V_i \in C_1^i \cup C_2^i, i = 0, \ldots, k, \text{ and } \tau_{(i)} \neq T, \text{ and } \tau_0 \neq \tau_{(i)} \wedge T) \\ &= \mathbb{P}(\tau_{(i)} \neq T, \text{ and } \tau_0 \neq \tau_{(i)} \wedge T) \text{ (Because } \mathbb{P}(V_i \in C_1^i \cup C_2^i, i = 0, \ldots, k) = 1 \\ &\qquad\qquad \text{due to Assumption 3 and [23, Proposition A.1]}) \\ &= 1 \qquad \text{(by Lemmas 19 and lemma 21)}.\end{aligned}$$

Similarly, one can show $\mathbb{P}(\omega : V(\omega) \in C_{F_2}) = 1$. With Assumptions 2 and 4, Theorem 17 says that $V^h \Rightarrow V$ as $h \to 0$. Using the mapping theorem, we conclude that

$$F_i\left(V^h\right) \Rightarrow F_i(V), \quad i = 1, 2, \quad \text{as } h \to 0.$$

Note that F_i, $i = 1, 2$, are bounded functions; hence $\{F_i(V^h) : h > 0\}$ are families of uniformly integrable random variables and

$$\lim_{h \to 0} \mathbb{E}\left[F_i\left(V^h\right)\right] = \mathbb{E}\left[F_i(V)\right], \quad i = 1, 2.$$

Therefore, $\lim_{h\to 0} \hat{c}_i^h = \hat{c}_i$

Remark 23 (extension to stochastic interest rate). Suppose the risk-free interest rate r is an $\mathcal{F}_t$-adapted continuous process. Let X be a $k+2$ dimensional $\mathcal{F}_t$-adapted continuous process which represents V_i, $i = 0, 1, \dots, k$, and r. Define $\mathcal{I}_1 : \mathbb{D}^{k+2} \times \mathbb{N} \mapsto \mathbb{R}$ and $\mathcal{I}_2 : \mathbb{D}^{k+2} \times [0, T] \mapsto \mathbb{R}$ as

$$\mathcal{I}_1(x, i) := \exp\left\{-\int_0^{S^k(\{\pi(x_n, L_n)\}_{n=1}^k, i)} x_{k+1}(u)du\right\}$$

and

$$\mathcal{I}_2(x, t) := \exp\left\{-\int_0^t x_{k+1}(u)du\right\}.$$

For $x_n \in \mathbb{D}$ and $x \in \mathbb{C}$ we have that x_n converging to x in Skorohod metric implies x_n converging to x in uniform metric which implies $\int_0^t x_n(u)du$ converging to $\int_0^t x(u)du$ for $0 \leq t \leq T$. This shows the mapping from $\mathbb{D}$ to $\mathbb{R}$, defined by $x \to \int_0^t x(u)du$, is continuous at each $x \in \mathbb{C}$. Then, the two mappings $\mathcal{I}_1(\cdot, i)$ and $\mathcal{I}_2(\cdot, t)$ are continuous at $\{x \in \mathbb{C}^{k+1} : x_i \in C_1^i \cup C_2^i, i = 0, 1, \dots, k\}$ under Skorohod topology. Since the discounting factors $\exp\{-\int_0^{\mathcal{T}(i)} r_u du\}$ and $\exp\{-\int_0^{t_j} r_u du\}$ can be rewritten as $\mathcal{I}_1(X, i)$ and $\mathcal{I}_2(X, t_j)$, respectively, Theorem 7 still holds for the process X under the same setting.

4. Conclusions.

We have derived the sufficient conditions for the convergence of the approximation of basket CDS with counterparty risk under a credit contagion model of multi-names by generalizing the known weak limit theorems with discontinuous coefficents under a non-Markovian setting. The method developed in this paper may be used to study other problems involving running maximal processes of correlated Brownian motions, the joint distribution of which is still unknown for dimension greater than two.

Acknowledgments. The authors thank Thomas G. Kurtz for useful discussions on weak convergence of stochastic processes. The authors are very grateful to the anonymous reviewers and the associate editor, whose instructive comments and suggestions helped greatly to improve the previous version of the paper.

REFERENCES

[1] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 2009.
[2] K. S. Chan and O. Stramer, *Weak consistency of the Euler method for numerically solving stochastic differential equations with discontinuous coefficients*, Stochastic Process. Appl., 76 (1998), pp. 33–34.
[3] R. Durrett, *Probability: Theory and Examples*, Cambridge University Press, Cambridge, UK, 2010.
[4] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 2005.
[5] J. W. Gu, W. K. Ching, T. K. Siu, and H. Zheng, *On pricing basket credit default swaps*, Quant. Finance, 13 (2013), pp. 1845–1854.
[6] H. Haworth and C. Reisinger, *Modeling basket credit defaults swaps with default contagion*, J. Credit Risk., 3 (2007), pp. 31–67.
[7] D. J. Higham, X. Mao, and A. M. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM J. Numer. Anal., 40 (2003), pp. 1041–1063.
[8] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients*, Ann. Appl. Probab., 22 (2012), pp. 1611–1641.

[9] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland/Kodansha, Amsterdam, 1989.
[10] S. Iyengar, *Hitting lines with two-dimensional Brownian motion*, SIAM J. Appl. Math., 45 (1985), pp. 983–989.
[11] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, 1991.
[12] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, New York, 2011.
[13] T. G. Kurtz and P. Protter, *Weak limit theorems for stochastic integrals and stochastic differential equations*, Ann. Appl. Probab., 19 (1991), pp. 1035–1070.
[14] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, New York, 2001.
[15] T. Leung, Q. Song, and J. Yang, *Outperformance portfolio optimization via the equivalence of pure and randomized hypothesis testing*, Finance Stoch., 17 (2013), pp. 839–870.
[16] X. Mao, C. Yuan, and G. Yin, *Approximations of Euler-Maruyama type for stochastic differential equations with Markovian switching, under non-Lipschitz conditions*, J. Comput. Appl. Math., 205 (2007), pp. 936–948.
[17] *A Non-Degenerate Martingale*, http://mathoverflow.net/questions/84216.
[18] A. Metzler, *On the first passage problem for correlated Brownian motion*, Statist. Probab. Lett., 80 (2010), pp. 277–284.
[19] P. Mörters and Y. Peres, *Brownian Motion*, Cambridge University Press, Cambridge, UK, 2010.
[20] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, New York, 1999.
[21] H. P. Rosenthal, *On the subspaces of* $\mathbb{L}^p$ $(p \geq 2)$ *spanned by sequences of independent random variables*, Israel J. Math., 8 (1970), pp. 273–303.
[22] Q. Song, G. Yin, and Q. Zhang, *Weak convergence methods for approximation of the evaluation of path-dependent functionals*, SIAM J. Control Optim., 51 (2013), pp. 4189–4210.
[23] Q. Song, G. Yin, and C. Zhu, *Optimal switching with constraints and utility maximization of an indivisible market*, SIAM J. Control Optim., 50 (2012), pp. 629–651.
[24] L. Yan, *The Euler scheme with irregular coefficients*, Ann. Probab., 30 (2002), pp. 1172–1194.
[25] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, New York, 1999.
[26] H. Zheng, *Contagion models a la carte: Which one to choose?*, Quant. Finance, 13 (2013), pp. 399–405.