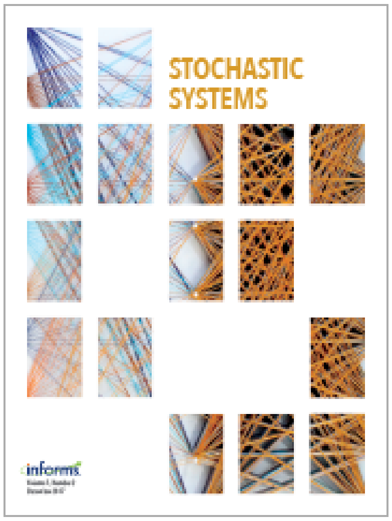


Stochastic Systems

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Monte Carlo Method for Estimating Sensitivities of Reflected Diffusions in Convex Polyhedral Domains

David Lipshutz, Kavita Ramanan

informs
http://pubsonline.informs.org/journal/stsy/

STOCHASTIC SYSTEMS
Vol. 9, No. 2, June 2019, pp. 101–140
ISSN 1946-5238 (online)

A Monte Carlo Method for Estimating Sensitivities of Reflected Diffusions in Convex Polyhedral Domains

David Lipshutz,[a] Kavita Ramanan[b]

[a] Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel; [b] Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

Contact: lipshutz@technion.ac.il, http://orcid.org/0000-0001-9347-8326 (DL); kavita_ramanan@brown.edu (KR)

Received: December 3, 2017
Revised: September 3, 2018; February 12, 2019
Accepted: February 25, 2019
Published Online in Articles in Advance: June 14, 2019

https://doi.org/10.1287/stsy.2019.0031

Abstract. We develop an effective Monte Carlo method for estimating sensitivities or gradients of expectations of sufficiently smooth functionals of a reflected diffusion in a convex polyhedral domain with respect to its defining parameters, namely its initial condition, drift, and diffusion coefficients and directions of reflection. Our method, which falls into the class of infinitesimal perturbation analysis (IPA) methods, uses a probabilistic representation for such sensitivities as the expectation of a functional of the reflected diffusion and its associated derivative process. The latter process is the unique solution to a constrained linear stochastic differential equation with jumps whose coefficients, domain, and directions of reflection are modulated by the reflected diffusion. We propose a consistent estimator for such sensitivities using an Euler approximation of the reflected diffusion and its associated derivative process. Proving that the Euler approximation converges is challenging because the derivative process jumps whenever the reflected diffusion hits the boundary of the domain. A key step in the proof is establishing a continuity property of the related derivative map, which is of independent interest. We compare the performance of our IPA estimator with a standard likelihood ratio estimator (whenever the latter is applicable) and provide numerical evidence that the variance of the former is substantially smaller than that of the latter. We illustrate our method with an example of a rank-based interacting diffusion model of equity markets. Interestingly, we show that estimating certain sensitivities of the rank-based interacting diffusion model using our method for a reflected Brownian motion description of the model outperforms a finite difference method for a stochastic differential equation description of the model.

Funding: D. Lipshutz was supported in part by the National Science Foundation (NSF) Divisions of Mathemetical Sciences [Grant 1148284] and Civil, Mechanical and Manufacturing Innovation (CMMI) [Grant 1407504]. K. Ramanan was supported in part by the NSF CMMI [Grant 1407504].

Keywords: reflected diffusion • reflected Brownian motion • Monte Carlo method • sensitivity analysis • infinitesimal perturbation analysis • pathwise differentiability • derivative process • derivative map • boundary jitter property • rank-based interacting diffusions • Atlas model

1. Introduction

1.1. Overview

Reflected diffusions in convex polyhedral domains arise in numerous applications. For instance, they arise in the study of rank-based interacting diffusion models in mathematical finance (Banner et al. 2005, Ichiba et al. 2011) and as diffusion approximations in queueing theory (Reiman 1984; Chen and Mandelbaum 1991; Peterson 1991; Mandelbaum and Pats 1998; Ramanan and Reiman 2003, 2008). For use in uncertainty quantification, stochastic optimization and other areas (see chapter VII of Asmussen and Glynn 2007 for a list of the numerous applications), it is of interest to estimate the gradient of the expectation of a functional of a reflected diffusion with respect to its defining parameters—namely, its initial condition, drift, and diffusion coefficients and directions of reflection along the boundary of its domain. (Henceforth, we simply write "sensitivities" to mean "gradients of expectations of functionals.") The main contribution of this work is to develop a broadly applicable consistent estimator of a large class of sensitivities of reflected diffusions, which can be used to approximate sensitivities of reflected diffusions via a Monte Carlo method.

Our estimator is based on a probabilistic representation for sensitivities of reflected diffusions that was obtained in Lipshutz and Ramanan (2019a) and reproduced in Theorem 1. The representation expresses the sensitivity as the expectation of a functional of a reflected diffusion and its associated derivative process. The derivative process satisfies a constrained linear stochastic differential equation (SDE) with jumps whose coefficients, domain, and directions of reflection are modulated by the reflected diffusion, and it was shown in theorem 3.13 of Lipshutz and Ramanan (2019a) that the pathwise derivatives of a reflected diffusion can be described via the derivative process. Although this representation provides an unbiased estimator for sensitivities of a reflected diffusion, computation of sensitivities via this representation would entail simulation of the reflected diffusion and its associated derivative process, which typically involves discrete-time approximations. There is a large literature devoted to approximating reflected diffusions in convex polyhedral domains (see, e.g., Liu 1993; Słomínski 1994, 2001; Lépingle 1995; Milshtein 1995; Pettersson 1995, 1997; Costantini et al. 1998; Gobet 2001; Bossy et al. 2004; and Blanchet and Murthy 2018); however, there is no method for approximating the derivative process. In this work, we propose an Euler scheme for the approximation of a reflected diffusion and its derivative process and prove that the associated estimators of sensitivities of the reflected diffusion are consistent as the discretization parameter goes to zero. The proof that the Euler scheme for the reflected diffusion is consistent follows an argument analogous to the one used in the proof of theorem 3.2 of Slominski (2001) which established the result in the case of normal reflection. The proof that the Euler scheme for the derivative process is consistent is much more challenging because the derivative process jumps whenever the reflected diffusion hits the boundary of the polyhedral domain. Establishing convergence of the discrete approximation is quite subtle and relies on a continuity property of the so-called derivative map (see Theorem 4 and Definition 5), which is a deterministic map introduced in Lipshutz and Ramanan (2018) to characterize directional derivatives of the associated extended Skorokhod map (ESM). The proof of this continuity property, which is carried out in Section 9, is the main technical challenge addressed in this work and requires a detailed understanding of the dynamics of the derivative map. This continuity property is also of independent interest; for example, it is used in Lipshutz and Ramanan (2019b) to prove that a reflected Brownian motion (RBM) in a convex polyhedral cone and its derivative process are jointly a Feller continuous Markov process.

Our method, which relies on pathwise derivatives of a reflected diffusion, falls into the class of infinitesimal perturbation analysis (IPA) methods used in sensitivity analysis (see, e.g., chapter VII.2 of Asmussen and Glynn 2007). We compare the performance of our IPA method to a standard likelihood ratio (LR) method and provide numerical evidence that the variance of the former is substantially smaller than that of the latter. It is also worth mentioning here that the LR method generally only applies to perturbations of the drift, and in many applications, it is of interest to study perturbations with respect to all of the parameters (e.g., when estimating sensitivities of diffusion approximations of queueing networks, which we plan to investigate in future work). We also apply our method to study a particular rank-based interacting diffusion model called the *Atlas model*, originally proposed by Fernholz (2002), see example 5.3.3, to model equity markets and subsequently generalized by Banner et al. (2005) and Ichiba et al. (2011). This model is described by an SDE with discontinuous drift coefficients, and its sensitivities can be estimated using a standard finite difference (FD) method (although this method remains biased as the time-discretization vanishes). On the other hand, the dynamics of this model can also be expressed in terms of an RBM in a convex polyhedral domain (see, e.g., Fernholz 2002 and Section 5.2). We estimate sensitivities by applying our method to this RBM representation and provide numerical evidence to show that the empirical variance of the FD estimator blows up as the derivative approximation parameter approaches zero (even when using common random numbers).

In summary, the main contributions of this work are as follows:

- An IPA method for estimating sensitivities of a reflected diffusion (Algorithm 1).
- Proof of convergence of Euler schemes for a reflected diffusion (Theorem 2) and its associated derivative process (Theorem 3).
- Comparison of our method with the LR method with numerical evidence that our method performs better, especially over long time horizons (Section 4).
- Illustration of the efficacy of our method on a one-dimensional RBM, for which analytic expressions for the sensitivities are available (Section 5.1).
- Application of our method to estimate certain sensitivities of the Atlas model and numerical evidence that our method applied to the RBM representation of the model performs better than the FD method applied to an SDE description of the model (Section 5.2).
- A continuity property of the derivative map (Theorem 4).

We conclude by making two observations. First, Algorithm 1 assumes existence of a unique *projection map* as well as certain *derivative projection matrices*. Even simulation of a reflected diffusion alone (i.e., without the derivative process) requires an expression for the projection map, which is often nontrivial. Despite this, we are unaware of a general explicit expression for the projection map in the simulation literature. In the appendix, we provide explicit expressions for both the projection map and the derivative projection matrices. In this way, our work also contributes to the literature on simulation of reflected diffusions. We note that the computational complexity of implementing our explicit expression for the projection map and derivative projection matrices increases exponentially with the number of faces on the polyhedral domain; however, as explained in Remark A.1, for most situations, we do not see this as a significant impediment to implementing our algorithm.

The second observation we make is that we do not prove a rate of convergence for our IPA method as the discretization parameter converges to zero. We defer this to future work. Proving a rate of convergence presents significant new challenges because the derivative map does not satisfy a certain *uniform* continuity property. For a further explanation of the technical difficulties, see Remark 16.

1.2. Outline

Following this introduction, the remainder of this paper is organized into eight sections and an appendix. Sections 2 and 3 and the appendix contain the relevant background information and main results required to implement our IPA algorithm, Sections 4 and 5 compare our IPA algorithm to existing algorithms, and Sections 6–9 contain the proofs of our results. Specifically, in Section 2, we give precise definitions of a reflected diffusion and its associated derivative process, and we recall the probabilistic representation of sensitivities of reflected diffusions from Lipshutz and Ramanan (2019a). In Section 3, we define an Euler approximation for a reflected diffusion and its derivative process, state our main convergence result, and describe a numerical algorithm for estimating sensitivities. In Section 4, we compare our algorithm with an LR algorithm for gradient estimation in cases in which the latter applies. In Section 5, we present numerical results from applying our method to a one-dimensional RBM and to the Atlas model. In Section 6, we define and state properties of the Skorokhod problem (SP) and the derivative problem, which are used to characterize a reflected diffusion and its derivative process, respectively. These are used in the proofs of our main results, which are presented in Sections 7–9.

1.3. Notation

Let $\mathbb{N} := \{1, 2, \ldots\}$ denote the set of positive integers $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$. Given $J \in \mathbb{N}$, we use $\mathbb{R}^J_+$ to denote the closed nonnegative orthant in J-dimensional Euclidean space $\mathbb{R}^J$. When $J = 1$, we suppress J and write $\mathbb{R}$ for $(-\infty, \infty)$ and $\mathbb{R}_+$ for $[0, \infty)$. For $r, s \in \mathbb{R}$, let $\lfloor r \rfloor = \max\{k \in \mathbb{N}_0 : k \leq r\}$, $r \wedge s = \min(r, s)$, and $r \vee s = \max(r, s)$. For a subset $A \subset \mathbb{R}$, let $\inf A$ and $\sup A$ denote the infimum and supremum, respectively, of A with the convention that the infimum and supremum of the empty set are, respectively, defined to be ∞ and $-\infty$. For a column vector $x \in \mathbb{R}^J$, let x^j denote the jth component of x for $j = 1, \ldots, J$ and let $|x|$ denote the usual Euclidean norm of x. We let $\{e_1, \ldots, e_J\}$ denote the standard normal basis in $\mathbb{R}^J$, where e_j is the column vector in $\mathbb{R}^J$ whose jth component is one and whose other components are zero for $j = 1, \ldots, J$. We let $\mathbb{S}^{J-1} = \{x \in \mathbb{R}^J : |x| = 1\}$ denote the unit sphere in $\mathbb{R}^J$. For $J, K \in \mathbb{N}$, let $\mathbb{R}^{J \times K}$ denote the set of real-valued matrices with J rows and K columns. We write $M^T \in \mathbb{R}^{K \times J}$ for the transpose of a matrix $M \in \mathbb{R}^{J \times K}$. Let E_J denote the $J \times J$ identity matrix. Let $\|\cdot\|$ denote the operator norm on $\mathbb{R}^{J \times K}$.Given $E \subseteq \mathbb{R}^J$, we let $\text{cone}(E)$ denote the convex cone generated by E; that is,

$$\text{cone}(E) := \left\{ \sum_{k=1}^{K} r_k x_k, K \in \mathbb{N}, x_k \in E, r_k \in \mathbb{R}_+ \right\},$$

and let $\text{span}(E)$ denote the set of all possible finite linear combinations of vectors in E with the convention that $\text{cone}(\emptyset)$ and $\text{span}(\emptyset)$ are equal to $\{0\}$. Given $E \subseteq \mathbb{R}^J$ or $E \subseteq \mathbb{R}^{J \times K}$, let $\mathbb{D}(E)$ denote the set of functions mapping $[0, \infty)$ to E that are right-continuous with finite left limits (RCLL). Let $\mathbb{C}(E)$ denote the subset of continuous functions in $\mathbb{D}(E)$. For a subset $A \subseteq E$, let $\mathbb{D}_A(E) := \{f \in \mathbb{D}(E) : f(0) \in A\}$ and $\mathbb{C}_A(E) := \{f \in \mathbb{C}(E) : f(0) \in A\}$. We endow $\mathbb{D}(E)$ and its subsets with the J_1-topology and recall that the J_1-topology relativized to $\mathbb{C}(E)$ coincides with the topology of uniform convergence on compact subsets of $[0, \infty)$. Given $f \in \mathbb{D}(E)$ and $t > 0$, we let $f(t-)$ denote the left limit of $f(\cdot)$ at t. Given a function $f : (0, \infty) \mapsto \mathbb{R}$, we say $f(\varepsilon) = o(\varepsilon)$ as $\varepsilon \downarrow 0$ if $|f(\varepsilon)|/\varepsilon \to 0$ as $\varepsilon \downarrow 0$, and we say $f(\varepsilon) = O(\varepsilon)$ as $\varepsilon \downarrow 0$ if $|f(\varepsilon)|/\varepsilon$ is bounded as $\varepsilon \downarrow 0$.

Throughout this paper, we fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, satisfying the usual conditions; that is, $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets in $\mathcal{F}$, and the filtration $\{\mathcal{F}_t\}$ is right-continuous. We write $\mathbb{E}$ to denote expectation under $\mathbb{P}$. We abbreviate “almost surely” as “a.s.” By a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$, we mean that $(W^1, \ldots, W^K)$ are independent, and for each $k = 1, \ldots, K$, $\{W^k(t), \mathcal{F}_t, t \geq 0\}$ is a continuous martingale with $W^k(0) = 0$ and quadratic variation $[W^k]_t = t$ for $t \geq 0$. We let $C_p < \infty$ for $p \geq 2$ denote the universal constants in the Burkholder–Davis–Gundy (BDG) inequalities (see, e.g., Rogers and Williams 2000, chapter IV, theorem 42.1).

2. Background on Reflected Diffusions and Their Sensitivities

2.1. A Parameterized Family of Reflected Diffusions

Let G be a closed polyhedron in $\mathbb{R}^J$ equal to the intersection of finitely many closed half spaces in $\mathbb{R}^J$; that is,

$$G := \bigcap_{i=1,\dots,N} \{x \in \mathbb{R}^J : \langle x, n_i \rangle \geq c_i\}, \tag{1}$$

for a positive integer $N \in \mathbb{N}$, unit vectors $n_i \in \mathbb{S}^{J-1}$, and constants $c_i \in \mathbb{R}$ for $i = 1, \dots, N$. Let $\mathcal{I} = \{1, \dots, N\}$. For $i \in \mathcal{I}$, we let $F_i := \{x \in \partial G : \langle x, n_i \rangle = c_i\}$ denote the ith face of G. For $x \in G$, we write $\mathcal{I}(x) := \{i \in \mathcal{I} : x \in F_i\}$ to denote the (possibly empty) set of indices associated with the faces that intersect at x.

Let $M \in \mathbb{N}$ and U be an open *parameter set* in $\mathbb{R}^M$. For each $i \in \mathcal{I}$, fix a continuously differentiable function $d_i : U \mapsto \mathbb{R}^J$ that satisfies $\langle d_i(\alpha), n_i \rangle = 1$ for all $\alpha \in U$. For $\alpha \in U$, $d_i(\alpha)$ denotes the (constant) direction of reflection along the face F_i associated with the parameter α. Because the direction of reflection $d_i(\alpha)$ can be renormalized, our assumption $\langle d_i(\alpha), n_i \rangle = 1$ for all $\alpha \in U$ is without loss of generality. We let

$$R : U \mapsto \mathbb{R}^{J \times N}$$

denote the continuous differentiable function defined by $R(\alpha) := \big(d_1(\alpha) \cdots d_N(\alpha)\big)$ for $\alpha \in U$ and let $R'(\alpha)$ denote the Jacobian of $R(\cdot)$ at $\alpha \in U$. We refer to $R(\alpha)$ as the reflection matrix associated with α. Fix continuously differentiable functions

$$x_0 : U \mapsto G, \qquad b : U \times G \mapsto \mathbb{R}^J, \qquad \sigma : U \times G \mapsto \mathbb{R}^{J \times K}.$$

For $\alpha \in U$, $x_0(\alpha)$, $b(\alpha, \cdot)$, and $\sigma(\alpha, \cdot)$ are, respectively, the initial condition, drift, and dispersion coefficients for the reflected diffusion associated with α. We refer to $a(\alpha, \cdot) := \sigma(\alpha, \cdot)\sigma^T(\alpha, \cdot)$ as the diffusion coefficient associated with $\alpha \in U$. For $\alpha \in U$ and $x \in G$, we let $x_0'(\alpha)$ denote the Jacobian of $x_0(\cdot)$ at $\alpha \in U$, $b_\alpha(\alpha, x)$ denote the Jacobian of $b(\cdot, x)$ at $\alpha \in U$, and $b_x(\alpha, x)$ denote the Jacobian of $b(\alpha, \cdot)$ at $x \in G$ and similarly define $\sigma_\alpha(\alpha, x)$ and $\sigma_x(\alpha, x)$.

Definition 1. Given $\alpha \in U$ and a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$, a reflected diffusion associated with α and driving Brownian motion W is a J-dimensional continuous $\{\mathcal{F}_t\}$-adapted process $Z^\alpha = \{Z^\alpha(t), t \geq 0\}$ such that a.s. for all $t \geq 0$, $Z^\alpha(t) \in G$ and Z^α satisfies

$$Z^\alpha(t) = x_0(\alpha) + \int_0^t b(\alpha, Z^\alpha(s))ds + \int_0^t \sigma(\alpha, Z^\alpha(s))dW(s) + R(\alpha)L^\alpha(t), \qquad t \geq 0, \tag{2}$$

where $L^\alpha = \{L^\alpha(t), t \geq 0\}$ is an N-dimensional continuous $\{\mathcal{F}_t\}$-adapted process such that a.s. $L^\alpha(0) = 0$, and for every $i \in \mathcal{I}$, the ith component of L^α, denoted $L^{\alpha,i}$, is nondecreasing and can only increase when Z^α lies in face F_i; that is,

$$\int_0^\infty 1_{\{Z^\alpha(s) \notin F_i\}} dL^{\alpha,i}(s) = 0.$$

We say that *pathwise uniqueness* holds if, given $\alpha \in U$, a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$ and any two reflected diffusions, Z^α and $\widetilde{Z}^\alpha$, associated with α and driving Brownian motion W, then $\mathbb{P}(Z^\alpha(t) = \widetilde{Z}^\alpha(t)$ for all $t \geq 0) = 1$.

Remark 1. In definition 2.1 of Lipshutz and Ramanan (2019a) the authors define a family of reflected diffusions in which the drift and dispersion coefficients and directions of reflection are parameterized by $\alpha \in U$, but the initial condition is parameterized by $x \in G$. In Lipshutz and Ramanan (2019a), this allows for a characterization of pathwise derivatives of flows of reflected diffusions and is a convenient representation in the proofs. In contrast, here we find it more convenient to assume that the initial condition is a continuously differentiable function $x_0(\cdot)$ on U, taking values in G.

Remark 2. Given $\alpha \in U$ and a reflected diffusion Z^α with L^α satisfying the conditions in Definition 1, define the J-dimensional $\{\mathcal{F}_t\}$-adapted constraining process $Y^\alpha = \{Y^\alpha(t), t \geq 0\}$ by $Y^\alpha := R(\alpha)L^\alpha$. It follows from the definition of $R(\alpha)$ and the conditions on L^α in Definition 1 that Y^α satisfies, for all $0 \leq s < t < \infty$,

$$Y^\alpha(t) - Y^\alpha(s) \in \text{cone}\left[\cup_{u \in (s,t]} d(\alpha, Z^\alpha(u))\right], \tag{3}$$

where

$$d(\alpha,x) := \text{cone}\{d_i(\alpha), i \in \mathcal{I}(x)\}, \qquad \alpha \in U,\ x \in G. \tag{4}$$

In particular, we see that Z^α satisfies definition 2.1 of Lipshutz and Ramanan (2019a) for a reflected diffusion (with initial condition $x = x_0(\alpha)$ there). The more general condition (3) allows for reflected diffusions that are not semimartingles. In this work, we impose a mild linear independence condition on the directions of reflection (see Assumption 1). As explained in remark 2.12 of Lipshutz and Ramanan (2019a) under this condition, Z^α satisfies Definition 1 if and only if Z^α satisfies definition 2.1 of Lipshutz and Ramanan (2019a).

Let ζ_1, ζ_2 lie in $C_b^1(G)$, the space of real-valued functions that are continuously differentiable with bounded first partial derivatives. Suppose that, for each $\alpha \in U$, there exists a unique reflected diffusion Z^α associated with α. Then, for $t \geq 0$, define $\Theta_t(\cdot)$ to be the mapping from U to $\mathbb{R}$ defined by

$$\Theta_t(\alpha) := \mathbb{E}\left[\int_0^t \zeta_1(Z^\alpha(s))ds + \zeta_2(Z^\alpha(t))\right], \qquad \alpha \in U. \tag{5}$$

In the next two sections, we state our main assumptions, introduce the derivative process along Z^α, and characterize the Jacobian of $\Theta_t(\cdot)$ in terms of Z^α and the derivative process along Z^α.

2.2. Main Assumptions

The assumptions stated in this section are assumed to hold, without restatement, throughout this work.

The first three assumptions on the data $\{(d_i(\cdot), n_i, c_i), i \in \mathcal{I}\}$ ensure the associated SP is well defined and the associated Skorokhod map (SM) is Lipschitz continuous (see Proposition 1), which is useful for proving strong existence of reflected diffusions and establishing pathwise uniqueness.

Assumption 1. *For each $\alpha \in U$ and $x \in \partial G$, $\{d_i(\alpha), i \in \mathcal{I}(x)\}$ is a set of linearly independent vectors.*

Given a convex set B, let $\nu_B(z) := \{\nu \in \mathbb{S}^{J-1} : \langle \nu, y - z\rangle \geq 0\ \forall\ y \in B\}$ denote the set of inward normal vectors to the set B at $z \in \partial B$.

Assumption 2. *For each $\alpha \in U$, there exists $\delta(\alpha) > 0$ and a compact, convex, symmetric set B^α in $\mathbb{R}^J$ with $0 \in (B^\alpha)^\circ$ such that, for $i \in \mathcal{I}$,*

$$\left\{\begin{array}{l} z \in \partial B^\alpha \\ |\langle z, n_i\rangle| < \delta(\alpha) \end{array}\right\} \qquad \Rightarrow \qquad \langle \nu, d_i(\alpha)\rangle = 0 \qquad \textit{for all} \ \ \nu \in \nu_{B^\alpha}(z). \tag{6}$$

Recall the definition of $d(\cdot,\cdot)$ given in (4). The following assumption guarantees that, for each $\alpha \in U$, there exists a map that projects points outside of G to the boundary ∂G in a way that is compatible with the vector field $d(\alpha,\cdot)$ on ∂G. Along with Assumption 2, the existence of such a map ensures that the directions of reflection along the boundary ∂G can constrain any path to remain in G.

Assumption 3. *For each $\alpha \in U$, there is a map $\pi^\alpha : \mathbb{R}^J \mapsto G$ satisfying $\pi^\alpha(x) = x$ for all $x \in G$ and $\pi^\alpha(x) - x \in d(\alpha, \pi^\alpha(x))$ for all $x \notin G$.*

Remark 3. Under Assumption 2, there can be at most one map $\pi^\alpha : \mathbb{R}^J \mapsto G$ that satisfies the conditions stated in Assumption 3 (see, e.g., the paragraph before lemma 4.3 of Dupuis and Ramanan 1999). See the appendix for an explicit expression for the projection map π^α in terms of the reflection matrix.

Remark 4. See example 2.14 of Lipshutz and Ramanan (2018) for a general set of algebraic conditions on $\{(d_i(\cdot), n_i, c_i), i \in \mathcal{I}\}$ that imply Assumptions 1–3 hold.

Along with the last three assumptions, the final two assumptions on the coefficients $b(\cdot,\cdot)$, $\sigma(\cdot,\cdot)$, and $R(\cdot)$ ensure existence and pathwise uniqueness of the reflected diffusion and the derivative process as well as the characterization of sensitivities of reflected diffusions in terms of the derivative process (see Theorem 1).

Assumption 4. *There exists $\kappa_1 < \infty$ such that, for all $\alpha \in U$ and $x \in G$,*

$$\max\{\|b_\alpha(\alpha,x)\|, \|b_x(\alpha,x)\|, \|\sigma_\alpha(\alpha,x)\|, \|\sigma_x(\alpha,x)\|, \|R'(\alpha)\|\} \leq \kappa_1.$$

In addition, there exist $\kappa_2 < \infty$ and $\gamma \in (0,1]$ such that, for all $\alpha, \beta \in U$ and $x, y \in G$,

$$\max\left\{\begin{array}{ll} \|b_\alpha(\alpha,x) - b_\alpha(\beta,y)\|, & \|b_x(\alpha,x) - b_x(\beta,y)\|, \\ \|\sigma_\alpha(\alpha,x) - \sigma_\alpha(\beta,y)\|, & \|\sigma_x(\alpha,x) - \sigma_x(\beta,y)\|, \\ \|R'(\alpha) - R'(\beta)\| & \end{array}\right\} \le \kappa_2 |(\alpha,x) - (\beta,y)|^\gamma.$$

Assumption 5. *For each $\alpha \in U$, there exists $c(\alpha) > 0$ such that, for all $x \in G$,*

$$y^T a(\alpha,x) y \ge c(\alpha)|y|^2, \qquad y \in \mathbb{R}^J. \tag{7}$$

2.3. Sensitivities of Reflected Diffusions in Terms of the Derivative Process

We first define the derivative process, which was introduced in definition 3.5 of Lipshutz and Ramanan (2019a) to characterize pathwise derivatives of a reflected diffusion. For $x \in \partial G$, define

$$\mathbb{H}_x := \bigcap_{i \in \mathcal{I}(x)} \{y \in \mathbb{R}^J : \langle y, n_i \rangle = 0\}, \tag{8}$$

and for $x \in G^\circ$, set $\mathbb{H}_x := \mathbb{R}^J$. Recall the definition of $d(\cdot,\cdot)$ in (4).

Definition 2. Let $\alpha \in U$, W be a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, and Z^α be a reflected diffusion associated with α and driving Brownian motion W. A derivative process along Z^α is an RCLL $\{\mathcal{F}_t\}$-adapted process $\mathcal{J}^\alpha = \{\mathcal{J}^\alpha(t), t \ge 0\}$ taking values in $\mathbb{R}^{J\times M}$ such that a.s., for all $t \ge 0$, $\mathcal{J}^\alpha_m(t) \in \mathbb{H}_{Z^\alpha(t)}$ for $m = 1, \ldots, M$ and $\mathcal{J}^\alpha$ satisfies

$$\begin{aligned}\mathcal{J}^\alpha(t) = x_0'(\alpha) &+ \int_0^t b_\alpha(\alpha, Z^\alpha(s))ds + \int_0^t b_x(\alpha, Z^\alpha(s))\mathcal{J}^\alpha(s)ds + \int_0^t \sigma_\alpha(\alpha, Z^\alpha(s))dW(s) \\ &+ \int_0^t \sigma_x(\alpha, Z^\alpha(s))\mathcal{J}^\alpha(s)dW(s) + R'(\alpha)L^\alpha(t) + \mathcal{K}^\alpha(t),\end{aligned} \tag{9}$$

where $\mathcal{K}^\alpha = \{\mathcal{K}^\alpha(t), t \ge 0\}$ is an RCLL $\{\mathcal{F}_t\}$-adapted process taking values in $\mathbb{R}^{J\times M}$ such that a.s. $\mathcal{K}^\alpha(0) = 0$, and for each $m = 1, \ldots, M$ and all $0 \le s < t < \infty$,

$$\mathcal{K}^\alpha_m(t) - \mathcal{K}^\alpha_m(s) \in \operatorname{span}\left[\cup_{u \in (s,t]} d(\alpha, Z^\alpha(u))\right]. \tag{10}$$

We say that *pathwise uniqueness* holds if, given $\alpha \in U$, a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W, a reflected diffusion Z^α associated with $\alpha \in U$ with driving Brownian motion W, and any two derivative processes $\mathcal{J}^\alpha$ and $\widetilde{\mathcal{J}}^\alpha$ along Z^α, then $\mathbb{P}(\mathcal{J}^\alpha(t) = \widetilde{\mathcal{J}}^\alpha(t)$ for all $t \ge 0) = 1$.

In Lipshutz and Ramanan (2019a), it was shown that a.s. the reflected diffusion $Z^\alpha(t)$ is differentiable with respect to α at each $t \ge 0$, and the right-continuous regularization of its so-called pathwise derivative is equal to the derivative process $\mathcal{J}^\alpha$ along Z^α. With this in mind, it is natural to view Equation (9) for the derivative process as a linearized version of Equation (2) for the reflected diffusion in the sense that (9) is obtained by formally differentiating each term in (2) with respect to α with $\mathcal{J}^\alpha$ serving as the appropriate linearization of the reflected diffusion Z^α and $R'(\alpha)L^\alpha + \mathcal{K}^\alpha$ serving as the appropriate linearization of the constraining process $R(\alpha)L^\alpha$.

Remark 5. Definition 2 for the derivative process is slightly different from the definition given in definition 3.5 of Lipshutz and Ramanan (2019a) because the initial condition here is a function of $\alpha \in U$. To clarify the relation, suppose $\mathcal{J}^\alpha$ satisfies Definition 2 and $\mathcal{J}^{\alpha,x}$ satisfies definition 3.5 of Lipshutz and Ramanan (2019a). Then, for $m = 1, \ldots, M$, the mth column vector of $\mathcal{J}^\alpha$ satisfies $\mathcal{J}^\alpha_m(t) = \mathcal{J}^{\alpha,x_0(\alpha)}_t[e_m, x_0'(\alpha)e_m]$ for all $t \ge 0$, where the right-hand side of the equality is written in the notation of Lipshutz and Ramanan (2019a).

In the following theorem, we state the probabilistic representation of sensitivities of reflected diffusions that was obtained in Lipshutz and Ramanan (2019a) and serves as the starting point for the method for sensitivity estimation that we introduce in this work. Recall that the assumptions stated in Section 2.2 hold.

Theorem 1. *For each $\alpha \in U$ and K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$, the following hold:*

i. *There exists a pathwise unique reflected diffusion Z^α associated with α and driving Brownian motion W, and it is a strong Markov process.*

ii. *There exists a pathwise unique derivative process $\mathcal{J}^\alpha$ along Z^α.*
iii. *Given $t \geq 0$, a.s. $\mathcal{J}^\alpha(\cdot)$ is continuous at t, and a.s. $\mathcal{J}^\alpha(\cdot)$ is continuous at almost every $s \in [0,t)$.*
iv. *Given $t \geq 0$, the function $\Theta_t(\cdot)$ defined in (5) is differentiable at α and its Jacobian satisfies*

$$\Theta_t'(\alpha) = \mathbb{E}\left[\int_0^t \zeta_1'(Z^\alpha(s))\mathcal{J}^\alpha(s)ds + \zeta_2'(Z^\alpha(t))\mathcal{J}^\alpha(t)\right]. \tag{11}$$

Proof. Part (i) follows from proposition 2.16 of Lipshutz and Ramanan (2019a) (see also theorem 4.3 of Ramanan 2006) and the equivalence between solutions of Definition 1 and Lipshutz and Ramanan (2019a), definition 2.1, stated in Remark 2. Parts (ii) and (iii) follow from corollary 3.15 of Lipshutz and Ramanan (2019a) and Remark 5. Part (iv) follows from corollary 3.16 of Lipshutz and Ramanan (2019a) and the chain rule. □

Although (11) provides an unbiased estimator for $\Theta_t'(\alpha)$, exact sampling of functionals of $(Z^\alpha, \mathcal{J}^\alpha)$ is complicated by the discontinuous dynamics when Z^α reaches the boundary ∂G. As is often the case when simulating diffusion processes, we sample from a discrete-time Euler approximation of $(Z^\alpha, \mathcal{J}^\alpha)$. Our main result (see Corollary 2) states that the Euler approximation can be used to construct a consistent estimator for $\Theta_t'(\alpha)$.

3. Main Results

Recall that the assumptions stated in Section 2.2 hold. Fix $\alpha \in U$ and a K-dimensional $\{\mathcal{F}_t\}$-Brownian motion W on $(\Omega, \mathcal{F}, \mathbb{P})$. Let Z^α denote the pathwise unique reflected diffusion associated with α and driving Brownian motion W and let $\mathcal{J}^\alpha$ denote the pathwise unique derivative process along Z^α. Throughout this section, given a time step $\Delta > 0$, define the sequence $\{t_n^\Delta\}_{n\in\mathbb{N}_0}$ by

$$t_n^\Delta := n\Delta, \qquad n \in \mathbb{N}_0, \tag{12}$$

and let $\{\delta_n^\Delta W\}_{n\in\mathbb{N}}$ be the sequence of independent and identically distributed K-dimensional Gaussian random variables with mean zero and diagonal covariance matrix ΔE_K given by

$$\delta_n^\Delta W := W(t_n^\Delta) - W(t_{n-1}^\Delta), \qquad n \in \mathbb{N}, \tag{13}$$

where we recall that E_K denotes the $K \times K$ identity matrix. In addition, let $\{\mathcal{F}_t^\Delta\}$ denote the discrete filtration defined by

$$\mathcal{F}_t^\Delta := \sigma\{\delta_k^\Delta W, k = 1, \ldots, n-1\}, \qquad t \in [t_{n-1}^\Delta, t_n^\Delta), \qquad n \in \mathbb{N}. \tag{14}$$

3.1. Euler Scheme for the Reflected Diffusion

In this section, we present an Euler scheme for approximating just the reflected diffusion Z^α. This is an extension of a result obtained in Słomínski (2001) to allow for reflected diffusions with oblique reflection. We also prove a convergence result for an Euler approximation of the process L^α that is needed in the next section. Let π^α denote the unique mapping satisfying the conditions in Assumption 3. To define the Euler scheme, we need the following lemma.

Lemma 1. *There exists a unique map $\xi^\alpha : \mathbb{R}^J \mapsto \mathbb{R}_+^N$ such that, for each $x \in \mathbb{R}^J$, $\pi^\alpha(x) - x = R(\alpha)\xi^\alpha(x)$, and the ith component of $\xi^\alpha(x)$ satisfies $\xi^{\alpha,i}(x) > 0$ only if $\pi^\alpha(x) \in F_i$ for $i \in \mathcal{I}$.*

Remark 6. In the case of a simple polyhedral cone with vertex at the origin (i.e., $N = J$ and $c_i = 0$ for $i \in \mathcal{I}$), Assumption 1 implies $R(\alpha)$ is invertible for each $\alpha \in U$, and so $\xi^\alpha(x) = (R(\alpha))^{-1}(\pi^\alpha(x) - x)$ for each $\alpha \in U$ and $x \in \mathbb{R}^J$. See the appendix for an explicit expression for the map ξ^α in terms of the reflection matrix.

Proof. According to Assumption 3 and Remark 3, $\pi^\alpha : \mathbb{R}^J \mapsto G$ is the unique mapping that satisfies $\pi^\alpha(x) - x \in d(\alpha, \pi^\alpha(x))$. By the definition of $d(\cdot,\cdot)$ in (4) and the linear independence of the vectors $\{d_i(\alpha), i \in \mathcal{I}(\pi^\alpha(x))\}$ guaranteed by Assumption 1, it follows that, for each $x \in \mathbb{R}^J$, there is a unique vector $\xi^\alpha(x) \in \mathbb{R}_+^N$ satisfying the conditions of the lemma. □

Given $\Delta > 0$, recall the sequence $\{\delta_n^\Delta W\}_{n\in\mathbb{N}}$ defined in (13). Let $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ denote the pair of piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted processes taking values in $G \times \mathbb{R}_+^N$ defined as follows: Set

$$(Z^{\alpha,\Delta}(0), L^{\alpha,\Delta}(0)) := (x_0(\alpha), 0) \tag{15}$$

and, for $n \in \mathbb{N}$, set

$$(Z^{\alpha,\Delta}(t), L^{\alpha,\Delta}(t)) := (Z^{\alpha,\Delta}(t_{n-1}^{\Delta}), L^{\alpha,\Delta}(t_{n-1}^{\Delta})), \qquad t \in [t_{n-1}^{\Delta}, t_n^{\Delta}), \tag{16}$$

and recursively define

$$(Z^{\alpha,\Delta}(t_n^{\Delta}), L^{\alpha,\Delta}(t_n^{\Delta})) := (\pi^{\alpha}(\Xi_n^{\alpha,\Delta}), L^{\alpha,\Delta}(t_{n-1}^{\Delta}) + \xi^{\alpha}(\Xi_n^{\alpha,\Delta})), \tag{17}$$

where $\Xi_n^{\alpha,\Delta}$ is the random J-dimensional vector defined by

$$\Xi_n^{\alpha,\Delta} := Z^{\alpha,\Delta}(t_{n-1}^{\Delta}) + b(\alpha, Z^{\alpha,\Delta}(t_{n-1}^{\Delta}))\Delta + \sigma(\alpha, Z^{\alpha,\Delta}(t_{n-1}^{\Delta}))\delta_n^{\Delta} W. \tag{18}$$

We have the following result on the convergence of the Euler approximations.

Theorem 2. *For each $p \geq 2$, $\alpha \in U$ and $t < \infty$, as $\Delta \downarrow 0$,*

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |Z^{\alpha,\Delta}(s) - Z^{\alpha}(s)|^p\right] = O\left(\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}}\right), \tag{19}$$

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |L^{\alpha,\Delta}(s) - L^{\alpha}(s)|^p\right] = O\left(\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}}\right). \tag{20}$$

The proof of Theorem 2 is given in Section 7. The proof is a straightforward adaptation of the proof of theorem 3.2(i) of Slominski (2001) which proves (19) in the case of normal reflection along the boundary. The main difference between the two results is that we allow for oblique reflection along the boundary and also prove convergence of the constraining process in (20). As a corollary of Theorem 2, we have the following result on the almost sure convergence of the Euler approximation. The corollary follows from Theorem 2 and an argument using the Borel–Cantelli lemma (see Słomínski 1994 for an example of such an argument).

Corollary 1. *For each $\alpha \in U$, $t < \infty$ and $\varepsilon > 0$, a.s.*

$$\lim_{\Delta\downarrow 0} \Delta^{\varepsilon-\frac{1}{2}} \sup_{0\leq s\leq t} |Z^{\alpha,\Delta}(s) - Z^{\alpha}(s)| = 0, \tag{21}$$

$$\lim_{\Delta\downarrow 0} \Delta^{\varepsilon-\frac{1}{2}} \sup_{0\leq s\leq t} |L^{\alpha,\Delta}(s) - L^{\alpha}(s)| = 0. \tag{22}$$

3.2. Euler scheme for the Derivative Process

We now construct an Euler scheme for the derivative process. We start with a lemma that introduces a linear projection map that can be interpreted as a linearization of π^{α}. Recall the definition of the linear subspace $\mathbb{H}_x$ for $x \in G$ given in (8).

Lemma 2. *For each $x \in G$, there is a unique map $\mathcal{L}_x^{\alpha} : \mathbb{R}^J \mapsto \mathbb{H}_x$ that satisfies $\mathcal{L}_x^{\alpha}(y) - y \in \operatorname{span}[d(\alpha, x)]$ for all $y \in \mathbb{R}^J$. Furthermore, $\mathcal{L}_x^{\alpha}$ is linear.*

Proof. By Assumption 1 and lemma 8.2 of Lipshutz and Ramanan (2018) the set $\mathcal{W}^{\alpha} := \{x \in G : \operatorname{span}(\mathbb{H}_x \cup d(\alpha, x)) \neq \mathbb{R}^J\}$ is empty. The lemma then follows from lemma 8.3 of Lipshutz and Ramanan (2018). □

Remark 7. For each $x \in G$, because $\mathcal{L}_x^{\alpha}$ is a linear map from $\mathbb{R}^J$ to $\mathbb{H}_x \subseteq \mathbb{R}^J$, we can write $\mathcal{L}_x^{\alpha}$ as a $J \times J$ matrix whose column vectors lie in $\mathbb{H}_x$. Throughout this work, we view $\mathcal{L}_x^{\alpha}$ as this $J \times J$ matrix and refer to $\mathcal{L}_x^{\alpha}$ as a "derivative projection" matrix because the matrix projects points in $\mathbb{R}^J$ to $\mathbb{H}_x$ and it is closely related to the derivative process. See the appendix for an explicit expression for the derivative projection matrices $\mathcal{L}_x^{\alpha}$, $x \in G$ in terms of the reflection matrix.

For $\Delta > 0$, recall the sequence $\{t_n^{\Delta}\}_{n\in\mathbb{N}_0}$, the random vectors $\{\delta_n^{\Delta} W\}_{n\in\mathbb{N}}$, the discrete filtration $\{\mathcal{F}_t^{\Delta}\}$, and the pair of processes $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ defined in the last section. Let $\mathcal{J}^{\alpha,\Delta} = \{\mathcal{J}^{\alpha,\Delta}(t), t \geq 0\}$ denote the piecewise constant RCLL $\{\mathcal{F}_t^{\Delta}\}$-adapted process taking values in $\mathbb{R}^{J\times M}$ defined as follows: Set

$$\mathcal{J}^{\alpha,\Delta}(0) := x_0'(\alpha) \tag{23}$$

and, for $n \in \mathbb{N}$, set

$$\mathcal{J}^{\alpha,\Delta}(t) := \mathcal{J}^{\alpha,\Delta}(t_{n-1}^{\Delta}), \qquad t \in [t_{n-1}^{\Delta}, t_n^{\Delta}), \tag{24}$$

and recursively define

$$\mathcal{J}^{\alpha,\Delta}(t_n^\Delta) := \mathcal{L}^{\alpha}_{Z^{\alpha,\Delta}(t_n^\Delta)} \mathcal{X}_n^{\alpha,\Delta}, \tag{25}$$

where $\mathcal{X}_n^{\alpha,\Delta}$ is the random element taking values in $\mathbb{R}^{J\times M}$ defined by

$$\begin{aligned}\mathcal{X}_n^{\alpha,\Delta} := \mathcal{J}^{\alpha,\Delta}(t_{n-1}^\Delta) + b_\alpha(\alpha, Z^{\alpha,\Delta}(t_{n-1}^\Delta))\Delta + b_x(\alpha, Z^{\alpha,\Delta}(t_{n-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{n-1}^\Delta)\Delta + \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(t_{n-1}^\Delta))\delta_n^\Delta W \\ + \sigma_x(\alpha, Z^{\alpha,\Delta}(t_{n-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{n-1}^\Delta)\delta_n^\Delta W + R'(\alpha)(L^{\alpha,\Delta}(t_n^\Delta) - L^{\alpha,\Delta}(t_{n-1}^\Delta)).\end{aligned} \tag{26}$$

The following establishes convergence of the Euler scheme for the derivative process.

Theorem 3. *For each $p \geq 2$, $\alpha \in U$, and $t < \infty$,*

$$\lim_{\Delta\downarrow 0} \mathbb{E}\left[\int_0^t \left\|\mathcal{J}^{\alpha,\Delta}(s) - \mathcal{J}^{\alpha}(s)\right\|^p ds + \left\|\mathcal{J}^{\alpha,\Delta}(t) - \mathcal{J}^{\alpha}(t)\right\|^p\right] = 0.$$

In addition, given any $t < \infty$, the family $\{\mathcal{J}^{\alpha,\Delta}(s)\}_{s\in[0,t],\Delta\in[0,t]}$ is uniformly integrable.

The proof of Theorem 3, which is given in Section 8.2, relies on a representation of the derivative process in terms of the deterministic derivative map evaluated at a related continuous process (see Remark 11) along with continuity properties of the derivative map, which are established in Theorem 4.

Corollary 2. *For all $t \geq 0$,*

$$\Theta_t'(\alpha) = \lim_{\Delta\downarrow 0} \mathbb{E}\left[\int_0^t \zeta_1'(Z^{\alpha,\Delta}(s))\mathcal{J}^{\alpha,\Delta}(s)ds + \zeta_2'(Z^{\alpha,\Delta}(t))\mathcal{J}^{\alpha,\Delta}(t)\right]. \tag{27}$$

Proof. The corollary follows from Theorems 1–3. □

Corollary 2 suggests a consistent estimator for $\Theta_t'(\alpha)$. In the next section, we describe an algorithm for estimating $\Theta_t'(\alpha)$ based on the Euler discretization. It is also of interest to investigate whether there is an exact sampling algorithm for the joint process $(Z^\alpha, \mathcal{J}^\alpha)$, which would avoid the bias introduced by the Euler discretization. Even in the setting of RBMs in the nonnegative orthant with reflection matrices that are so-called $\mathcal{M}$-matrices, for which an exact sampling method has been proposed in Blanchet and Murthy (2018), it is a challenging problem to exactly sample the associated derivative process because the derivative process jumps whenever the RBM reaches the boundary of the domain. We leave this as an interesting open problem for future work.

3.3. Numerical Algorithm

Fix $0 < \Delta < t < \infty$. Because Δ is typically much smaller than t, for simplicity, we can assume $N := t/\Delta$ is a positive integer. Because $Z^{\alpha,\Delta}$ and $\mathcal{J}^{\alpha,\Delta}$ are constant on intervals of the form $[t_{n-1}^\Delta, t_n^\Delta)$, $n = 1, \ldots, N$, we have

$$\int_0^t \zeta_1'(Z^{\alpha,\Delta}(s))\mathcal{J}^{\alpha,\Delta}(s)ds + \zeta_2'(Z^{\alpha,\Delta}(t))\mathcal{J}^{\alpha,\Delta}(t) = \sum_{n=0}^{N-1} \zeta_1'(Z^{\alpha,\Delta}(t_n^\Delta))\mathcal{J}^{\alpha,\Delta}(t_n^\Delta)\Delta + \zeta_2'(Z^{\alpha,\Delta}(t_N^\Delta))\mathcal{J}^{\alpha,\Delta}(t_N^\Delta). \tag{28}$$

Algorithm 1 (Our IPA Method)

Construct the projection map π^α, the map ξ^α, and the derivative projection matrices $\mathcal{L}_x^\alpha$, $x \in G$ (as described in the appendix). Set $Z^{\alpha,\Delta}(t_0^\Delta) := x_0(\alpha)$ and $\mathcal{J}^{\alpha,\Delta}(t_0^\Delta) := x_0'(\alpha)$. For $n = 1, \ldots, N$, recursively complete the following steps:

1. Generate a K-dimensional Gaussian random variable $\delta_n^\Delta W$ with mean zero and diagonal covariance matrix ΔE_K.
2. Define $(Z^{\alpha,\Delta}(t_n^\Delta), L^{\alpha,\Delta}(t_n^\Delta))$ as in (17) and (18).
3. Define $\mathcal{J}^{\alpha,\Delta}(t_n^\Delta)$ as in (25) and (26).

Set the (IPA) estimator $\widehat{\Theta_t'(\alpha)}_{\mathrm{IPA}}$ of $\Theta_t'(\alpha)$ equal to the right-hand side of (28).

In the next section, we compare our algorithm with other methods for sensitivity analysis. In Section 5, we illustrate our algorithm with an example of a one-dimensional RBM and an example of an RBM that arises in the study of a rank-based interacting diffusion model of equity markets.

4. Comparison with Existing Methods

The main existing (consistent) estimator for sensitivities of reflected diffusions is the LR method with the caveat that the LR method is applicable only when the law of the perturbed process is absolutely continuous with respect to the law of the original process. In the context of multidimensional (reflected) diffusions, this only holds for perturbations to the drift. In this case, the LR method uses a change of measure argument to recast expectations of functionals of the perturbed process as the expectation, under the law of the original process, of the functional multiplied by the likelihood ratio or the Radon–Nikodym derivative. The LR method for sensitivities of diffusions with respect to drift was introduced in Yang and Kushner (1991), and the authors describe an extension of their method to reflected diffusions (see section 9 of Yang and Kushner 1991). Here we provide a brief summary of the method in the context of reflected diffusions in convex polyhedral domains and refer the reader to Yang and Kushner (1991) for the details.

Because we can only consider perturbations to the drift in this section, in addition to the assumptions stated in Section 2.2, we also assume that the initial condition, dispersion coefficient, and directions of reflection are constant in $\alpha \in U$; that is, $x_0(\cdot) \equiv x_0$, $\sigma(\cdot,\cdot) \equiv \sigma(\cdot)$, and $R(\cdot) \equiv R$. Fix $t < \infty$ and $\alpha \in U$. For $m = 1, \ldots, M$ and $\varepsilon > 0$, let $\mu_m^{\alpha,\varepsilon} : G \mapsto \mathbb{R}^K$ be the function given by

$$\mu_m^{\alpha,\varepsilon}(x) := \sigma^T(x)a^{-1}(x)(b(\alpha + \varepsilon e_m, x) - b(\alpha, x)), \qquad x \in G.$$

A standard argument using the Lipschitz continuity of $b(\cdot,\cdot)$ and Girsanov's transformation (see, e.g., chapter IV.38 of Rogers and Williams 2000) shows that there exists a family of probability measures $\{\mathbb{P}_m^{\alpha,\varepsilon}\}_{\varepsilon>0}$ on the measurable space $(\Omega, \mathcal{F}_t)$ that are mutually absolutely continuous with respect to the underlying measure $\mathbb{P}$ with Radon–Nikodym derivatives

$$\frac{d\mathbb{P}_m^{\alpha,\varepsilon}}{d\mathbb{P}} := \exp\left\{\int_0^t \langle \mu_m^{\alpha,\varepsilon}(Z^\alpha(s)), dW(s)\rangle - \frac{1}{2}\int_0^t |\mu_m^{\alpha,\varepsilon}(Z^\alpha(s))|^2 ds\right\}, \qquad \varepsilon > 0, \tag{29}$$

and such that $\{Z^\alpha(s), 0 \le s \le t\}$ under $\mathbb{P}_m^{\alpha,\varepsilon}$ is equal in distribution to $\{Z^{\alpha+\varepsilon e_m}(s), 0 \le s \le t\}$ under $\mathbb{P}$. Let D^α denote the M-dimensional random variable whose mth component, denoted $D^{\alpha,m}$ for $m = 1, \ldots, M$, is defined by

$$D^{\alpha,m} := \lim_{\varepsilon\to 0}\frac{(d\mathbb{P}_m^{\alpha,\varepsilon}/d\mathbb{P}) - 1}{\varepsilon} = \int_0^t \langle \sigma^T(Z^\alpha(s))a^{-1}(Z^\alpha(s))b_{\alpha^m}(\alpha, Z^\alpha(s)), dW(s)\rangle,$$

where $b_{\alpha^m}(\alpha, x)$ denotes the partial derivative of $b(\cdot, x)$ at α with respect to the mth component of α. According to theorem 3.1 of Yang and Kushner (1991),

$$\begin{aligned}\partial_m\Theta_t(\alpha) &= \lim_{\varepsilon\to 0}\mathbb{E}\left[\left(\int_0^t \zeta_1(Z^\alpha(s))ds + \zeta_2(Z^\alpha(t))\right)\frac{(d\mathbb{P}_m^{\alpha,\varepsilon}/d\mathbb{P}) - 1}{\varepsilon}\right]\\ &= \mathbb{E}\left[\left(\int_0^t \zeta_1(Z^\alpha(s))ds + \zeta_2(Z^\alpha(t))\right)D^{\alpha,m}\right].\end{aligned}$$

As described in (15)–(17), we have an Euler approximation $Z^{\alpha,\Delta}$ for Z^α, which yields the approximation $D^{\alpha,\Delta,m}$ for $D^{\alpha,m}$ for $m = 1, \ldots, M$, given by

$$D^{\alpha,\Delta,m} = \sum_{n=0}^{N-1}\langle \sigma^T(Z^{\alpha,\Delta}(t_n^\Delta))a^{-1}(Z^{\alpha,\Delta}(t_n^\Delta))b_{\alpha^m}(\alpha, Z^{\alpha,\Delta}(t_n^\Delta)), \delta_n^\Delta W\rangle. \tag{30}$$

According to theorem 4.1 of Yang and Kushner (1991),

$$\Theta_t'(\alpha) = \lim_{\Delta\to 0}\mathbb{E}\left[\left(\int_0^t \zeta_1(Z^{\alpha,\Delta}(s))ds + \zeta_2(Z^{\alpha,\Delta}(t))\right)D^{\alpha,\Delta}\right]. \tag{31}$$

We now propose a numerical algorithm for estimating $\Theta_t'(\alpha)$ based on (30) and (31). Fix $\Delta \in (0, t)$ so that $N := t/\Delta$ is a positive integer. Then we have

$$\left(\int_0^t \zeta_1(Z^{\alpha,\Delta}(s))ds + \zeta_2(Z^{\alpha,\Delta}(t))\right)D^{\alpha,\Delta} = \left(\sum_{n=0}^{N-1}\zeta_1(Z^{\alpha,\Delta}(t_n^\Delta))\Delta + \zeta_2(Z^{\alpha,\Delta}(t_N^\Delta))\right)D^{\alpha,\Delta}. \tag{32}$$

Algorithm 2 (LR Method)
Construct the projection map π^α (as described in the appendix). Set $Z^{\alpha,\Delta}(t_0^\Delta) = x_0$. For $n = 1, \ldots, N$, recursively complete the following two steps:

1. Generate a K-dimensional Gaussian random variable $\delta_n^\Delta W$ with mean zero and diagonal covariance matrix ΔE_K.
2. Define $Z^{\alpha,\Delta}(t_n^\Delta)$ as in (17) and (18).

Define $D^{\alpha,\Delta}$ as in (30) and set the (LR) estimator $\widehat{\Theta_t'(\alpha)}_{\mathrm{LR}}$ of $\Theta_t'(\alpha)$ equal to the right-hand side of (32).

We briefly mention some of the advantages and drawbacks of the LR method. When applicable, the main advantage of the LR method is that it applies to a broad class of functionals of reflected diffusions, requiring only that ζ_1 and ζ_2 be measurable (and integrable) without imposing additional smoothness conditions that are required for our IPA method (Algorithm 1). In addition, because there are exact sampling methods for reflected diffusions (see Blanchet and Murthy 2018), they can potentially be adopted to obtain an unbiased estimator of $\Theta_t'(\alpha)$. On the other hand, as mentioned, the main drawback of the LR method is that it is generally only applicable to perturbations of the drift. In section 8 of Yang and Kushner (1991) the authors show that, in the one-dimensional setting, the approach can be adapted to allow for perturbations of the diffusion coefficient, but those methods do not extend to higher dimensions. This is quite restrictive as it is often of interest to compute sensitivities with respect to parameters other than the drift (e.g., in diffusion approximations of queueing networks, which we plan to study in future work). Furthermore, even when considering perturbations to the drift, numerical evidence suggests that the variance of the LR estimator is substantially larger than the variance our estimator (see Section 3). This supports the observation made in chapter VII.5 of Asmussen and Glynn (2007) that IPA estimators generally outperform LR estimators when both methods are applicable.

Remark 8. It is worthwhile to compare our IPA algorithm with the class of FD methods, which can be used to estimate sensitivities. These are based on approximating the derivative by a finite difference and include forward, backward, and central difference approximations (see, e.g., chapter VII.1 of Asmussen and Glynn 2007). For example, the forward difference estimator associated with fixed $\varepsilon, \Delta > 0$ (and assuming once again that $N := t/\Delta$ is a positive integer) for $\partial_m \Theta_t(\alpha)$ for $m = 1, \ldots, M$ is given by

$$\widehat{\partial_m \Theta_t(\alpha)}_{\mathrm{FD}} = \sum_{n=0}^{N-1} \frac{\zeta_1(Z^{\alpha+\varepsilon e_m,\Delta}(t_n^\Delta)) - \zeta_1(Z^{\alpha,\Delta}(t_n^\Delta))}{\varepsilon}\Delta + \frac{\zeta_2(Z^{\alpha+\varepsilon e_m,\Delta}(t_N^\Delta)) - \zeta_2(Z^{\alpha,\Delta}(t_N^\Delta))}{\varepsilon}, \tag{33}$$

with t_n^Δ defined as in (12). The appeal of an FD method is that it is simple to implement and can be used to approximate sensitivities with respect to all of the parameters. The main drawback is that it introduces an extra source of bias from the derivative approximation. Our IPA method can be viewed as a limiting version of the FD method that eliminates the bias from the derivative approximation without any increase in the computational complexity of implementation. Because there is no computational advantage to using the FD method and, in general, there do not appear to be any other advantages to using an FD method over an IPA method when both are applicable (see, e.g., the first paragraph at the the top of p. 219 of Asmussen and Glynn 2007), we do not provide a numerical comparison of our method and an FD method.

5. Examples

We present two examples to illustrate the efficacy of our method (Algorithm 1): a one-dimensional RBM in which analytic formulas for the transition density are available and a multidimensional RBM that arises in the study of rank-based diffusion models for equity markets. We chose these two examples because they are useful for illustrating the efficacy of our method. We emphasize that our algorithm can be used to estimated sensitivities for broad classes of reflected diffusions with oblique reflection, such as RBMs with oblique reflection and reflected Ornstein–Uhlenbeck processes that arise as heavy traffic limits of queueing networks (see, e.g., Reiman 1984, Chen and Mandelbaum 1991, Peterson 1991, Mandelbaum and Pats 1998, and Ward and Glynn 2003). We used the program R for all computations, which were performed on an IBM machine with a 3.40 GHz Intel Core i7 processor.

5.1. One-Dimensional RBM

Let $J = 1$, $U := (0,\infty) \times \mathbb{R} \times (0,\infty)$, and W be a one-dimensional $\{\mathcal{F}_t\}$-Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. For each $\alpha = (x, b, \sigma) \in U$, let $Z^{(x,b,\sigma)}$ be a one-dimensional RBM in $\mathbb{R}_+$ with initial condition x, constant drift b, and variance σ^2 and driving Brownian motion W. It is well known (see, e.g., chapter X.8 of Asmussen and Glynn 2007) that $Z^{(x,b,\sigma)}$ satisfies

$$Z^{(x,b,\sigma)}(t) = x + bt + \sigma W(t) + Y^{(x,b,\sigma)}(t), \qquad t \geq 0,$$

where $Y^{(x,b,\sigma)}$ is given by

$$Y^{(x,b,\sigma)}(t) = \sup_{0\le s\le t}(-x - bs - \sigma W(s)) \vee 0, \qquad t \ge 0.$$

For $t \ge 0$, define $\Theta_t : U \mapsto \mathbb{R}_+$ by

$$\Theta_t(x,b,\sigma) := \mathbb{E}\Big[Z^{(x,b,\sigma)}(t)\Big], \qquad (x,b,\sigma) \in U. \tag{34}$$

We use our algorithm (Algorithm 1) to estimate the first partial derivatives $\partial_x\Theta_1(1,-1,1)$, $\partial_b\Theta_1(1,-1,1)$, and $\partial_\sigma\Theta_1(1,-1,1)$. In Table 1, we compare the estimates obtained using our IPA method with the analytic values obtained using the formula for the cumulative distribution function of a one-dimensional RBM given in equation 3.63 of Harrison (2013) and, in the case of perturbations to the drift, with the estimates obtained using the LR method (Algorithm 2).

In Figure 1, we plot confidence intervals (CIs) for the relative biases of our IPA estimator and the LR estimator (when applicable) of the partial derivatives of $\Theta_1(\cdot,\cdot,\cdot)$ for different time steps Δ. As anticipated, the magnitudes of the relative biases decrease with Δ. (The "relative bias" of an estimator is equal to the difference between the expected value of the estimator and the analytic value divided by the magnitude of the analytic value.)

In Table 2, we compare CIs for the relative biases of the LR estimator and our IPA estimator of $\partial_b\Theta_t(1,b,1)$ for different values of b, t, and Δ. As seen in Table 2, there is no significant difference between the estimates of the relative biases. In theorem 4.1 of Yang and Kushner (1991) the authors state that (under our stated assumptions and also assuming that only the drift b depends on α)

$$\mathbb{E}\Big[\widehat{\Theta_t'(\alpha)}_{\mathrm{LR}}\Big] = \frac{d}{d\alpha}\mathbb{E}\bigg[\int_0^t \zeta_1(Z^{\alpha,\Delta}(s))ds + \zeta_2(Z^{\alpha,\Delta}(t))\bigg].$$

(Specifically, they prove the preceding holds in the case that Z^α is a diffusion and describe an extension to the case of a reflected diffusion.) We conjecture that

$$\mathbb{E}\Big[\widehat{\Theta_t'(\alpha)}_{\mathrm{IPA}}\Big] = \frac{d}{d\alpha}\mathbb{E}\bigg[\int_0^t \zeta_1(Z^{\alpha,\Delta}(s))ds + \zeta_2(Z^{\alpha,\Delta}(t))\bigg]$$

also holds, and so our IPA estimator and the LR estimator have the same bias for each time step $\Delta > 0$. The proof of this conjecture would require significant additional work that entails the proof that $\mathcal{J}^{\alpha,\Delta}$ is the appropriate linearization of $Z^{\alpha,\Delta}$ with respect to α. Because this is beyond the scope of this paper, we leave our conjecture as an interesting problem for future work.

In Table 3, we compare the empirical standard deviations of the LR estimator and our IPA estimator of $\partial_b\Theta_t(1,b,1)$ for fixed $\Delta > 0$ and different values of b and t. Our method appears much more efficient in all cases although the degree of variance reduction depends strongly on b and t.

5.2. Rank-Based Interacting Diffusion Model of Equity Markets

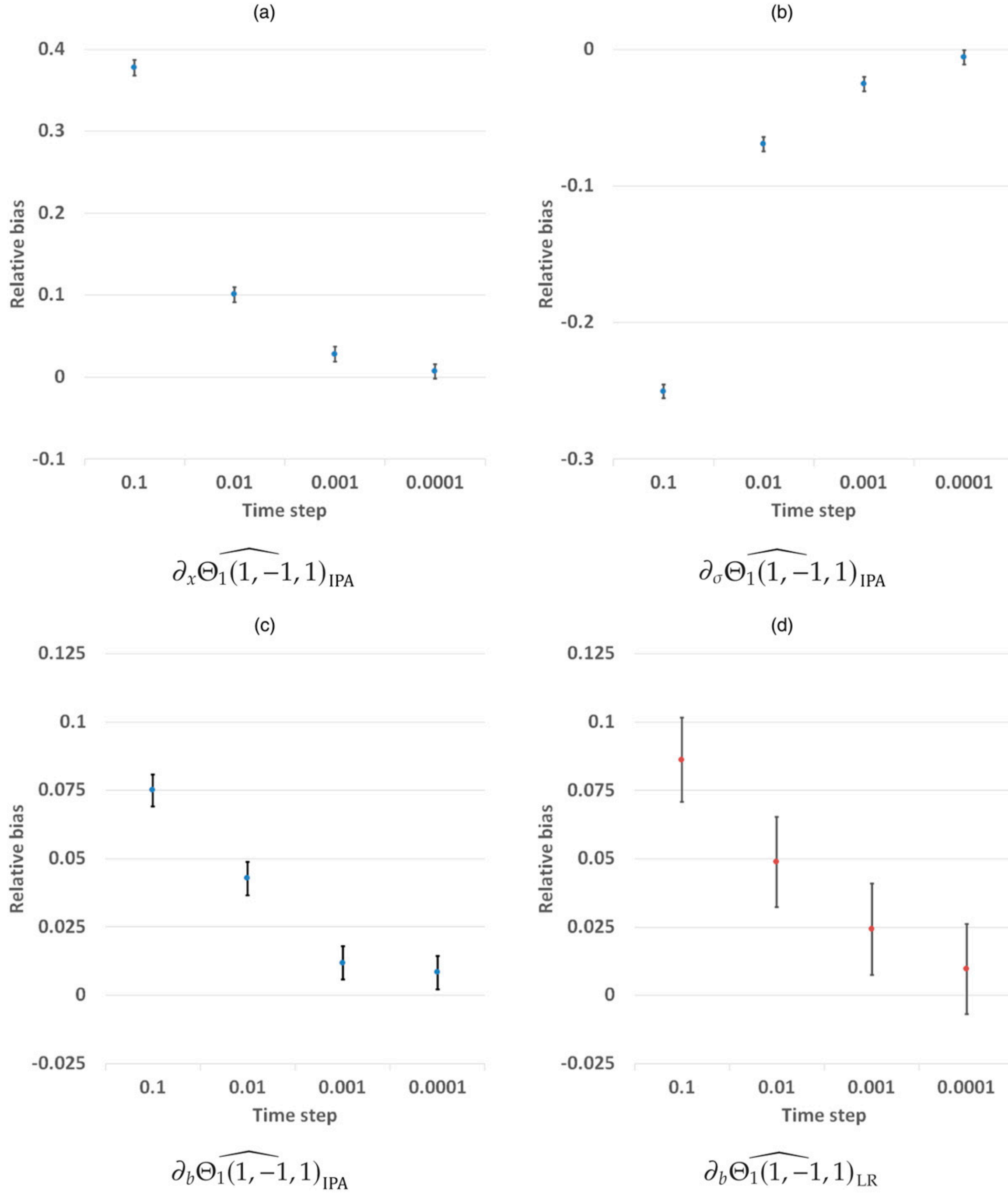
We consider a rank-based interacting diffusion model, referred to as the Atlas model, which was first introduced by example 5.3.3 of Fernholz (2002) to study equity markets and subsequently generalized by

Table 1. Analytic Values, the LR Estimates, and Our IPA Estimates for Sensitivities of an RBM

Method	Δ	$\partial_x\Theta_1(1,-1,1)$	$\partial_b\Theta_1(1,-1,1)$	$\partial_\sigma\Theta_1(1,-1,1)$
Analytic	N/A	0.3319	0.4351	0.6681
LR	10^{-1}	N/A	.4726 ± .0067 (2.4 seconds)	N/A
IPA	10^{-1}	.4572 ± .0031 (2.2 seconds)	.4677 ± .0025 (2.2 seconds)	.5008 ± .0034 (2.2 seconds)
LR	10^{-2}	N/A	.4563 ± .0072 (27.9 seconds)	N/A
IPA	10^{-2}	.3653 ± .0030 (23.9 seconds)	.4537 ± .0027 (24.3 seconds)	.6217 ± .0035 (24.3 seconds)
LR	10^{-3}	N/A	.4456 ± .0073 (282.1 seconds)	N/A
IPA	10^{-3}	.3412 ± .0029 (233.4 seconds)	.4402 ± .0026 (238.2 seconds)	.6510 ± .0034 (282.1 seconds)
LR	10^{-4}	N/A	.4393 ± .0072 (2,700.3 seconds)	N/A
IPA	10^{-4}	.3342 ± .0029 (2,359.8 seconds)	.4387 ± .0026 (2,358.7 seconds)	.6641 ± .0035 (2,308.5 seconds)

Notes. Analytic values (approximated up to four decimal places) of the first partial derivatives of $\Theta_1(\cdot,\cdot,\cdot)$ defined as in (34) and evaluated at $(1,-1,1)$, as well as 95% confidence intervals for the LR estimates (when applicable) and our IPA estimates of the first partial derivatives of $\Theta_1(\cdot,\cdot,\cdot)$. Each estimate was obtained using 10^5 trials. The CPU time for each estimate is indicated in parenthesis.

Figure 1. Relative Biases of Our IPA Estimators and of the LR Estimators for an RBM



Notes. 95% CI for the relative biases of IPA estimators and (when applicable) LP estimators for the first partial derivatives of $\Theta_1(\cdot,\cdot,\cdot)$ defined in (34) and evaluated at $(1,-1,1)$ using different time steps Δ. Each estimate was obtained using 10^5 trials.

Banner et al. (2005) and Ichiba et al. (2011). In the usual model of an equity market, the growth and volatility of each stock are fixed and depend on the identity of the stock, whereas in the class of Atlas models, the growth and volatility of each stock are allowed to depend on its relative rank within the market and, thus, are time-varying (and discontinuous). The hybrid Atlas models considered in Ichiba et al. (2011) allow for the drift and diffusion coefficients to depend on both the identity of the stock and its relative rank within the market. Here we restrict our consideration to the class of Atlas models originally introduced in Fernholz (2002) and studied in Banner et al. (2005), whose coefficients depend only on the relative ranks of the stocks within the market because their associated "ranked processes" are RBMs.

Table 2. Relative Biases of Our IPA Estimators and of the LR Estimators for an RBM

Method	b	Δ	$t = 2$	$t = 6$	$t = 10$	CPU time
LR	−1	0.1	.0602 ± .0020	−.0306 ± .0026	−.0411 ± .0030	2,601.9 seconds
IPA	−1	0.1	.0603 ± .0009	−.0303 ± .0011	−.0436 ± .0011	2,329.1 seconds
LR	−1	0.01	.0253 ± .0021	−.0019 ± .0028	−.0064 ± .0033	24,248.9 seconds
IPA	−1	0.01	.0246 ± .0009	−.0018 ± .0011	−.0057 ± .0011	24,283.2 seconds
LR	0	0.1	.0586 ± .0014	.0363 ± .0015	.0265 ± .0015	2,492.8 seconds
IPA	0	0.1	.0589 ± .0004	.0350 ± .0004	.0247 ± .0004	2,539.8 seconds
LR	0	0.01	.0224 ± .0015	.0127 ± .0016	.0091 ± .0016	24,520.9 seconds
IPA	0	0.01	.0204 ± .0004	.0119 ± .0004	.0081 ± .0004	23,958.8 seconds
LR	1	0.1	.0179 ± .0017	.0084 ± .0021	.0052 ± .0024	2,609.8 seconds
IPA	1	0.1	.0173 ± .0001	.0076 ± .0001	.0047 ± .0001	2,475.1 seconds
LR	1	0.01	.0056 ± .0017	.0033 ± .0021	.0032 ± .0024	24,546.3 seconds
IPA	1	0.01	.0062 ± .0001	.0026 ± .0001	.0016 ± .0001	23,897.3 seconds

Notes. 95% confidence intervals for the relative biases of the LR estimator and our IPA estimator of the first derivative of $\Theta_t(1,\cdot,1)$ defined in (34) evaluated at different values of b and t, and using different time steps Δ. For each method, drift b and time step Δ, the CPU time to estimate the first derivative of $\Theta_{10}(1,\cdot,1)$ is listed. Each estimate was obtained using 10^7 trials.

In the simplest version of the Atlas model (with $J \geq 2$ stocks), the stocks have equal volatility, the smallest stock (referred to as the *Atlas stock*) has positive growth rate, and all other stocks have zero growth. To be precise, let $g > 0$, $\sigma > 0$, and W be a J-dimensional Brownian motion. Let $S = \{S(t), t \geq 0\}$ denote the J-dimensional process whose jth component at time $t \geq 0$, denoted $S^j(t)$, is equal to the value of the jth stock at time t. Then S evolves according to the SDE (with discontinuous drift)

$$d(\log S^j(t)) = \gamma^j(t)dt + \sigma dW^j(t), \qquad t \geq 0, \qquad j = 1, \ldots, J, \tag{35}$$

where $\gamma^j(t)$ is equal to Jg if $S^j(t) < S^k(t)$ for all $k \neq j$ and equal to zero otherwise for $j = 1, \ldots, J$. Strong existence and pathwise uniqueness of solutions to (35) follows from Ichiba et al. (2013). Let $Z = \{Z(t), t \geq 0\}$ be the *rank-ordered process*, where $Z(t)$ is obtained by permuting the components of $\log S(t)$ (the logarithm is applied componentwise) so that they are in descending order for $t \geq 0$. Then according to Banner et al. (2005) (see the paragraph following the proof of proposition 2.3), Z is a J-dimensional RBM in the Weyl chamber

$$G := \{x \in \mathbb{R}^J : x^1 \geq x^2 \geq \cdots \geq x^J\}$$

with drift vector $(0, \ldots, 0, Jg)^T$, diagonal covariance matrix $\sigma^2 E_J$, and normal reflection vector $d_i = n_i = (e_i - e_{i+1})/\sqrt{2}$ along face $F_i := \{x \in \partial G : x^i = x^{i+1}\}$ for $i = 1, \ldots, J-1$.

In many applications in mathematical finance, it is of interest to compute sensitivities of functionals of the stocks in a market with respect to the underlying parameters of the market. For example, sensitivities of derivative prices with respect to the underlying market parameters, commonly referred to as the "Greeks," play an important role in risk management and hedging strategies (see, e.g., chapter 7 of Glasserman 2003). Here, we study sensitivities related to the Atlas model, which have not yet received much attention. A performance measure of interest in this context is the so-called diversity of the equity market, which is expressed in terms of the J-dimensional process $\mu = \{\mu(t), t \geq 0\}$ of *relative market capitalizations* defined by $\mu(t) := C(S(t))$ for $t \geq 0$, where $C : (0,\infty)^J \mapsto \{x \in (0,1)^J : x^1 + \cdots + x^J = 1\}$ is defined by

$$C^j(x) := \frac{x^j}{x^1 + \cdots + x^J}, \qquad x \in (0,\infty)^J, \qquad j = 1, \ldots, J. \tag{36}$$

Table 3. Relative Performance of the LR Estimator and Our IPA Estimator

b	$t = 2$	$t = 4$	$t = 6$	$t = 8$	$t = 10$
−3	2.61	3.16	3.71	4.18	4.59
−2	2.10	2.88	3.21	3.63	3.85
−1	2.17	2.26	2.44	2.61	2.82
0	3.96	3.66	3.61	3.58	3.59
1	3.91	4.75	5.69	6.66	7.59
2	1.01×10^2	2.41×10^2	4.20×10^2	6.30×10^2	8.53×10^2
3	7.18×10^2	1.85×10^3	3.30×10^3	4.99×10^3	6.80×10^3

Notes. Ratios between the empirical standard deviations of the LR estimate and our IPA estimate of the first derivative of $\Theta_t(1,\cdot,1)$ defined in (34), evaluated at different values of t and b. Each estimate was obtained using 10^5 trials and time step $\Delta = 0.1$.

Fix $p \in (0,1)$ and define $f : \mathbb{R}^J_+ \to \mathbb{R}_+$ by

$$f(x) := (x^1)^p + \cdots + (x^J)^p, \qquad x \in \mathbb{R}^J_+. \tag{37}$$

Then f is the pth power of the so-called *diversity function*, and $f(\mu(t))$ is a measure of the diversity of the market (see example 3.4.4 of Fernholz 2002). Because $f(x)$ is invariant under permutations of the components of x and Z is obtained by permuting the components of $\log S$, it follows from (36) and (37) that $f(\mu(t)) = \zeta(Z(t))$, where $\zeta : G \mapsto [J^{p-1}, 1)$ is the continuously differentiable function defined by

$$\zeta(x) := \frac{e^{px^1} + \cdots + e^{px^J}}{(e^{x^1} + \cdots + e^{x^J})^p}, \qquad x \in G.$$

A straightforward computation shows that the gradient of ζ is bounded and continuous. Here, we first use an FD method based on the SDE representation (35) and then our IPA method for the reflected diffusion representation to estimate the sensitivity of $\mathbb{E}[f(\mu(t))]$ to an underlying parameter of the market.

Fix $\sigma > 0$ and set $U := (0, \infty)$. For each $\alpha \in U$, set the growth rate to be $g(\alpha) := \sigma^2/2\alpha$. Here, the form of $g(\alpha)$ is chosen to be consistent with the growth rate of interest in section 6 of Banner et al. (2005). Let $S^\alpha = \{S^\alpha(t), t \geq 0\}$ be the J-dimensional process satisfying (35) with $g = g(\alpha)$, let $\mu^\alpha = \{\mu^\alpha(t), t \geq 0\}$ be the process of relative market capitalizations defined by $\mu^\alpha(t) := C(S^\alpha(t))$ for $t \geq 0$, and define

$$\Theta_t(\alpha) := \mathbb{E}\big[f(\mu^\alpha(t))\big], \qquad t \geq 0. \tag{38}$$

Let $t \geq 0$ and $\alpha \in U$. Assuming that $\Theta_t(\cdot)$ is differentiable, it follows that, as $\varepsilon \to 0$,

$$\Theta_t'(\alpha) = \frac{\mathbb{E}\big[f(\mu^{\alpha+\varepsilon}(t))\big] - \mathbb{E}\big[f(\mu^\alpha(t))\big]}{\varepsilon} + o(1).$$

Yan (2002) proposed an Euler scheme for approximating solutions to SDEs with discontinuous drift and proved convergence properties of the Euler scheme. Using the algorithm described in section 4 of Yan (2002) one can obtain estimators $\widehat{S^{\alpha+\varepsilon}}$ and $\widehat{S^\alpha}$ (using "common random numbers," see chapter VII.1 of Asmussen and Glynn 2007) for $S^{\alpha+\varepsilon}$ and S^α, respectively, and use the (forward-difference) approximation

$$\Theta_t'(\alpha) \approx \frac{f(C(\widehat{S^{\alpha+\varepsilon}}(t))) - f(C(\widehat{S^\alpha}(t)))}{\varepsilon}.$$

Alternatively, letting Z^α denote the associated rank-ordered process so that $\Theta_t(\alpha) = \mathbb{E}[\zeta(Z^\alpha(t))]$, we can use our IPA method (applied to the reflected diffusion Z^α) to estimate $\Theta_t'(\alpha)$. We note that although Z^α is an RBM in the Weyl chamber with *normal* reflection, the transformed process $\widetilde{Z}^\alpha := R^T Z^\alpha$ is a $(J-1)$-dimensional RBM in the orthant with *oblique* directions of reflection $\widetilde{d}_i = [R^T R]^i$ along the face $\widetilde{F}_i = \{x \in \mathbb{R}^{J-1}_+, x^i = 0\}$ for $i = 1, \ldots, J-1$. Thus, we believe the numerical results presented in this example are also representative of the performance of the algorithm for a large class of oblique RBMs in the orthant.

Let $\sigma^2 = 3 \times 10^{-4}$, $Z^\alpha(0) = 0$, and $p = 1/2$. In Table 4, we list the ratio of the empirical standard deviations of the FD estimator (described in Yan 2002 for the SDE description of the model) and our IPA estimator (for the

Table 4. Relative Performance of Our IPA Estimator, the LR estimator, and the FD Estimator

Method	ε	95% CI	Ratio of standard deviations	CPU time
IPA	N/A	$-.01037 \pm .00005$	1.0	1,165.0 seconds
LR	N/A	$-.00624 \pm .00865$	161.3	533.7 seconds
FD	10^{-2}	$-.01053 \pm .00006$	1.1	763.6 seconds
FD	10^{-3}	$-.01063 \pm .00007$	1.3	852.6 seconds
FD	10^{-4}	$-.01061 \pm .00014$	2.6	858.9 seconds
FD	10^{-5}	$-.01052 \pm .00038$	7.2	809.8 seconds
FD	10^{-6}	$-.01065 \pm .00070$	13.0	786.7 seconds

Notes. 95% CI for our IPA estimator, the LR estimator, and the FD estimator (for different $\varepsilon > 0$) of the first derivative of $\mathbf{\Theta}_{1000}(\cdot)$, defined in (38) with $J = 3$, and evaluated at $\alpha = 1$. For each method, the ratio between the standard deviation of the estimator and the standard deviation of our IPA estimator is listed as well as the CPU time required to compute the estimate. Each estimate was obtained using 10^5 trials and time step $\Delta = 1$.

RBM description of the model) for $\Theta'_{1000}(1)$ when $J = 3$. We also include the ratio of the empirical standard deviations of the LR estimator and our IPA estimator (both for the RBM description of the model) for $\Theta'_{1000}(1)$ when $J = 3$. As seen in the table, we obtain significant variance reduction by using our IPA method over the LR method and over the FD method for small $\varepsilon > 0$.

To illustrate the performance of our method in higher dimensions, in Table 5, we compare the empirical standard deviations and CPU times of our IPA estimator, the LR estimator, and the FD estimator (with $\varepsilon = .01$) for different dimensions J. As expected, the CPU time required to compute each estimate increases with J. The empirical standard deviation of the LR estimate appears to be diverging as J increases. This can be explained by the fact that the LR estimator depends on the derivative of the drift b with respect to α as seen in (30) and (31) and the fact that the magnitude of the drift grows linearly with the dimension J—in particular, $b(\alpha) = (0, \ldots, 0, Jg(\alpha))^T$.

6. The Skorokhod Problem and the Derivative Problem

In this section, we recall the deterministic SP, extended Skorokhod problem (ESP), and derivative problem. We state some useful properties and also state a new continuity result for the derivative problem (see Theorem 4) that is needed for the proof of Theorem 3 and whose proof is deferred to Section 9. Throughout this section, we fix $\alpha \in U$. Recall that Assumptions 1–3 hold.

6.1. The Skorokhod Problem and Extended Skorokhod Problem

Recall that $R(\alpha) := \big(d_1(\alpha) \cdots d_N(\alpha)\big)$ denotes the reflection matrix associated with α.

Definition 3. Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, a pair $(h, \ell) \in \mathbb{D}(G) \times \mathbb{D}(\mathbb{R}^N_+)$ is a solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f if

$$h(t) = f(t) + R(\alpha)\ell(t), \qquad t \geq 0,$$

where $\ell(0) = 0$, and for each $i \in \mathcal{I}$, the ith component of ℓ, denoted ℓ^i, is nondecreasing and can only increase when h lies in face F_i; that is,

$$\int_0^\infty 1_{\{h(s) \notin F_i\}} d\ell^i(s) = 0.$$

If there exists a unique solution (h, ℓ) to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f, we write $h = \Gamma^\alpha(f)$ and refer to Γ^α as the associated SM.

In the following, we state the ESP, which is an extension of the SP that allows for a constraining term with unbounded variation. We introduce the ESP formulation because it is more convenient to work with in subsequent proofs.

Definition 4. Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, a pair $(h, g) \in \mathbb{D}(\mathbb{R}^J) \times \mathbb{D}(\mathbb{R}^J)$ is a solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f if $h(0) = f(0)$ and the following conditions hold for all $t \geq 0$:

1. $h(t) = f(t) + g(t)$;
2. $h(t) \in G$;
3. for all $s \in [0, t)$,
$$g(t) - g(s) \in \text{cone}\left[\cup_{u \in (s,t]} d(\alpha, h(u))\right]; \text{ and}$$
4. $g(t) - g(t-) \in \text{cone}\,[d(\alpha, h(t))]$, where $g(0-) := 0$.

If there exists a unique solution (h, g) to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f, we write $h = \bar{\Gamma}^\alpha(f)$ and refer to $\bar{\Gamma}^\alpha$ as the associated ESM.

Table 5. Performance of Our IPA Estimator, the LR Estimator, and the FD Estimator as the Number of Particles Increases

J	IPA	LR	FD ($\varepsilon = .01$)
2	7.257×10^{-3} (104.4 seconds)	0.7677 (45.5 seconds)	7.774×10^{-3} (88.8 seconds)
3	8.713×10^{-3} (112.4 seconds)	1.407 (55.4 seconds)	9.444×10^{-3} (91.5 seconds)
4	9.053×10^{-3} (130.6 seconds)	2.146 (60.0 seconds)	1.043×10^{-2} (97.0 seconds)
5	9.821×10^{-3} (143.7 seconds)	3.043 (65.9 seconds)	1.101×10^{-2} (104.4 seconds)
6	1.105×10^{-2} (171.7 seconds)	3.923 (74.5 seconds)	1.158×10^{-2} (110.1 seconds)

Notes. Empirical standard deviations for our IPA estimator, the LR estimator, and the FD estimator (with $\varepsilon = .01$) of the first derivative of $\mathbf{\Theta}_{1000}(\cdot)$, defined in (38), evaluated at $\alpha = 1$ for increasing dimension J. The CPU time required to compute each estimate is listed in parenthesis. Each estimate was obtained using 10^4 trials and time step $\Delta = 1$.

Remark 9. Given a reflected diffusion Z^α with L^α satisfying the conditions in Definition 1, define the $\{\mathcal{F}_t\}$-adapted J-dimensional continuous process $X^\alpha = \{X^\alpha(t), t \geq 0\}$ by

$$X^\alpha(t) := x_0(\alpha) + \int_0^t b(\alpha, Z^\alpha(s))ds + \int_0^t \sigma(\alpha, Z^\alpha(s))dW(s), \qquad t \geq 0. \tag{39}$$

Then, by the properties of (Z^α, L^α) stated in Definition 1, (Z^α, L^α) is a solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for X^α. Define the constraining process Y^α by $Y^\alpha := R(\alpha)L^\alpha$. Then, by the properties of Y^α stated in Remark 2, (Z^α, Y^α) is a solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for X^α.

The following is a time-shift property to the ESP. A similar property holds for the SP; however, the latter is not needed in this work.

Lemma 3 (Ramanan 2006, Lemma 2.3). *Suppose (h, g) is a solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for $f \in \mathbb{D}_G(\mathbb{R}^J)$. Let $s \geq 0$ and define $h^s \in \mathbb{D}(G)$, $g^s \in \mathbb{D}(\mathbb{R}^J)$, and $f^s \in \mathbb{D}_G(\mathbb{R}^J)$ by*

$$h^s(\cdot) := h(s + \cdot), \qquad g^s(\cdot) := g(s + \cdot) - g(s), \qquad f^s(\cdot) := h(s) + f(s + \cdot) - f(s). \tag{40}$$

Then (h^s, g^s) is a solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f^s.

We close this section with the following proposition that ensures, under our standing assumptions, that the SP and ESP are well defined and the associated SM and ESM are Lipschitz continuous.

Proposition 1. *Let $\alpha \in U$.*

i. *There exists $\kappa_\Gamma(\alpha) < \infty$ such that if, for $k = 1, 2$, $f_k \in \mathbb{D}_G(\mathbb{R}^J)$ and (h_k, ℓ_k) is the solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f_k, then for all $t \geq 0$,*

$$\sup_{0\leq s\leq t} |h_1(s) - h_2(s)| + \sup_{0\leq s\leq t} |\ell_1(s) - \ell_2(s)| \leq \kappa_\Gamma(\alpha) \sup_{0\leq s\leq t} |f_1(s) - f_2(s)|. \tag{41}$$

ii. *There exists $\kappa_{\bar{\Gamma}}(\alpha) < \infty$ such that if, for $k = 1, 2$, $f_k \in \mathbb{D}_G(\mathbb{R}^J)$ and (h_k, g_k) is the solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f_k, then for all $t \geq 0$,*

$$\sup_{0\leq s\leq t} |h_1(s) - h_2(s)| + \sup_{0\leq s\leq t} |g_1(s) - g_2(s)| \leq \kappa_{\bar{\Gamma}}(\alpha) \sup_{0\leq s\leq t} |f_1(s) - f_2(s)|. \tag{42}$$

iii. *Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, there exists a unique solution (h, ℓ) to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f.*
iv. *Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, there exists a unique solution (h, g) to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f.*

Proof. The Lipschitz continuity of the ESM (42) stated in part (ii) follows from Assumption 2 and theorem 3.3 of Ramanan (2006). Part (iv) follows from part (ii) and lemma 2.6 of Ramanan (2006). Parts (i) and (iii) then follow from parts (ii) and (iv), Assumption 1, and lemma 2.1 of Lipshutz and Ramanan (2019a). □

6.2. The Derivative Problem

The derivative problem was first introduced in Lipshutz and Ramanan (2018) as an axiomatic framework for studying directional derivatives of the ESM. The formulation of the derivative problem can be thought of as a linearization to the ESP (note the similarities between Definitions 4 and 5). Recall the definition of $\mathbb{H}_x$ given in (8).

Definition 5. Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, suppose (h, g) is a solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f. Let $\psi \in \mathbb{D}(\mathbb{R}^J)$. Then $(\phi, \eta) \in \mathbb{D}(\mathbb{R}^J) \times \mathbb{D}(\mathbb{R}^J)$ is a solution to the derivative problem along h for ψ if the following conditions hold for all $t \geq 0$:

1. $\phi(t) = \psi(t) + \eta(t)$;
2. $\phi(t) \in \mathbb{H}_{h(t)}$;
3. for all $s \in [0, t)$,

$$\eta(t) - \eta(s) \in \operatorname{span}\left[\cup_{u\in(s,t]} d(\alpha, h(u))\right]; \text{ and}$$

4. $\eta(t) - \eta(t-) \in \operatorname{span}[d(\alpha, h(t))]$, where $\eta(0-) := 0$.

If there exists a unique solution (ϕ, η) to the derivative problem along h for ψ, we write $\phi = \Lambda_h^\alpha(\psi)$ and refer to Λ_h^α as the derivative map associated with h.

Remark 10. Definition 5 is slightly different from definition 3.4 of Lipshutz and Ramanan (2018), where h was assumed to be continuous. Here we allow for h to lie in $\mathbb{D}(\mathbb{R}^J)$ and impose condition 4, which follows from condition 3 when h is continuous. When h is not continuous, condition 3 only guarantees that $\eta(t) - \eta(t-) \in \text{span}[d(\alpha, h(t-)) \cup d(\alpha, h(t))]$.

Remark 11. Given a derivative process $\mathcal{J}^\alpha$ along a reflected diffusion Z^α, define the continuous $\{\mathcal{F}_t\}$-adapted process $\mathcal{H}^\alpha = \{\mathcal{H}^\alpha(t), t \geq 0\}$ taking values in $\mathbb{R}^{J\times M}$, for $t \geq 0$, by

$$\begin{aligned}\mathcal{H}^\alpha(t) := x_0'(\alpha) + \int_0^t b_\alpha(\alpha, Z^\alpha(s))ds + \int_0^t b_x(\alpha, Z^\alpha(s))\mathcal{J}^\alpha(s)ds \\ + \int_0^t \sigma_\alpha(\alpha, Z^\alpha(s))dW(s) + \int_0^t \sigma_x(\alpha, Z^\alpha(s))\mathcal{J}^\alpha(s)dW(s) + R'(\alpha)L^\alpha(t).\end{aligned} \tag{43}$$

For each $m = 1, \ldots, M$, it follows from the properties of $(\mathcal{J}_m^\alpha, \mathcal{H}_m^\alpha)$ stated in Definition 2 that a.s. $(\mathcal{J}_m^\alpha, \mathcal{H}_m^\alpha)$ is a solution to the derivative problem along Z^α for $\mathcal{H}_m^\alpha$. In other words, if $\Lambda_{Z^\alpha}^\alpha$ is a.s. well defined, then a.s. $\mathcal{J}_m^\alpha = \Lambda_{Z^\alpha}^\alpha(\mathcal{H}_m^\alpha)$.

In the following, we state some important properties of the derivative problem and its associated derivative map. The first lemma states a time-shift property of the derivative problem that is the analogue of the time-shift property of the ESP stated in Lemma 3.

Lemma 4. *Let (h, g) be a solution to the ESP $\{(d_i, n_i, c_i), i \in \mathcal{I}\}$ for $f \in \mathbb{D}_G(\mathbb{R}^J)$. Suppose (ϕ, η) is a solution to the derivative problem along h for $\psi \in \mathbb{D}(\mathbb{R}^J)$. Let $s \geq 0$ and define $h^s \in \mathbb{D}(G)$ as in (40), and define $\phi^s, \eta^s, \psi^s \in \mathbb{D}(\mathbb{R}^J)$ by*

$$\phi^s(\cdot) := \phi(s + \cdot), \qquad \eta^s(\cdot) := \eta(s + \cdot) - \eta(s), \qquad \psi^s(\cdot) := \phi(s) + \psi(s + \cdot) - \psi(s). \tag{44}$$

Then (ϕ^s, η^s) is a solution to the derivative problem along h^s for ψ^s.

Proof. This was shown for $f \in \mathbb{C}_G(\mathbb{R}^J)$ in lemma 5.2 of Lipshutz and Ramanan (2018). A similar argument can be used to prove the lemma in the case in which $f \in \mathbb{D}_G(\mathbb{R}^J)$. The only difference is that condition 4 must also be verified. To see that condition 4 holds, fix $s \geq 0$ and let $t > 0$. By the definition of η^s, condition 4 to the DP and the definition of h^s,

$$\eta^s(t) - \eta^s(t-) = \eta(s + t) - \eta((s + t)-) \in \text{span}[d(\alpha, h(s + t))] = \text{span}[d(\alpha, h^s(t))].$$

This proves condition 4 of the DP holds. □

Proposition 2. *Let $\alpha \in U$. There exists $\kappa_\Lambda(\alpha) < \infty$ such that, if (h, g) is the solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for $f \in \mathbb{D}_G(\mathbb{R}^J)$, and for $k = 1, 2$, (ϕ_k, η_k) is a solution to the derivative problem along h for $\psi_k \in \mathbb{D}(\mathbb{R}^J)$, then, for all $t < \infty$,*

$$\sup_{0\leq s\leq t} |\phi_1(s) - \phi_2(s)| \leq \kappa_\Lambda(\alpha) \sup_{0\leq s\leq t} |\psi_1(s) - \psi_2(s)|. \tag{45}$$

Proof. When $f \in \mathbb{C}_G(\mathbb{R}^J)$, the proposition follows from theorem 5.4 of Lipshutz and Ramanan (2018). When f is not continuous, the proof follows exactly analogously to the proof of theorem 5.4 of Lipshutz and Ramanan (2018) except that equation 3.5 of Lipshutz and Ramanan (2018) does not follow from condition 3 of the derivative problem, but rather is due to condition 4 of the derivative problem. □

Our final two results in this section state conditions under which the derivative problem along h is well defined and the associated derivative map is continuous. Both results require that (h, g) satisfy a version of the boundary jitter property, which was originally introduced in Lipshutz and Ramanan (2018) to characterize directional derivatives of ESMs. For the following, let $\mathcal{N}$ denote the nonsmooth part of the boundary; that is,

$$\mathcal{N} := \{x \in \partial G : |\mathcal{I}(x)| \geq 2\}. \tag{46}$$

Definition 6. A pair $(h, g) \in \mathbb{C}(G) \times \mathbb{C}(\mathbb{R}^J)$ satisfies the boundary jitter property if the following hold:

1. If $h(t) \in \partial G$ for some $t_1 < t < t_2$, then g is nonconstant on $(t_1 \vee 0, t_2)$.
2.′ h does not spend positive Lebesgue time in the boundary ∂G; that is,

$$\int_0^\infty 1_{\partial G}(h(s))ds = 0.$$

3. If $h(t) \in \mathcal{N}$ for some $t > 0$, then, for each $i \in \mathcal{I}(h(t))$ and all $\delta \in (0, t)$, there exists $s \in (t - \delta, t)$ such that $\mathcal{I}(h(s)) = \{i\}$.
4. If $h(0) \in \mathcal{N}$, then, for each $i \in \mathcal{I}(h(0))$ and all $\delta > 0$, there exists $s \in (0, \delta)$ such that $\mathcal{I}(h(s)) = \{i\}$.

Remark 12. Condition 2′ of the boundary jitter property is slightly stronger than condition 2 of definition 3.1 of Lipshutz and Ramanan (2018), which only requires that h does not spend positive Lebesgue time in the *nonsmooth* part of the boundary (i.e., where two or more faces intersect). The stronger version of condition 2 stated in Definition 6 is used in the proof of Theorem 4.

In Lipshutz and Ramanan (2019a), it was shown that a reflected diffusion and its constraining process a.s. satisfy the boundary jitter property. For the following, given $\alpha \in U$ and the associated reflected diffusion Z^α, let Y^α denote the constraining process defined in Remark 9.

Proposition 3. *Let $\alpha \in U$. Then a.s. (Z^α, Y^α) satisfies the boundary jitter property.*

Proof. Conditions 1, 3, and 4 of the jitter property hold because of theorem 3.3 of Lipshutz and Ramanan (2019a). Condition 2′ follows from lemma 4.3 of Lipshutz and Ramanan (2019a). □

Proposition 4. *Let $\alpha \in U$. Given $f \in \mathbb{C}_G(\mathbb{R}^J)$, suppose the solution (h, g) to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f satisfies the boundary jitter property (Definition 6). Then, given $\psi \in \mathbb{C}(\mathbb{R}^J)$, there exists a unique solution (ϕ, η) to the derivative problem along h for ψ. Furthermore, ϕ is continuous at times $t > 0$ such that $h(t) \in G^\circ \cup \mathcal{N}$.*

Proof. The proposition follows from theorem 3.11 and lemma 8.2 of Lipshutz and Ramanan (2018). □

The following continuity property of the derivative map is instrumental in the proof of our main result. In addition to its use in this work, it is also used in Lipshutz and Ramanan (2019b) to show that the joint reflected diffusion and derivative process is Feller continuous. In contrast to the other results in this section, we explicitly state exactly which assumptions are required for the following theorem. Recall that we have equipped $\mathbb{D}(\mathbb{R}^J)$ with the Skorokhod J_1-topology.

Theorem 4. *Suppose Assumptions 2 and 3 hold. Given $f \in \mathbb{C}_G(\mathbb{R}^J)$, suppose the solution (h, g) to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f satisfies the boundary jitter property (Definition 6). Let $\{f_l\}_{l \in \mathbb{N}}$ be a sequence in $\mathbb{D}(\mathbb{R}^J)$ such that f_l converges to f in $\mathbb{D}(\mathbb{R}^J)$ as $l \to \infty$, and for each $l \in \mathbb{N}$, let (h_l, g_l) denote the solution to the ESP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for f_l. Suppose $\psi \in \mathbb{C}(\mathbb{R}^J)$ satisfies $\psi(0) \in \mathbb{H}_{h(0)}$ and $\{\psi_l\}_{l \in \mathbb{N}}$ is a sequence in $\mathbb{D}(\mathbb{R}^J)$ such that ψ_l converges to ψ in $\mathbb{D}(\mathbb{R}^J)$ as $l \to \infty$. Then $\Lambda^\alpha_{h_l}(\psi_l)$ converges to $\Lambda^\alpha_h(\psi)$ in $\mathbb{D}(\mathbb{R}^J)$ as $l \to \infty$.*

Remark 13. The continuity result in Theorem 4 is somewhat subtle and cannot be strengthened in the sense explained as follows. Note that Theorem 4 establishes continuity of the mapping $(f, \psi) \mapsto \Lambda_{\bar{\Gamma}(f)}(\psi)$, where we recall that $\bar{\Gamma}$ denotes the ESM, only at continuous functions (f, ψ) for which the pair $(\bar{\Gamma}(f), \bar{\Gamma}(f) - f)$ satisfies the boundary jitter property and $\psi(0) \in \mathbb{H}_{\bar{\Gamma}(f)(0)}$. Because convergence in the Skorokhod J_1-topology to a continuous function implies uniform convergence on compact intervals, this continuity holds when the domain of the mapping is equipped with the topology of uniform convergence on compact intervals, and the range is equipped with the Skorokhod J_1-topology. It is not possible to strengthen the topology on the range to the topology of uniform convergence on compact intervals as illustrated by the following simple one-dimensional example with data $\{(e_1, e_1, 0)\}$ and with $f, \{f_l\}_{l \in \mathbb{N}}, \psi$ defined by $f(t) := 1 - t$, $f_l(t) := 1 + l^{-1} - t$ for $l \in \mathbb{N}$ and $\psi(t) := 1$ for $t \geq 0$. Furthermore, the condition that the pair $(\bar{\Gamma}(f), \bar{\Gamma}(f) - f)$ satisfies the boundary jitter property can also not be relaxed. This can be seen by analyzing the two-dimensional example with data $\{(e_i, e_i, 0), i = 1, 2\}$ and with $f, \{f_l\}_{l \in \mathbb{N}}, \psi$ defined by $f(t) := (1 - t, 1 - t)^T$, $f_l(t) := (1 - t, 1 + l^{-1} - t)^T$ for $l \in \mathbb{N}$ and $\psi(t) := (1, 1)^T$ for $t \geq 0$. We leave a detailed verification of these claims to the interested reader.

The proof of Theorem 4 is deferred to Section 9.

7. Convergence of the Euler Scheme for Reflected Diffusions

For $\Delta > 0$, recall the sequence $\{t^\Delta_n\}_{n \in \mathbb{N}_0}$ in $\mathbb{R}_+$ defined in (12), define the RCLL step function $\rho^\Delta : \mathbb{R}_+ \mapsto \mathbb{R}_+$ by

$$\rho^\Delta(t) = t^\Delta_n, \qquad t \in [t^\Delta_n, t^\Delta_{n+1}), \qquad n \in \mathbb{N}_0, \tag{47}$$

recall the definition of the discrete filtration $\{\mathcal{F}^\Delta_t\}$ given in (14), and define the piecewise constant RCLL $\{\mathcal{F}^\Delta_t\}$-adapted process $W^\Delta = \{W^\Delta(t), t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ by

$$W^\Delta(t) = W(t^\Delta_n), \qquad t \in [t^\Delta_n, t^\Delta_{n+1}), \qquad n \in \mathbb{N}_0. \tag{48}$$

Observe that ρ^Δ and W^Δ are constant on the interval $[t_n^\Delta, t_{n+1}^\Delta)$ for each $n \in \mathbb{N}_0$ and

$$\rho^\Delta(t_n^\Delta) - \rho^\Delta(t_n^\Delta -) = \Delta, \tag{49}$$

$$W^\Delta(t_n^\Delta) - W^\Delta(t_n^\Delta -) = \delta_n^\Delta W, \tag{50}$$

for each $n \in \mathbb{N}$, where $\delta_n^\Delta W$ is defined as in (13). Recall the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted processes $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ taking values in $G \times \mathbb{R}_+^N$ defined in (15)–(18). By (15)–(18) and (12), we have, for each $n \in \mathbb{N}$,

$$\begin{aligned}
Z^{\alpha,\Delta}(t_n^\Delta) &= Z^{\alpha,\Delta}(t_0^\Delta) + \sum_{j=1}^n \left(Z^{\alpha,\Delta}(t_j^\Delta) - Z^{\alpha,\Delta}(t_{j-1}^\Delta)\right) \\
&= x_0(\alpha) + \sum_{j=1}^n \left(\Xi_j^{\alpha,\Delta} - Z^{\alpha,\Delta}(t_{j-1}^\Delta) + \pi^\alpha(\Xi_j^{\alpha,\Delta}) - \Xi_j^{\alpha,\Delta}\right) \\
&= x_0(\alpha) + \sum_{j=1}^n b(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\Delta + \sum_{j=1}^n \sigma(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\delta_j^\Delta W + R(\alpha)\sum_{j=1}^n \xi^\alpha(\Xi_j^{\alpha,\Delta}) \\
&= x_0(\alpha) + \sum_{j=1}^n b(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\Delta + \sum_{j=1}^n \sigma(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\delta_j^\Delta W + R(\alpha)L^{\alpha,\Delta}(t_n^\Delta),
\end{aligned}$$

where the third equality uses the fact that, by Lemma 1, $R(\alpha)\xi^\alpha(\Xi_j^\alpha) = \pi^\alpha(\Xi_j^{\alpha,\Delta}) - \Xi_j^{\alpha,\Delta}$ for each $j \in \mathbb{N}$. Because $Z^{\alpha,\Delta}$, $L^{\alpha,\Delta}$, ρ^Δ, and W^Δ are constant on intervals of the form $[t_n^\Delta, t_{n+1}^\Delta)$ for $n \in \mathbb{N}_0$, it follows from (15), (49), (50), and the last display that, for all $t \geq 0$,

$$Z^{\alpha,\Delta}(t) = x_0(\alpha) + \int_0^t b(\alpha, Z^{\alpha,\Delta}(s-))d\rho^\Delta(s) + \int_0^t \sigma(\alpha, Z^{\alpha,\Delta}(s-))dW^\Delta(s) + R(\alpha)L^{\alpha,\Delta}(t). \tag{51}$$

In addition, by (15)–(17) and Lemma 1, we see that a.s. $L^{\alpha,\Delta}(0) = 0$, and for each $i \in \mathcal{I}$, the ith component of $L^{\alpha,\Delta}$ is nondecreasing and can only increase when $Z^{\alpha,\Delta}$ lies in F_i.

Remark 14. Define the J-dimensional piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted process $X^{\alpha,\Delta} = \{X^{\alpha,\Delta}(t), t \geq 0\}$, for $t \geq 0$, by

$$\begin{aligned}
X^{\alpha,\Delta}(t) &:= x_0(\alpha) + \int_0^t b(\alpha, Z^{\alpha,\Delta}(s-))d\rho^\Delta(s) + \int_0^t \sigma(\alpha, Z^{\alpha,\Delta}(s-))dW^\Delta(s) \\
&= x_0(\alpha) + \int_0^{\rho^\Delta(t)} b(\alpha, Z^{\alpha,\Delta}(s))ds + \int_0^{\rho^\Delta(t)} \sigma(\alpha, Z^{\alpha,\Delta}(s))dW(s).
\end{aligned} \tag{52}$$

Then, by the properties of $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ stated previously and Definition 3, we see that $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ is a solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for $X^{\alpha,\Delta}$.

We now prove the convergence of the Euler scheme for the reflected diffusion. The argument is analogous to the one given in Słomínski (2001) for reflected diffusions in convex polyhedral domains with normal reflection. Given a function $f : \mathbb{R}_+ \mapsto \mathbb{R}^J$ and $0 < \Delta < t$, define

$$\mathrm{Osc}(f, \Delta, t) := \sup\{|f(u) - f(s)| : s, u \in [0, t], |u - s| \leq \Delta\}. \tag{53}$$

Lemma 5 (Słomínski 2001, Lemma A.4). *For each $\alpha \in U$, $p \geq 2$, and $t < \infty$,*

$$\mathbb{E}[|\mathrm{Osc}(W, \Delta, t)|^p] = O\left(\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}}\right) \qquad \textit{as } \Delta \downarrow 0.$$

Proposition 5. *For each $\alpha \in U$, $p \geq 2$, and $t < \infty$,*

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} |X^{\alpha,\Delta}(s) - X^\alpha(s)|^p\right] = O\left(\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}}\right) \qquad \textit{as } \Delta \downarrow 0.$$

Proof. Fix $\alpha \in U$, $p \geq 2$, and $t < \infty$. The proof is analogous to the proof of theorem 3.2 of Slominki (2001), which assumes normal reflection along the boundary. In particular, following equations 3.3 and 3.4 of Slominski (2001), there exists a constant $C < \infty$ (depending on $\alpha \in U$ and $t < \infty$) such that

$$\mathbb{E}\left[\sup_{0\leq u\leq t} |X^{\alpha,\Delta}(u) - X^\alpha(u)|^p\right] \leq C\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}} + C\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |Z^{\alpha,\Delta}(u) - Z^\alpha(u)|^p\right] ds. \tag{54}$$

Because $Z^{\alpha,\Delta} = \Gamma^\alpha(X^{\alpha,\Delta})$ and $Z^\alpha = \Gamma^\alpha(X^\alpha)$ by Remarks 9 and 14 and the Lipschitz continuity of the SM shown in Proposition 1(i), we have, for $t \geq 0$,

$$\mathbb{E}\left[\sup_{0\leq u\leq t} |Z^{\alpha,\Delta}(u) - Z^\alpha(u)|^p\right] \leq (\kappa_\Gamma(\alpha))^p \mathbb{E}\left[\sup_{0\leq u\leq t} |X^{\alpha,\Delta}(u) - X^\alpha(u)|^p\right].$$

Substituting the last inequality into the integrand on the right-hand side of (54) and applying Gronwall's inequality yields

$$\mathbb{E}\left[\sup_{0\leq u\leq t} |X^{\alpha,\Delta}(u) - X^\alpha(u)|^p\right] \leq C\left(\Delta \log \frac{1}{\Delta}\right)^{\frac{p}{2}} \exp(C(\kappa_\Gamma(\alpha))^p t), \qquad t \geq 0,$$

which completes the proof of the lemma. □

Proof of Theorem 2. Recall that $\alpha \in U$ is given. Fix $p \geq 2$. Recall that (Z^α, L^α) is a solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for X^α by Remark 9, and $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ is a solution to the SP $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$ for $X^{\alpha,\Delta}$ by Remark 14. The theorem then follows from the Lipschitz continuity of the SM (see Proposition 1) and Proposition 5. □

8. Convergence of the Euler Scheme for the Derivative Process

In this section, we prove Theorem 3, our main result on the convergence of the discrete approximation of the derivative process. In Section 8.1, we express the discretized derivative process $\mathcal{J}^{\alpha,\Delta}$ in terms of the derivative map along $Z^{\alpha,\Delta}$ applied to an associated discretized process $\mathcal{H}^{\alpha,\Delta}$, see (63) and (64), that is more tractable to analyze. In Proposition 6 of Section 8.2, we state that the discretized process $\mathcal{H}^{\alpha,\Delta}$ converges to $\mathcal{H}^\alpha$ in an appropriate sense as the time step Δ approaches zero, where $\mathcal{H}^\alpha$ is a process associated with the derivative process. In particular, the derivative process can be expressed in terms of the derivative map along Z^α applied to $\mathcal{H}^\alpha$ (see Remark 11). We close Section 8.2 by combining the convergence result (Proposition 6) and the continuity property of the derivative map (Theorem 4) to prove Theorem 3. Finally, in Section 8.3, we prove Proposition 6.

For $\Delta > 0$, recall the sequence $\{t_n^\Delta\}_{n\in\mathbb{N}_0}$ in $\mathbb{R}_+$ defined in (12), the RCLL step function ρ^Δ defined in (47), the discrete filtration $\{\mathcal{F}_t^\Delta\}$ defined in (14), the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted process W^Δ defined in (48), the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted processes $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ taking values in $G \times \mathbb{R}_+^N$ defined in (15)–(18), and the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted process $\mathcal{J}^{\alpha,\Delta}$ taking values in $\mathbb{R}^{J\times M}$ defined in (23)–(25).

8.1. Relation Between the Discretized Derivative Process and the Derivative Map

Recall the definition of $\mathbb{H}_x$ given in (8) for $x \in G$. Suppose $x_0(\alpha) \in F_i$ for some $i \in \mathcal{I}$. Then, for all $\beta \in \mathbb{R}^M$ and $\varepsilon > 0$ sufficiently small so that $\alpha + \varepsilon\beta \in U$, we have

$$\varepsilon^{-1}\langle x_0(\alpha + \varepsilon\beta) - x_0(\alpha), n_i\rangle = \varepsilon^{-1}(\langle x_0(\alpha + \varepsilon\beta), n_i\rangle - c_i) \geq 0,$$

where the final inequality is because $x_0(\cdot)$ takes values in G. Because the last display holds for all $\beta \in \mathbb{R}^M$, by considering limits as $\varepsilon \to 0$ from the left and right, this implies that the column vectors of $x_0'(\alpha)$ take values in $\{y \in \mathbb{R}^J : \langle y, n_i\rangle = 0\}$. In particular, it follows that given $\alpha \in U$, the column vectors of $x_0'(\alpha)$ take values in $\mathbb{H}_{x_0(\alpha)}$. Thus, by (12), (23), and (15),

$$\mathcal{J}_m^{\alpha,\Delta}(t_0^\Delta) = [x_0'(\alpha)]_m \in \mathbb{H}_{x_0(\alpha)} = \mathbb{H}_{Z^{\alpha,\Delta}(t_0^\Delta)}, \qquad m = 1, \ldots, M. \tag{55}$$

Here $[x_0'(\alpha)]_m$ denotes the mth column vector of $x_0'(\alpha)$. By (23)–(26) and (15), we have, for each $n \in \mathbb{N}$,

$$\mathcal{J}_m^{\alpha,\Delta}(t_n^\Delta) \in \mathbb{H}_{Z^{\alpha,\Delta}(t_n)}, \qquad m = 1, \ldots, M, \tag{56}$$

and

$$\begin{aligned}\mathcal{J}^{\alpha,\Delta}(t_n^\Delta) &= \mathcal{J}^{\alpha,\Delta}(t_0^\Delta) + \sum_{j=1}^n \left(\mathcal{J}^{\alpha,\Delta}(t_j^\Delta) - \mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta)\right) \\ &= x_0'(\alpha) + \sum_{j=1}^n \left(\mathcal{X}_j^{\alpha,\Delta} - \mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta) + \mathcal{L}^\alpha_{Z^{\alpha,\Delta}(t_j^\Delta)}\mathcal{X}_j^{\alpha,\Delta} - \mathcal{X}_j^{\alpha,\Delta}\right) \\ &= x_0'(\alpha) + \sum_{j=1}^n b_\alpha(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\Delta + \sum_{j=1}^n b_x(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta)\Delta + \sum_{j=1}^n \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\delta_j^\Delta W \\ &\quad + \sum_{j=1}^n \sigma_x(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta)\delta_j^\Delta W + R'(\alpha)\sum_{j=1}^n (L^{\alpha,\Delta}(t_j^\Delta) - L^{\alpha,\Delta}(t_{j-1}^\Delta)) + \sum_{j=1}^n \left(\mathcal{L}^\alpha_{Z^{\alpha,\Delta}(t_j^\Delta)}\mathcal{X}_j^{\alpha,\Delta} - \mathcal{X}_j^{\alpha,\Delta}\right) \\ &= x_0'(\alpha) + \sum_{j=1}^n b_\alpha(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\Delta + \sum_{j=1}^n b_x(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta)\Delta + \sum_{j=1}^n \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\delta_j^\Delta W \\ &\quad + \sum_{j=1}^n \sigma_x(\alpha, Z^{\alpha,\Delta}(t_{j-1}^\Delta))\mathcal{J}^{\alpha,\Delta}(t_{j-1}^\Delta)\delta_j^\Delta W + R'(\alpha)L^{\alpha,\Delta}(t_n^\Delta) + \mathcal{K}^{\alpha,\Delta}(t_n^\Delta),\end{aligned} \tag{57}$$

where $\mathcal{K}^{\alpha,\Delta} = \{\mathcal{K}^{\alpha,\Delta}(t), t \geq 0\}$ is the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted process taking values in $\mathbb{R}^{J\times M}$ defined by $\mathcal{K}^{\alpha,\Delta}(0) = 0$ and, for each $n \in \mathbb{N}$, $\mathcal{K}^{\alpha,\Delta}(t) := \mathcal{K}^{\alpha,\Delta}(t_{n-1}^\Delta)$ for all $t \in [t_{n-1}^\Delta, t_n^\Delta)$ and

$$\mathcal{K}^{\alpha,\Delta}(t_n^\Delta) := \sum_{j=1}^n \left(\mathcal{L}^\alpha_{Z^{\alpha,\Delta}(t_j^\Delta)}\mathcal{X}_j^{\alpha,\Delta} - \mathcal{X}_j^{\alpha,\Delta}\right). \tag{58}$$

Because $\mathcal{J}^{\alpha,\Delta}$, $Z^{\alpha,\Delta}$, $L^{\alpha,\Delta}$, ρ^Δ, W^Δ, and $\mathcal{K}^{\alpha,\Delta}$ are constant on intervals of the form $[t_n^\Delta, t_{n+1}^\Delta)$ for $n \in \mathbb{N}_0$, it follows from (55)–(57), (23), (47), and (48) that, for all $t \geq 0$,

$$\mathcal{J}_m^{\alpha,\Delta}(t) \in \mathbb{H}_{Z^{\alpha,\Delta}(t)}, \qquad m = 1, \ldots, M, \tag{59}$$

and

$$\begin{aligned}\mathcal{J}^{\alpha,\Delta}(t) &= x_0'(\alpha) + \int_0^t b_\alpha(\alpha, Z^{\alpha,\Delta}(s-))d\rho^\Delta(s) + \int_0^t b_x(\alpha, Z^{\alpha,\Delta}(s-))\mathcal{J}^{\alpha,\Delta}(s-)d\rho^\Delta(s) \\ &\quad + \int_0^t \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(s-))dW^\Delta(s) + \int_0^t \sigma_x(\alpha, Z^{\alpha,\Delta}(s-))\mathcal{J}^{\alpha,\Delta}(s-)dW^\Delta(s) + R'(\alpha)L^{\alpha,\Delta}(t) + \mathcal{K}^{\alpha,\Delta}(t).\end{aligned} \tag{60}$$

By (58) and Lemma 2, we see that, for each $m = 1, \ldots, M$ and $n \in \mathbb{N}$,

$$\mathcal{K}_m^{\alpha,\Delta}(t_n^\Delta) - \mathcal{K}_m^{\alpha,\Delta}(t_n^\Delta-) = \mathcal{L}^\alpha_{Z^{\alpha,\Delta}(t_n^\Delta)}[\mathcal{X}_n^{\alpha,\Delta}]_m - [\mathcal{X}_n^{\alpha,\Delta}]_m \in \operatorname{span}\left[d(\alpha, Z^{\alpha,\Delta}(t_n^\Delta))\right].$$

Here $[\mathcal{X}_n^{\alpha,\Delta}]_m$ denotes the mth column vector of the $J \times M$ matrix $\mathcal{X}_n^{\alpha,\Delta}$. Because $\mathcal{K}^{\alpha,\Delta}$ and $Z^{\alpha,\Delta}$ are piecewise constant, it follows that, for all $m = 1, \ldots, M$ and $0 \leq s < t < \infty$,

$$\mathcal{K}_m^{\alpha,\Delta}(t) - \mathcal{K}_m^{\alpha,\Delta}(s) \in \operatorname{span}\left[\cup_{u\in(s,t]} d(\alpha, Z^{\alpha,\Delta}(u))\right] \tag{61}$$

and for all $m = 1, \ldots, M$ and $t > 0$,

$$\mathcal{K}_m^{\alpha,\Delta}(t) - \mathcal{K}_m^{\alpha,\Delta}(t-) \in \operatorname{span}\left[d(\alpha, Z^{\alpha,\Delta}(t))\right]. \tag{62}$$

Remark 15. Define the piecewise constant RCLL $\{\mathcal{F}_t^\Delta\}$-adapted process $\mathcal{H}^{\alpha,\Delta} = \{\mathcal{H}^{\alpha,\Delta}(t), t \geq 0\}$ taking values in $\mathbb{R}^{J\times M}$, for $t \geq 0$, by

$$\begin{aligned}\mathcal{H}^{\alpha,\Delta}(t) &= x_0'(\alpha) + \int_0^t b_\alpha(\alpha, Z^{\alpha,\Delta}(s-))d\rho^\Delta(s) + \int_0^t b_x(\alpha, Z^{\alpha,\Delta}(s-))\mathcal{J}^{\alpha,\Delta}(s-)d\rho^\Delta(s)\\ &\quad + \int_0^t \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(s-))dW^\Delta(s) + \int_0^t \sigma_x(\alpha, Z^{\alpha,\Delta}(s-))\mathcal{J}^{\alpha,\Delta}(s-)dW^\Delta(s) + R'(\alpha)L^{\alpha,\Delta}(t)\\ &= x_0'(\alpha) + \int_0^{\rho^\Delta(t)} b_\alpha(\alpha, Z^{\alpha,\Delta}(s))ds + \int_0^{\rho^\Delta(t)} b_x(\alpha, Z^{\alpha,\Delta}(s))\mathcal{J}^{\alpha,\Delta}(s)ds\\ &\quad + \int_0^{\rho^\Delta(t)} \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(s))dW(s) + \int_0^{\rho^\Delta(t)} \sigma_x(\alpha, Z^{\alpha,\Delta}(s))\mathcal{J}^{\alpha,\Delta}(s)dW(s) + R'(\alpha)L^{\alpha,\Delta}(t).\end{aligned} \tag{63}$$

It follows from (59), (57), and (61)–(63) that, for each $m = 1, \ldots, M$, a.s. $(\mathcal{J}_m^{\alpha,\Delta}, \mathcal{H}_m^{\alpha,\Delta})$ is a solution to the derivative problem along $Z^{\alpha,\Delta}$ for $\mathcal{H}_m^{\alpha,\Delta}$. In other words, a.s.

$$\mathcal{J}_m^{\alpha,\Delta} = \Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha,\Delta}), \qquad m = 1, \ldots, M. \tag{64}$$

8.2. Proof of Theorem 3

In preparation for proving Theorem 3, we first state some preliminary results.

Lemma 6. *For each $\alpha \in U$, $p \geq 2$, and $t < \infty$, there exists $C_1^{(\alpha,p,t)} < \infty$ such that for all $\Delta > 0$,*

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |L^{\alpha,\Delta}(s)|^p\right] \leq C_1^{(\alpha,p,t)}. \tag{65}$$

Proof. Fix $\alpha \in U$, $p \geq 2$, and $t < \infty$. The fact that

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |X^{\alpha,\Delta}(s) - x_0(\alpha)|^p\right] < \infty \tag{66}$$

follows from a standard argument using (52), Hölder's inequality, the BDG inequalities, Tonelli's theorem, Assumption 4, the fact that $Z^{\alpha,\Delta} = \Gamma^\alpha(X^{\alpha,\Delta})$ by Remark 14 and $h = \Gamma^\alpha(f)$ where $h(\cdot) = f(\cdot) \equiv x_0(\alpha)$, the Lipschitz continuity of the SM (Proposition 1) and Gronwall's inequality. Then, using the fact that $(Z^{\alpha,\Delta}, L^{\alpha,\Delta})$ is the solution to the SP for $X^{\alpha,\Delta}$ by Remark 14, the fact that $(h, 0)$ is the solution to the SP for f, (66), and the Lipschitz continuity of the SM, we obtain (65). □

Lemma 7. *For each $\alpha \in U$, $p \geq 2$, and $t < \infty$, there exists $C_2^{(\alpha,p,t)} < \infty$ such that, for all $\Delta > 0$,*

$$\mathbb{E}\left[\sup_{0\leq s\leq t} \|\mathcal{J}^{\alpha,\Delta}(s)\|^p\right] \leq C_2^{(\alpha,p,t)}. \tag{67}$$

Proof. Fix $\alpha \in U$, $p \geq 2$, and $t < \infty$. Let $\Delta > 0$. For brevity, let $\kappa := \max(\kappa_1, \kappa_\Lambda(\alpha)) < \infty$, where $\kappa_1 < \infty$ is the constant from Assumption 4. By (63), Hölder's inequality, the BDG inequalities, Tonelli's theorem, and bounds on the coefficients stated in Assumption 4,

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq u\leq t} \|\mathcal{H}^{\alpha,\Delta}(u)\|^p\right] &\leq 4^{p-1}\|x_0'(\alpha)\|^p + 4^{p-1}\mathbb{E}\left[\sup_{0\leq s\leq t} |R'(\alpha)L^{\alpha,\Delta}(s)|^p\right]\\ &\quad + 4^{p-1}(t^{p-1} + C_p)\kappa^p \int_0^t \left(1 + \mathbb{E}\left[\sup_{0\leq u\leq s} \|\mathcal{J}^{\alpha,\Delta}(u)\|^p\right]\right) ds.\end{aligned}$$

By (64), the Lipschitz continuity of the derivative map $\Lambda_{Z^{\alpha,\Delta}}$ (Proposition 2), and Lemma 6, we have

$$\mathbb{E}\left[\sup_{0\le u\le t}\left\|\mathcal{J}^{\alpha,\Delta}(u)\right\|^p\right] \le 4^{p-1}\kappa^p\left\|x_0'(\alpha)\right\|^p + 4^{p-1}\kappa^p\|R'(\alpha)\|^p C_1^{(\alpha,p,t)}$$
$$+ 4^{p-1}(t^{p-1}+C_p)\kappa^{2p}\int_0^t\left(1+\mathbb{E}\left[\sup_{0\le u\le s}\left\|\mathcal{J}^{\alpha,\Delta}(u)\right\|^p\right]\right)ds.$$

An application of Gronwall's inequality yields (67) with

$$C_2^{(\alpha,p,t)} := \left(1 + 4^{p-1}\kappa^p\left\|x_0'(\alpha)\right\|^p + 4^{p-1}\kappa^p\|R'(\alpha)\|^p C_1^{(\alpha,p,t)}\right)e^{4^{p-1}(t^{p-1}+C_p)\kappa^{2p}t}.$$

Recall the derivative processes $\mathcal{J}^\alpha$ introduced in Definition 2 and the associated process $\mathcal{H}^\alpha$ introduced in Remark 11. We recall a basic estimate.

Lemma 8 (Lipshutz and Ramanan 2019a, Lemma 5.7). *For each $\alpha\in U$, $p\ge 2$, and $t<\infty$,*

$$\mathbb{E}\left[\sup_{0\le s\le t}\|\mathcal{J}^\alpha(s)\|^p\right]<\infty \qquad \text{and} \qquad \mathbb{E}\left[\sup_{0\le s\le t}\|\mathcal{H}^\alpha(s)\|^p\right]<\infty.$$

Proposition 6. *For each $\alpha\in U$, $p\ge 2$, and $t<\infty$,*

$$\lim_{\Delta\downarrow 0}\mathbb{E}\left[\sup_{0\le s\le t}\left\|\mathcal{H}^{\alpha,\Delta}(s)-\mathcal{H}^\alpha(s)\right\|^p\right]=0. \tag{69}$$

We first show how Theorem 3 and Corollary 2 can be deduced from Proposition 6. The proof of Proposition 6 is given in Section 8.3.

Proof of Theorem 3. Let $m=1,\dots,M$. By Theorem 2, Proposition 6, and an application of the Skorokhod representation theorem, we can assume that a.s. $(Z^{\alpha,\Delta},\mathcal{H}_m^{\alpha,\Delta})$ converges to $(Z^\alpha,\mathcal{H}_m^\alpha)$ uniformly on compact time intervals as $\Delta\downarrow 0$. By Proposition 3, a.s. (Z^α,Y^α) satisfies the boundary jitter property. Then, by the fact that $\mathcal{J}_m^\alpha=\Lambda_{Z^\alpha}(\mathcal{H}_m^\alpha)$ and $\mathcal{J}_m^{\alpha,\Delta}=\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha,\Delta})$ for $\Delta>0$, and the continuity of the derivative map shown in Theorem 4, a.s. $\mathcal{J}_m^{\alpha,\Delta}$ converges to $\mathcal{J}_m^\alpha$ in the J_1-topology as $\Delta\downarrow 0$. By part (iii) of Theorem 1, Proposition 3, condition 2′ of the boundary jitter property and lemma 4.13 of Lipshutz and Ramanan (2019a) a.s. $\mathcal{J}^\alpha$ is continuous at t and at almost every $s\in(0,t)$. Thus, a.s. $\mathcal{J}^{\alpha,\Delta}$ converges to $\mathcal{J}^\alpha$ at t and at almost every $s\in(0,t)$, which, along with Lemmas 7 and 8, completes the proof of Theorem 3. □

8.3. Proof of Proposition 6

Fix $\alpha\in U$. For brevity, throughout this section, we let

$$\kappa := \max(\kappa_1,\kappa_2,\kappa_\Gamma(\alpha),\kappa_\Lambda(\alpha))<\infty,$$

where κ_1 and κ_2 denote the constants in Assumption 4, $\kappa_\Gamma(\alpha)$ denotes the constant in Proposition 1, and $\kappa_\Lambda(\alpha)$ denotes the constant in Proposition 2.

We first prove an estimate that relates the derivative map along the reflected diffusion with the derivative map along the Euler discretization of the reflected diffusion.

Lemma 9. *For each $p\ge 2$, $m=1,\dots,M$, and $t<\infty$,*

$$\lim_{\Delta\downarrow 0}\mathbb{E}\left[\int_0^t|\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^\alpha)(s)-\Lambda_{Z^\alpha}(\mathcal{H}_m^\alpha)(s)|^p ds\right]=0. \tag{70}$$

Proof. Fix $p\ge 2$, $m=1,\dots,M$, and $t<\infty$. Recall the constraining process Y^α introduced in Remark 2 and the process X^α defined in (39). By Remark 9 and Proposition 3, a.s.

a. (Z^α,Y^α) is the solution to the ESP for X^α, and
b. (Z^α,Y^α) satisfies the boundary jitter property.

For $\Delta>0$, recall the process $X^{\alpha,\Delta}$ defined in (52). Then by Remark 14, Proposition 5, and an argument using the Borel–Cantelli lemma (see Słomínski 1994 for an example of such an argument), a.s.

c. $Z^{\alpha,\Delta}=\bar\Gamma(X^{\alpha,\Delta})$, and
d. $X^{\alpha,\Delta}$ converges to X^α in $\mathbb{D}(\mathbb{R}^J)$ as $\Delta\downarrow 0$.

Because (a)–(d) hold a.s., it follows from Theorem 4 that a.s. $\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^\alpha)$ converges to $\Lambda_{Z^\alpha}(\mathcal{H}_m^\alpha)$ in $\mathbb{D}(\mathbb{R}^J)$ as $\Delta \downarrow 0$. Thus, a.s.

$$\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^\alpha)(s) \to \Lambda_{Z^\alpha}(\mathcal{H}_m^\alpha)(s) \text{ as } \Delta \downarrow 0 \text{ for almost every } s \in [0,t]. \tag{71}$$

By the Lipschitz continuity of the derivative map (Proposition 2),

$$\sup_{0\le s\le t} |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^\alpha)(s)|^p + \sup_{0\le s\le t} |\Lambda_{Z^\alpha}(\mathcal{H}_m^\alpha)(s)|^p \le 2\kappa^p \sup_{0\le s\le t} |\mathcal{H}_m^\alpha(s)|^p.$$

Therefore, by Lemma 8, the dominated convergence theorem, and (71), (70) holds. □

Proof of Proposition 6. Let $p \ge 2$ and $t < \infty$. By (63) and (43), for each $s \in [0,t]$,

$$\left\|\mathcal{H}^{\alpha,\Delta}(s) - \mathcal{H}^\alpha(s)\right\| \tag{72}$$
$$\le \left\|\int_{\rho^\Delta(s)}^s b_\alpha(\alpha, Z^{\alpha,\Delta}(u))du\right\| \tag{73}$$
$$+ \left\|\int_0^s \{b_\alpha(\alpha, Z^{\alpha,\Delta}(u)) - b_\alpha(\alpha, Z^\alpha(u))\}du\right\| \tag{74}$$
$$+ \left\|\int_{\rho^\Delta(s)}^s b_x(\alpha, Z^{\alpha,\Delta}(u))\mathcal{J}^{\alpha,\Delta}(u)du\right\| \tag{75}$$
$$+ \left\|\int_0^s \{b_x(\alpha, Z^{\alpha,\Delta}(u)) - b_x(\alpha, Z^\alpha(u))\}\mathcal{J}^{\alpha,\Delta}(u)du\right\| \tag{76}$$
$$+ \left\|\int_0^s b_x(\alpha, Z^\alpha(u))(\mathcal{J}^{\alpha,\Delta}(u) - \mathcal{J}^\alpha(u))du\right\| \tag{77}$$
$$+ \left\|\int_{\rho^\Delta(s)}^s \sigma_\alpha(\alpha, Z^{\alpha,\Delta}(u))dW(u)\right\| \tag{78}$$
$$+ \left\|\int_0^s \{\sigma_\alpha(\alpha, Z^{\alpha,\Delta}(u)) - \sigma_\alpha(\alpha, Z^\alpha(u))\}dW(u)\right\| \tag{79}$$
$$+ \left\|\int_{\rho^\Delta(s)}^s \sigma_x(\alpha, Z^{\alpha,\Delta}(u))\mathcal{J}^{\alpha,\Delta}(u)dW(u)\right\| \tag{80}$$
$$+ \left\|\int_0^s \{\sigma_x(\alpha, Z^{\alpha,\Delta}(u)) - \sigma_x(\alpha, Z^\alpha(u))\}\mathcal{J}^{\alpha,\Delta}(u)dW(u)\right\| \tag{81}$$
$$+ \left\|\int_0^s \sigma_x(\alpha, Z^\alpha(u))\left(\mathcal{J}^{\alpha,\Delta}(u) - \mathcal{J}^\alpha(u)\right)dW(u)\right\| \tag{82}$$
$$+ \left\|R'(\alpha)(L^{\alpha,\Delta}(s) - L^\alpha(s))\right\|. \tag{83}$$

We treat each of the 11 terms on the right-hand side of the last display separately.

Term (73): By Assumption 4 and the fact that $|s - \rho^\Delta(s)| \le \Delta$ for all $s \in [0,t]$,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_{\rho^\Delta(s)}^s b_\alpha(\alpha, Z^{\alpha,\Delta}(u))du\right|^p\right] \le \Delta^p\kappa^p. \tag{84}$$

Term (74): By Hölder's inequality and Assumption 4, with the associated constant γ,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s \{b_\alpha(\alpha, Z^{\alpha,\Delta}(u)) - b_\alpha(\alpha, Z^\alpha(u))\}du\right\|^p\right] \le t^p\kappa^p\mathbb{E}\left[\sup_{0\le s\le t} |Z^{\alpha,\Delta}(s) - Z^\alpha(s)|^{\gamma p}\right]. \tag{85}$$

Term (75): By Hölder's inequality, Assumption 4, and the fact that $|s-\rho^\Delta(s)|\le\Delta$ for all $s\in[0,t]$,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_{\rho^\Delta(s)}^s b_x(\alpha,Z^{\alpha,\Delta}(u))\mathcal{J}^{\alpha,\Delta}(u)du\right\|^p\right]\le\Delta^p\kappa^p\mathbb{E}\left[\sup_{0\le s\le t}\left\|\mathcal{J}^{\alpha,\Delta}(s)\right\|^p\right]. \tag{86}$$

Term (76): By Hölder's inequality, Assumption 4, and the Cauchy–Schwarz inequality,

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s\{b_x(\alpha,Z^{\alpha,\Delta}(u))-b_x(\alpha,Z^\alpha(u))\}\mathcal{J}^{\alpha,\Delta}(u)du\right\|^p\right]\\&\quad\le t^p\kappa^p\left\{\mathbb{E}\left[\sup_{0\le s\le t}|Z^{\alpha,\Delta}(s)-Z^\alpha(s)|^{2\gamma p}\right]\mathbb{E}\left[\sup_{0\le s\le t}\left\|\mathcal{J}^{\alpha,\Delta}(s)\right\|^{2p}\right]\right\}^{\frac{1}{2}}.\end{aligned} \tag{87}$$

Term (77): By Hölder's inequality, Tonelli's theorem, and Assumption 4,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s b_x(\alpha,Z^\alpha(u))(\mathcal{J}^{\alpha,\Delta}(u)-\mathcal{J}^\alpha(u))du\right\|^p\right]\le t^{p-1}\kappa^p\int_0^t\mathbb{E}\left[\left\|\mathcal{J}^{\alpha,\Delta}(s)-\mathcal{J}^\alpha(s)\right\|^p\right]ds. \tag{88}$$

Term (78): By Assumption 4, (53), and the fact that $|s-\rho^\Delta(s)|\le\Delta$ for all $s\in[0,t]$,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_{\rho^\Delta(s)}^s\sigma_\alpha(\alpha,Z^{\alpha,\Delta}(u))dW(u)\right\|^p\right]\le\kappa^p\mathbb{E}[|\mathrm{Osc}(W,\Delta,t)|^p]. \tag{89}$$

Term (79): By the BDG inequalities and Assumption 4,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s\{\sigma_\alpha(\alpha,Z^{\alpha,\Delta}(u))-\sigma_\alpha(\alpha,Z^\alpha(u))\}dW(u)\right\|^p\right]\le C_p\kappa^pt^{\frac{p}{2}}\mathbb{E}\left[\sup_{0\le s\le t}|Z^{\alpha,\Delta}(s)-Z^\alpha(s)|^{\gamma p}\right]. \tag{90}$$

Term (80): By Assumption 4, the Cauchy–Schwarz inequality, (53), and the fact that $|s-\rho^\Delta(s)|\le\Delta$ for all $s\in[0,t]$,

$$E\left[\sup_{0\le s\le t}\left\|\int_{\rho^\Delta(s)}^s\sigma_x(\alpha,Z^{\alpha,\Delta}(u))\mathcal{J}^{\alpha,\Delta}(u)dW(u)\right\|^p\right]\le\kappa^p\left\{\mathbb{E}\left[\sup_{0\le s\le t}\left\|\mathcal{J}^{\alpha,\Delta}(s)\right\|^{2p}\right]\mathbb{E}[|\mathrm{Osc}(W,\Delta,t)|^{2p}]\right\}^{\frac{1}{2}}. \tag{91}$$

Term (81): By the BDG inequalities, Assumption 4, and the Cauchy–Schwarz inequality,

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s\{\sigma_x(\alpha,Z^{\alpha,\Delta}(u))-\sigma_x(\alpha,Z^\alpha(u))\}\mathcal{J}^{\alpha,\Delta}(u)dW(u)\right\|^p\right]\\&\quad\le C_p\kappa^pt^{\frac{p}{2}}\left\{\mathbb{E}\left[\sup_{0\le s\le t}|Z^{\alpha,\Delta}(s)-Z^\alpha(s)|^{2\gamma p}\right]\mathbb{E}\left[\sup_{0\le s\le t}\left\|\mathcal{J}^{\alpha,\Delta}(s)\right\|^{2p}\right]\right\}^{\frac{1}{2}}.\end{aligned} \tag{92}$$

Term (82): By the BDG inequalities, Tonelli's theorem, and Assumption 4,

$$\mathbb{E}\left[\sup_{0\le s\le t}\left\|\int_0^s\sigma_x(\alpha,Z^\alpha(u))\left(\mathcal{J}^{\alpha,\Delta}(u)-\mathcal{J}^\alpha(u)\right)dW(u)\right\|^p\right]\le C_p\kappa^pt^{\frac{p}{2}-1}\int_0^t\mathbb{E}[\left\|\mathcal{J}^{\alpha,\Delta}(s)-\mathcal{J}^\alpha(s)\right\|^p]ds. \tag{93}$$

Term (83): By Assumption 4, the facts that $(Z^{\alpha,\Delta},L^{\alpha,\Delta})$ solves the SP for $X^{\alpha,\Delta}$ and (Z^α,L^α) solves the SP for X^α, and the Lipschitz continuity of the SM (Proposition 1),

$$E\left[\left\|\sup_{0\le s\le t}R'(\alpha)(L^{\alpha,\Delta}(s)-L^\alpha(s))\right\|^p\right]\le\kappa^{2p}\mathbb{E}\left[\sup_{0\le s\le t}|X^{\alpha,\Delta}(s)-X^\alpha(s)|^p\right]. \tag{94}$$

Before combining the terms, we observe that, because $\mathcal{J}_m^{\alpha,\Delta} = \Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha,\Delta})$ and $\mathcal{J}_m^{\alpha} = \Lambda_{Z^{\alpha}}(\mathcal{H}_m^{\alpha})$ for each $m = 1,\dots,M$, it follows from the Lipschitz continuity of the derivative map that

$$
\begin{aligned}
|\mathcal{J}_m^{\alpha,\Delta}(s) - \mathcal{J}_m^{\alpha}(s)| &\le |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha,\Delta})(s) - \Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha})(s)| + |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha})(s) - \Lambda_{Z^{\alpha}}(\mathcal{H}_m^{\alpha})(s)| \\
&\le \kappa \sup_{0\le u\le s} |\mathcal{H}_m^{\alpha,\Delta}(u) - \mathcal{H}_m^{\alpha}(u)| + |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha})(s) - \Lambda_{Z^{\alpha}}(\mathcal{H}_m^{\alpha})(s)|.
\end{aligned}
$$

Thus,

$$
\|\mathcal{J}^{\alpha,\Delta}(s) - \mathcal{J}^{\alpha}(s)\|^p \le (2M)^{p-1}\kappa^p \sup_{0\le u\le s} \|\mathcal{H}^{\alpha,\Delta}(u) - \mathcal{H}^{\alpha}(u)\|^p + (2M)^{p-1} \sum_{m=1}^{M} |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha})(s) - \Lambda_{Z^{\alpha}}(\mathcal{H}_m^{\alpha})(s)|^p. \tag{95}
$$

Now combining (72)–(95) and applying Gronwall's inequality, we obtain

$$
\mathbb{E}\left[\sup_{0\le s\le t} \|\mathcal{H}^{\alpha,\Delta}(s) - \mathcal{H}^{\alpha}(s)\|^p\right] \le 11^{p-1}\kappa^p C_\Delta \exp(11^{p-1}(M+1)^{p-1}\kappa^{2p}(t^p + C_p t)),
$$

where

$$
\begin{aligned}
C_\Delta := \Delta^p &\left(1 + \mathbb{E}\left[\sup_{0\le s\le t} \|\mathcal{J}^{\alpha,\Delta}(s)\|^p\right]\right) + \kappa^p \mathbb{E}[|\mathrm{Osc}(W,\Delta,t)|^p] + \left\{\mathbb{E}\left[\sup_{0\le s\le t} \|\mathcal{J}^{\alpha,\Delta}(s)\|^{2p}\right] \mathbb{E}[|\mathrm{Osc}(W,\Delta,t)|^{2p}]\right\}^{\frac{1}{2}} \\
&+ (t^p + C_p t^{\frac{p}{2}})\mathbb{E}\left[\sup_{0\le s\le t} |Z^{\alpha,\Delta}(s) - Z^{\alpha}(s)|^{\gamma p}\right] + (t^p + C_p t^{\frac{p}{2}})\left\{\mathbb{E}\left[\sup_{0\le s\le t} |Z^{\alpha,\Delta}(s) - Z^{\alpha}(s)|^{2\gamma p}\right] \mathbb{E}\left[\sup_{0\le s\le t} \|\mathcal{J}^{\alpha,\Delta}(s)\|^{2p}\right]\right\}^{\frac{1}{2}} \\
&+ \kappa^p \mathbb{E}\left[\sup_{0\le s\le t} |X^{\alpha,\Delta}(s) - X^{\alpha}(s)|^p\right] + (2M)^{p-1}(t^{p-1} + C_p t^{\frac{p}{2}-1}) \sum_{m=1}^{M} \mathbb{E}\left[\int_0^t |\Lambda_{Z^{\alpha,\Delta}}(\mathcal{H}_m^{\alpha})(s) - \Lambda_{Z^{\alpha}}(\mathcal{H}_m^{\alpha})(s)|^p ds\right].
\end{aligned}
$$

By Lemma 7, Lemma 5, Theorem 2, Proposition 5, and Lemma 9, $C_\Delta \to 0$ as $\Delta \downarrow 0$, thus completing the proof of (69). □

Remark 16. From the final display in the proof of Proposition 6 and in view of Theorem 2, Lemma 5, and Proposition 5, we see that a challenge to obtaining a rate of convergence for our IPA method as the discretization parameter converges to zero is the lack of *uniform* continuity for the derivative map; specifically, uniform continuity of the map $f \mapsto \Lambda^{\alpha}_{\bar{\Gamma}(f)}(\psi)$ at $f \in \mathbb{C}_G(\mathbb{R}^J)$ such that $(\bar{\Gamma}(f), \bar{\Gamma}(f) - f)$ satisfies the boundary jitter property. In fact, we strongly suspect that such a uniform continuity property does not hold because the proof of continuity of Theorem 4 relies heavily on path properties of $(\bar{\Gamma}(f), \bar{\Gamma}(f) - f)$.

9. Continuity of the Derivative Map

In this section, we prove Theorem 4. In Section 9.1, we show that it suffices to consider the case when ψ lies in a certain class of simple functions. In Section 9.2, we characterize solutions to the derivative problem in terms of the derivative projection matrices (introduced in Lemma 2) when ψ is constant or when ψ is simple and h does not hit the nonsmooth part of the boundary. In Section 9.3, we prove the continuity property of the derivative map when ψ lies in a certain class of simple functions.

Throughout this section, we fix $\alpha \in U$. For brevity, we drop the α notation and write $d(\cdot)$, Λ_h, and $\mathcal{L}_x$ for $d(\alpha,\cdot)$, Λ_h^{α}, and $\mathcal{L}_x^{\alpha}$, respectively.

9.1. Proof of Theorem 4

In this section, we first prove Theorem 4 when $\psi_l = \psi$ for all $l \in \mathbb{N}$ and ψ lies in a class of simple functions (see Lemma 11). With that in mind, let $\mathbb{D}_s(\mathbb{R}^J)$ denote the set of simple functions in $\mathbb{D}(\mathbb{R}^J)$; that is,

$$
\mathbb{D}_s(\mathbb{R}^J) := \left\{\sum_{j=1}^{m} y_j 1_{[s_j, s_{j+1})} : m \in \mathbb{N}, y_1, \dots, y_m \in \mathbb{R}^J, 0 \le s_1 < \cdots < s_{m+1} < \infty\right\}.
$$

For $h \in \mathbb{C}(G)$, let

$$
B_h := \{t \ge 0 : h(t) \in \partial G\}
$$

denote the times in $[0,\infty)$ that h lies in the boundary ∂G, and for $\psi\in\mathbb{D}(\mathbb{R}^J)$, let

$$J_\psi := \{t>0 : \psi(t)\neq\psi(t-)\}$$

denote the jump or discontinuity points of ψ in $(0,\infty)$. Recall the definition of $\mathbb{H}_x$, $x\in G$, given in (8). For $h\in\mathbb{C}(G)$, define

$$\mathbb{D}_{\mathrm{s}}^h(\mathbb{R}^J) := \{\psi\in\mathbb{D}_{\mathrm{s}}(\mathbb{R}^J) : \psi(0)\in\mathbb{H}_{h(0)} \text{ and } B_h\cap J_\psi=\emptyset\}. \tag{96}$$

Lemma 10. *Given $h\in\mathbb{C}(G)$, suppose B_h has zero Lebesgue measure. Then, for $\psi\in\mathbb{C}(\mathbb{R}^J)$ with $\psi(0)\in\mathbb{H}_{h(0)}$, there is a sequence $\{\psi_k\}_{k\in\mathbb{N}}$ in $\mathbb{D}_{\mathrm{s}}^h(\mathbb{R}^J)$ such that ψ_k converges to ψ uniformly on compact intervals as $k\to\infty$.*

Proof. Fix $\psi\in\mathbb{C}(\mathbb{R}^J)$ with $\psi(0)\in\mathbb{H}_{h(0)}$. Let $t<\infty$ and $\varepsilon>0$ be arbitrary. It suffices to show that there exists $\tilde\psi\in\mathbb{D}_{\mathrm{s}}^h(\mathbb{R}^J)$ such that $\sup_{0\le s\le t}|\psi(s)-\tilde\psi(s)|<\varepsilon$. Because $\mathbb{D}_{\mathrm{s}}(\mathbb{R}^J)$ is dense in $\mathbb{C}(\mathbb{R}^J)$ under the topology of uniform convergence on compact intervals (see, e.g., theorem 6.2.2 of Whitt 2002), we can choose $\hat\psi\in\mathbb{D}_{\mathrm{s}}(\mathbb{R}^J)$ such that $\hat\psi(0)=\psi(0)$ and $\sup_{0\le s\le t}|\psi(s)-\hat\psi(s)|<\varepsilon/2$. By the definition of $\mathbb{D}_{\mathrm{s}}(\mathbb{R}^J)$ and the fact that $\hat\psi(0)=\psi(0)$, there exist $m\in\mathbb{N}$, $y_1,\dots,y_m\in\mathbb{R}^J$, and $0=s_1<\cdots<s_{m+1}$ such that $y_1=\psi(0)$ and

$$\hat\psi=\sum_{j=1}^m y_j 1_{[s_j,s_{j+1})}.$$

Without loss of generality, we can assume $s_{m+1}>t$. Because ψ is uniformly continuous on $[0,t]$, we can choose $\delta>0$ sufficiently small so that $s_{m+1}>t+\delta$, $s_j+\delta<s_{j+1}-\delta$ for each $j=1,\dots,m$, and

$$\sup\{|\psi(u)-\psi(s)| : s,u\in[0,t], |u-s|<\delta\}<\varepsilon/2.$$

Set $\tilde s_1=s_1=0$. For each $j=2,\dots,m+1$, let $\tilde s_j\in(s_j-\delta,s_j+\delta)\setminus B_h$, which is nonempty because B_h has zero Lebesgue measure. Define $\tilde\psi\in\mathbb{D}_{\mathrm{s}}^h(\mathbb{R}^J)$ by

$$\tilde\psi=\sum_{j=1}^m y_j 1_{[\tilde s_j,\tilde s_{j+1})}.$$

Then $\tilde s_{m+1}>t$. Let $s\in[0,t]$ and $j\in\{1,\dots,m\}$ be such that $s\in[\tilde s_j,\tilde s_{j+1})\subset(s_j-\delta,s_{j+1}+\delta)$. Then there exists $u\in(s_j,s_{j+1})$ such that $|u-s|<\delta$ and so $\tilde\psi(s)=y_j=\hat\psi(u)$ and

$$|\psi(s)-\tilde\psi(s)|\le|\psi(s)-\psi(u)|+|\psi(u)-\hat\psi(u)|<\varepsilon.$$

Because $s\in[0,t]$ was arbitrary, this completes the proof. □

We now state a key convergence lemma, Lemma 11, and show how it, together with Lemma 10, can be used to prove Theorem 4. The proof of Lemma 11 is deferred to Section 9.3 after first establishing some preliminary results in Section 9.2.

Lemma 11. *Given $f\in\mathbb{C}_G(\mathbb{R}^J)$, suppose that the solution (h,g) to the ESP for f satisfies the boundary jitter property. Let $\{f_l\}_{l\in\mathbb{N}}$ be a sequence in $\mathbb{D}(\mathbb{R}^J)$ such that f_l converges to f in $\mathbb{D}(\mathbb{R}^J)$ as $l\to\infty$, and for each $l\in\mathbb{N}$, let (h_l,g_l) denote the solution to the ESP for f_l. Then, for all $\psi\in\mathbb{D}_{\mathrm{s}}^h(\mathbb{R}^J)$, $\Lambda_{h_l}(\psi)$ converges to $\Lambda_h(\psi)$ in $\mathbb{D}(\mathbb{R}^J)$ as $l\to\infty$.*

Proof of Theorem 4. Recall that $\alpha\in U$ is fixed and, for any $h\in\mathbb{D}(\mathbb{R}^J)$, denote Λ_h^α just by Λ_h. Fix $f\in\mathbb{C}_G(\mathbb{R}^J)$ and a sequence $\{f_l\}_{l\in\mathbb{N}}$ in $\mathbb{D}(\mathbb{R}^J)$ such that f_l converges to f in $\mathbb{D}(\mathbb{R}^J)$ as $l\to\infty$. Let (h,g) denote the solution to the ESP for f and, for each $l\in\mathbb{N}$, let (h_l,g_l) denote the solution to the ESP for f_l. Fix $\psi\in\mathbb{C}(\mathbb{R}^J)$ and a sequence $\{\psi_l\}_{l\in\mathbb{N}}$ in $\mathbb{D}(\mathbb{R}^J)$ such that ψ_l converges to ψ in $\mathbb{D}(\mathbb{R}^J)$ as $l\to\infty$. Let $d_{J_1}(\cdot,\cdot)$ and $d_U(\cdot,\cdot)$ denote metrics on $\mathbb{D}(\mathbb{R}^J)$ that are respectively compatible with the Skorokhod J_1-topology and the topology of uniform convergence on compact intervals. Because the topology of uniform convergence on compact intervals is a stronger topology than the Skorokhod J_1-topology, we can assume the metrics are chosen such that $d_{J_1}(\cdot,\cdot)\le d_U(\cdot,\cdot)$. Then, by the triangle inequality and Lipschitz continuity of the derivative map (Proposition 2),

$$\begin{aligned} d_{J_1}(\Lambda_{h_l}(\psi_l),\Lambda_h(\psi)) &\le d_U(\Lambda_{h_l}(\psi_l),\Lambda_{h_l}(\psi))+d_{J_1}(\Lambda_{h_l}(\psi),\Lambda_h(\psi))\\ &\le \kappa_\Lambda d_U(\psi_l,\psi)+d_{J_1}(\Lambda_{h_l}(\psi),\Lambda_h(\psi)). \end{aligned}$$

Because $\psi \in \mathbb{C}(\mathbb{R}^J)$ and the Skorokhod J_1-topology relativized to $\mathbb{C}(\mathbb{R}^J)$ coincides with the topology of uniform convergence on compact intervals there, $d_U(\psi_l, \psi)$ converges to zero as $l \to \infty$. We are left to show $d_{J_1}(\Lambda_{h_l}(\psi), \Lambda_h(\psi))$ converges to zero as $l \to \infty$. According to Lemma 10, there is a sequence $\{\psi_k\}_{k\in\mathbb{N}}$ in $\mathbb{D}_s^h(\mathbb{R}^J)$ such that ψ_k converges to ψ in $\mathbb{D}(\mathbb{R}^J)$ as $k \to \infty$. For each $k \in \mathbb{N}$, by the Lipschitz continuity of the derivative map,

$$\begin{aligned} d_{J_1}(\Lambda_{h_l}(\psi), \Lambda_h(\psi)) &\le d_{J_1}(\Lambda_{h_l}(\psi), \Lambda_{h_l}(\psi_k)) + d_{J_1}(\Lambda_{h_l}(\psi_k), \Lambda_h(\psi_k)) + d_{J_1}(\Lambda_h(\psi_k), \Lambda_h(\psi)) \\ &\le 2\kappa_\Lambda d_U(\psi, \psi_k) + d_{J_1}(\Lambda_{h_l}(\psi_k), \Lambda_h(\psi_k)). \end{aligned}$$

By Lemma 11, $d_{J_1}(\Lambda_{h_l}(\psi_k), \Lambda_h(\psi_k))$ converges to zero as $l \to \infty$. Then letting $k \to \infty$, $d_U(\psi, \psi_k)$ converges to zero. This completes the proof. □

9.2. Characterization of Solutions to the Derivative Problem

In Lemma 12, we characterize solutions to the DP along h for simple ψ up until the first time h hits the intersection of two or more faces. In the next section, we use an inductive argument (on the first time h hits n or more faces) to prove continuity properties of the derivative map. With that in mind, given a solution (h, g) to the ESP for $f \in \mathbb{D}_G(\mathbb{R}^J)$, define

$$\theta_n := \inf\{t \ge 0 : |\mathcal{I}(h(t))| \vee |\mathcal{I}(h(t-))| \ge n\}, \qquad n = 1, \dots, N, \tag{97}$$

and set $\theta_{N+1} := \infty$. Suppose $\psi \in \mathbb{D}_s(\mathbb{R}^J)$ is of the form

$$\psi = \sum_{j=1}^{m} y_j 1_{[s_j, s_{j+1})} \tag{98}$$

for some $m \in \mathbb{N}$, $y_1, \dots, y_m \in \mathbb{R}^J$ such that $y_j \ne y_{j+1}$ for $1 \le j < m$, and $0 = s_1 < \cdots < s_{m+1} = \theta_2$. For each $1 \le j \le m$, define $K_j \in \mathbb{N}_\infty$ and the sequence $\{\rho_j(k)\}_{k=1,\dots,K_j}$ in $[s_j, s_{j+1}]$ as follows: set $\rho_j(1) := s_j$ and, for $k \in \mathbb{N}$ such that $\rho_j(k) < s_{j+1}$, recursively define

$$\rho_j(k+1) := \inf\{t > \rho_j(k) : \mathcal{I}(h(t)) \not\subseteq \mathcal{I}(h(\rho_j(k)))\} \wedge s_{j+1}. \tag{99}$$

If $\rho_j(k) = s_{j+1}$ for some $k \in \mathbb{N}$, then set $K_j := k$ and end the sequence so that

$$s_j = \rho_j(1) < \cdots < \rho_j(K_j) = s_{j+1}. \tag{100}$$

On the other hand, if $\rho_j(k) < s_{j+1}$ for all $k \in \mathbb{N}$, then set $K_j := \infty$. In the case that $K_j = \infty$, because $\rho_j(k)$ is increasing and bounded above by s_{j+1}, there exists $\rho_j(\infty) \le s_{j+1}$ such that $\rho_j(k) \to \rho_j(\infty)$ as $k \to \infty$. If $\rho_j(\infty) = \infty$, then we must have $s_{j+1} = \infty$, and hence, $\theta_2 = \infty$. Alternatively, if $\rho_j(\infty) < \infty$, then the definition of $\rho_j(k)$ given in (99) implies that $|\mathcal{I}(h(\rho_j(\infty)-))| \ge 2$, which, along with the fact that $\rho_j(k) \le \theta_2$ for all $k \in \mathbb{N}$ and the definition of θ_2 given in (97), implies $\rho_j(\infty) = \theta_2$. In either case, we see that if $K_j = \infty$, then $\rho_j(k) \to \theta_2$ as $k \to \infty$. Hence, because $s_{m+1} = \theta_2$, we have $K_j = \infty$ only if $j = m$. In this case,

$$s_m = \rho_m(1) < \rho_m(2) < \cdots < s_{m+1} = \theta_2. \tag{101}$$

Lemma 12. *Given $f \in \mathbb{D}_G(\mathbb{R}^J)$, let (h, g) denote the solution to the ESP for f and define θ_2 as in (97). Suppose $\psi \in \mathbb{D}_s(\mathbb{R}^J)$ satisfies (98). For each $1 \le j \le m$, define $K_j \in \mathbb{N}_\infty$ and the sequence $\{\rho_j(k)\}_{k=1,\dots,K_j}$ as in (99). Then the solution (ϕ, η) to the derivative problem along h for ψ on $[0, \theta_2)$ satisfies, for each $1 \le j \le m$,*

$$\phi(t) = \mathcal{L}_{h(\rho_j(k))} \cdots \mathcal{L}_{h(\rho_j(1))}(\phi(s_j-) + y_j) \qquad \text{for } t \in [\rho_j(k), \rho_j(k+1)),\ 1 \le k < K_j, \tag{102}$$

where $\phi(0-) := 0$.

Proof. To prove the lemma, it suffices to show that $(\phi, \phi - \psi)$, where ϕ is defined as in (102), satisfies conditions 1–4 of the derivative problem along h for ψ on the interval $[0, \theta_2)$. Condition 1 holds automatically. We next show that condition 2 holds. Given $t \in [0, \theta_2)$, let $1 \le j \le m$ and $1 \le k < K_j$ be such that $t \in [\rho_j(k), \rho_j(k+1))$. By (99), $\mathcal{I}(h(t)) \subseteq \mathcal{I}(h(\rho_j(k)))$. Therefore, because $\mathcal{L}_{h(\rho_j(k))}$ maps $\mathbb{R}^J$ into $\mathbb{H}_{h(\rho_j(k))}$, it follows that

$$\phi(t) = \mathcal{L}_{h(\rho_j(k))} \cdots \mathcal{L}_{h(\rho_j(1))}(\phi(s_j-) + y_j) \in \mathbb{H}_{h(\rho_j(k))} \subseteq \mathbb{H}_{h(t)}.$$

Therefore, condition 2 of the derivative problem holds. We are left to show that conditions 3 and 4 of the derivative problem hold. First, observe that, for $1 \le j \le m$ and $1 \le k < K_j$, (98) and (102) imply that ψ and ϕ are constant on the interval $[\rho_j(k), \rho_j(k+1))$, and hence, $\eta := \phi - \psi$ is constant there as well. Therefore, to prove $(\phi, \phi - \psi)$ satisfies both conditions 3 and 4 of the derivative problem, it suffices to show that

$$\eta(\rho_j(k)) - \eta(\rho_j(k)-) \in \text{span}[d(h(\rho_j(k)))] \tag{103}$$

for each $1 \le j \le m$ and $1 \le k < K_j$. Suppose $1 \le j \le m$ and $k = 1$. Then $\rho_j(1) = s_j$ by definition, so the facts that $\eta := \phi - \psi$, $\phi(s_j) = \mathcal{L}_{h(s_j)}(\phi(s_j-) + y_j)$ by (102), $\psi(s_j) = \psi(s_j-) + y_j$ by (98), and $\mathcal{L}_{h(s_j)}(\phi(s_j-) + y_j) - (\phi(s_j-) + y_j) \in \text{span}[d(h(s_j))]$ by Lemma 2, together imply

$$\begin{aligned}\eta(\rho_j(1)) - \eta(\rho_j(1)-) &= \phi(s_j) - \phi(s_j-) - (\psi(s_j) - \psi(s_j-))\\ &= \mathcal{L}_{h(s_j)}(\phi(s_j-) + y_j) - (\phi(s_j-) + y_j)\\ &\in \text{span}[d(h(\rho_j(1)))].\end{aligned}$$

Next, suppose $1 \le j \le m$ and $1 < k < K_j$. Then the facts that $\eta := \phi - \psi$, $\phi(\rho_j(k)) = \mathcal{L}_{h(\rho_j(k))}(\phi(\rho_j(k)-))$ by (102), ψ is continuous at $\rho_j(k)$ by (98), and $\mathcal{L}_{h(\rho_j(k))}(\phi(\rho_j(k)-)) - \phi(\rho_j(k)-) \in \text{span}[d(h(\rho_j(k))]$ together imply

$$\begin{aligned}\eta(\rho_j(k)) - \eta(\rho_j(k)-) &= \phi(\rho_j(k)) - \phi(\rho_j(k)-)\\ &= \mathcal{L}_{h(\rho_j(k))}\phi(\rho_j(k)-) - \phi(\rho_j(k)-)\\ &\in \text{span}[d(h(\rho_j(k)))].\end{aligned}$$

This verifies (103) and, thus, proves the lemma. □

9.3. Proof of Lemma 11

We use the principle of mathematical induction to prove the following statement holds for $n = 2, \ldots, N+1$. Because $\theta_{N+1} := \infty$ by definition, this completes the proof of Lemma 11. Given $T < \infty$, we say that a function $\lambda : [0,T) \to [0,T)$ is a *time change on* $[0,T)$ if λ is nondecreasing and maps $[0,T)$ onto itself.

Statement 1. Given $f \in \mathbb{C}_G(\mathbb{R}^J)$, suppose that the solution (h,g) to the ESP for f satisfies the boundary jitter property (Definition 6). Let $\{f_l\}_{l\in\mathbb{N}}$ be a sequence in $\mathbb{D}(\mathbb{R}^J)$ such that f_l converges to f in $\mathbb{D}(\mathbb{R}^J)$ as $l \to \infty$, and for each $l \in \mathbb{N}$, let (h_l, g_l) denote the solution to the ESP for f_l. Then given $\psi \in \mathbb{D}^h_s(\mathbb{R}^J)$, there is a sequence $\{\lambda_l\}_{l\in\mathbb{N}}$ of time changes on $[0,\theta_n)$ such that for all $T < \infty$,

$$\lim_{l\to\infty} \sup_{t\in[0,\theta_n\wedge T)} |\lambda_l(t) - t| = 0, \tag{104}$$

and for all $T < \theta_n$,

$$\lim_{l\to\infty} \sup_{t\in[0,T]} \left|\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)\right| = 0. \tag{105}$$

We first prove the base case $n = 2$.

Lemma 13. *Statement* 1 *holds with* $n = 2$.

Because the proof of the base case is lengthy, we first provide a brief outline of the argument. We show that, roughly speaking, because (h,g) satisfies the boundary jitter property and f_l converges to f as $l \to \infty$, it follows that, on a compact interval $[0,T]$ contained in $[0,\theta_2)$ and for l sufficiently large, h_l hits the same boundary faces that h hits, and h_l hits those boundary faces in the same order that h does. The time change λ_l on $[0,\theta_2)$ is then constructed so that, on $[0,T]$, the paths $\psi \circ \lambda_l$ and ψ are equal, and the paths $h_l \circ \lambda_l$ and h hit the same boundary faces in the same order and at the same times. The paths $\Lambda_{h_l}(\psi) \circ \lambda_l$ and $\Lambda_h(\psi)$ are then equal on $[0,T]$ because the jumps of both paths depend only on the paths $\psi \circ \lambda_l$ and ψ, respectively, and on the times and the order in which $h_l \circ \lambda_l$ and h, respectively, hit the boundary faces.

The following is an upper semicontinuity property of the set-valued function $\mathcal{I}(\cdot)$ that is needed in the proof of Lemma 13.

Lemma 14 (Kang and Williams 2007, lemma 2.1). *For each $x \in G$, there is an open neighborhood U_x of x in $\mathbb{R}^J$ such that $\mathcal{I}(y) \subseteq \mathcal{I}(x)$ for all $y \in U_x \cap G$.*

Proof of Lemma 13. Let f, (h,g), $\{f_l\}_{l\in\mathbb{N}}$, $\{(h_l,g_l)\}_{l\in\mathbb{N}}$, and ψ be as in Statement 1. Define θ_2 as in (97). If $\theta_2 = 0$, then (104) and (105) hold trivially. For the remainder of the proof, we assume that $\theta_2 > 0$. Because f_l converges to f in $\mathbb{D}(\mathbb{R}^J)$ as $l \to \infty$, $f \in \mathbb{C}_G(\mathbb{R}^J)$, and the Skorokhod J_1-topology relativized to $\mathbb{C}(\mathbb{R}^J)$ coincides with the topology of uniform convergence on compact intervals, the Lipschitz continuity of the ESM stated in Proposition 1(ii) implies the following:

a. (h_l, g_l) converges to (h, g) uniformly on compact intervals as $l \to \infty$.

Because we are only concerned with the interval $[0, \theta_2)$ in this proof and ψ lies in $\mathbb{D}_s^h(\mathbb{R}^J)$, we can assume without loss of generality that, as specified in (98), $\psi = \sum_{j=1}^m y_j 1_{[s_j, s_{j+1})}$ for some $m \in \mathbb{N}$, $y_1 \in \mathbb{H}_{h(0)}$, $y_2, \ldots, y_m \in \mathbb{R}^J$ such that $y_j \neq y_{j+1}$ for $1 \le j < m$, and $0 = s_1 < \cdots < s_{m+1} < \infty$ with $s_{m+1} \le \theta_2$. For each $1 \le j \le m$, define $K_j \in \mathbb{N}_\infty$ and the sequence $\{\rho_j(k)\}_{k=1,\ldots,K_j}$ in $[s_j, s_{j+1}]$ as in (99) and the subsequent two sentences. As noted, following (99), $K_j < \infty$ for each $1 \le j < m$ and

$$\rho_m(K_m) = \theta_2 = \infty \qquad \text{if } K_m < \infty, \tag{106}$$

$$\lim_{k\to\infty} \rho_m(k) = \theta_2 \qquad \text{if } K_m = \infty. \tag{107}$$

By the definition of $\rho_j(k)$ in (99), the definition of θ_2 in (97) and condition 1 of the boundary jitter property, for each $1 \le j \le m$ and $1 < k < K_j$, there is a unique index $i_j(k) \in \mathcal{I}$ such that

b. $\mathcal{I}(h(\rho_j(k))) = \{i_j(k)\}$ and g is nonconstant on every neighborhood of $\rho_j(k)$.

By the upper semicontinuity of $\mathcal{I}(\cdot)$ shown in Lemma 14, the continuity of h, the definition of $\rho_j(k)$ in (99) and (b), for each $1 \le j \le m$ and $1 \le k < K_j$, there exists $\xi_j(k) \in (\rho_j(k), \rho_j(k+1))$, which is not necessarily unique, such that

$$h(t) \in G^\circ, \qquad t \in [s_j, \xi_j(1)], \qquad 1 \le j \le m, \tag{108}$$

$$h(t) \in G^\circ, \qquad t \in [\xi_j(k-1), \rho_j(k)), \qquad 1 \le j \le m,\ 1 < k < K_j, \tag{109}$$

$$\mathcal{I}(h(t)) \subseteq \{i_j(k)\}, \qquad t \in [\rho_j(k), \xi_j(k)], \qquad 1 \le j \le m,\ 1 < k < K_j, \tag{110}$$

$$h(t) \in G^\circ, \qquad t \in [\xi_j(K_j - 1), s_{j+1}], \qquad 1 \le j < m. \tag{111}$$

Set $l_1 := 1$. By (108)–(111), the upper semicontinuity of $\mathcal{I}(\cdot)$, (b) and (a), for each $1 < \hat{k} < K_m$, we can choose $l_{\hat{k}} \ge l_{\hat{k}-1}$ such that for all $l \ge l_{\hat{k}}$,

$$h_l(t) \in G^\circ, \qquad t \in [s_j, \xi_j(1)], \qquad 1 \le j \le m, \tag{112}$$

$$\mathcal{I}(h_l(t)) \subseteq \{i_j(k)\}, \qquad t \in [\xi_j(k-1), \xi_j(k)], \qquad 1 \le j < m,\ 1 < k < K_j, \tag{113}$$

$$h_l(t) \in G^\circ, \qquad t \in [\xi_j(K_j - 1), s_{j+1}], \qquad 1 \le j < m, \tag{114}$$

$$\mathcal{I}(h_l(t)) \subseteq \{i_m(k)\}, \qquad t \in [\xi_m(k-1), \xi_m(k)], \qquad 1 < k \le \hat{k}, \tag{115}$$

and

c. g_l is nonconstant on $(\xi_j(k-1), \xi_j(k))$ for $1 \le j < m$ and $1 < k < K_j$.

d. g_l is nonconstant on $(\xi_m(k-1), \xi_m(k))$ for $1 < k \le \hat{k}$.

By (113), (115), (c), (d), and the fact that g_l can only increase when h_l lies in ∂G, we have

e. For each $1 \le j < m$ and $1 < k < K_j$, $\mathcal{I}(h_l(t)) = \{i_j(k)\}$ for some $t \in (\xi_j(k-1), \xi_j(k))$.

f. For each $1 < k \le \hat{k}$, $\mathcal{I}(h_l(t)) = \{i_m(k)\}$ for some $t \in (\xi_m(k-1), \xi_m(k))$.

Given $l \in \mathbb{N}$, define $\theta_2^l := \inf\{t \ge 0 : |\mathcal{I}(h_l(t))| \ge 2\}$, and for each $1 \le j \le m$, define $K_j^l \in \mathbb{N}_\infty$ and the sequence $\{\rho_j^l(k)\}_{k=1,\ldots,K_j^l}$ in $[s_j, s_{j+1}]$ as follows: set $\rho_j^l := s_j$ and, for $k \in \mathbb{N}$ such that $\rho_j^l(k) < s_{j+1}$, define

$$\rho_j^l(k+1) := \inf\{t > \rho_j^l(k) : \mathcal{I}(h_l(t)) \not\subseteq \mathcal{I}(h_l(\rho_j(k)))\} \wedge s_{j+1}. \tag{116}$$

If $\rho_j^l(k) = s_{j+1}$ for some $k \in \mathbb{N}$, then set $K_j^l := s_{j+1}$. Alternatively, if $\rho_j^l(k) < s_{j+1}$ for all $k \in \mathbb{N}$, then set $K_j^l := \infty$. Let $1 \le \hat{k} < K_m$ and suppose $l \ge l_{\hat{k}}$. Then (116), the definition of K_j^l for $1 \le j \le m$, (112)–(115), (e) and (f) imply that

g. $K_j^l = K_j$ for $1 \le j < m$.

h. $\rho_j^l(k) \in (\xi_j(k-1), \xi_j(k))$ and $\mathcal{I}(h_l(\rho_j^l(k))) = \{i_j(k)\} = \mathcal{I}(h(\rho_j(k)))$ for $1 \le j < m$ and $1 < k < K_j$.

i. $\rho_m^l(k) \in (\xi_m(k-1), \xi_m(k))$ and $\mathcal{I}(h_l(\rho_m^l(k))) = \{i_m(k)\} = \mathcal{I}(h(\rho_m(k)))$ for $1 < k \le \hat{k}$.

Now fix $1 \le j \le m$ and $1 < k < K_j$. We show that

$$\lim_{l\to\infty} \rho_j^l(k) = \rho_j(k). \tag{117}$$

Because $\xi_j(k-1) < \rho_j(k) < \xi_j(k)$, we can choose $\delta > 0$ sufficiently small so that $(\rho_j(k) - \delta, \rho_j(k) + \delta) \subseteq [\xi_j(k-1), \xi_j(k)]$. Let $1 \le \hat{k} < K_m$ be such that $\rho_j(k) \le \rho_m(\hat{k})$ (e.g., if $j < m$, choose $\hat{k} = 1$, and if $j = m$, choose $\hat{k} = k$). By (a), (109), and (b)–(d), there exists a positive integer $l^\delta \ge l_{\hat{k}}$ sufficiently large such that, if $l \ge l^\delta$, then $h_l(t) \in G^\circ$ for

$t \in [\xi_j(k-1), \rho_j(k)-\delta]$ and g_l is nonconstant on $(\rho_j(k)-\delta, \rho_j(k)+\delta)$. This, along with (e) and (f), implies that $\rho_j^l(k) \in (\rho_j(k)-\delta, \rho_j(k)+\delta)$ for all $l \geq l^\delta$. Because $\delta > 0$ was arbitrary, (117) holds.

We now define the time changes λ_l on $[0, \theta_2)$. Let $\{l_{\hat{k}}\}_{1\leq \hat{k} < K_m}$ be such that, for each $1 \leq \hat{k} < K_m$ and all $l \geq l_{\hat{k}}$, (112)–(114) hold. Fix $l \in \mathbb{N}$. If $l < l_1$, then set $\lambda_l(t) := t$ for all $t \in [0, \theta_2)$. Now suppose $l \geq l_1$. Let $1 \leq \hat{k} < K_m$ be such that $l_{\hat{k}} \leq l < l_{\hat{k}+1}$. For $1 \leq j < m$ and $1 \leq k < K_j$ and for $j = m$ and $1 \leq k < \hat{k}$, define

$$\lambda_l(t) := \rho_j^l(k) + \frac{\rho_j^l(k+1) - \rho_j^l(k)}{\rho_j(k+1) - \rho_j(k)}(t - \rho_j(k)), \qquad t \in [\rho_j(k), \rho_j(k+1)), \tag{118}$$

and, if $\theta_2 < \infty$, define

$$\lambda_l(t) := \rho_m^l(\hat{k}+1) + \frac{\theta_2 - \rho_m^l(\hat{k}+1)}{\theta_2 - \rho_m(\hat{k}+1)}(t - \rho_m(\hat{k}+1)), \qquad t \in [\rho_m(\hat{k}+1), \theta_2), \tag{119}$$

whereas if $\theta_2 = \infty$, define

$$\lambda_l(t) := t + (\rho_m^l(\hat{k}+1) - \rho_m(\hat{k}+1))e^{-(t-\rho_m(\hat{k}+1))}, \qquad t \in [\rho_m(\hat{k}+1), \infty). \tag{120}$$

It is readily verified from (118)–(120) that, for each $l \geq l_1$, λ_l is a nondecreasing, continuous, piecewise linear function mapping $[0, \theta_2)$ onto $[0, \theta_2)$ such that for each $1 \leq j \leq m$,

$$\lambda_l(\rho_j(k)) = \rho_j^l(k), \qquad 1 < k < K_j, \tag{121}$$

$$\lambda_l(\rho_m(k)) = \rho_m^l(k), \qquad 1 < k \leq \hat{k}. \tag{122}$$

We now turn to the proofs of (104) and (105).

We begin with the proof of (104). Let $T < \infty$. We first treat the case that $T < \theta_2$. By (106) and (107), we can choose $1 \leq \hat{k} < K_m$ such that $T < \rho_m(\hat{k})$. Then by (118)–(120), for all $l \geq l_{\hat{k}}$,

$$\sup_{t\in[0,T)} |\lambda_l(t) - t| \leq \max_{1<k\leq \hat{k}} |\rho_j^l(k) - \rho_j(k)|. \tag{123}$$

Letting $l \to \infty$, it follows from (117) that (104) holds. Next, consider the case in which $\theta_2 \leq T < \infty$. Let $\varepsilon \in (0, \theta_2)$ be arbitrary. Because $\theta_2 - \varepsilon < \theta_2$ and (104) holds whenever $T < \theta_2$, we have

$$\lim_{l\to\infty} \sup_{t\in[0,\theta_2-\varepsilon)} |\lambda_l(t) - t| = 0. \tag{124}$$

Therefore, it suffices to show that

$$\lim_{\varepsilon\downarrow 0} \lim_{l\to\infty} \sup_{t\in[\theta_2-\varepsilon,\theta_2)} |\lambda_l(t) - t| = 0. \tag{125}$$

By the triangle inequality and the fact that $\lambda_l : [0, \theta_2) \mapsto [0, \theta_2)$ is nondecreasing, we have, for all $\varepsilon \in (0, \theta_2)$ and $t \in [\theta_2 - \varepsilon, \theta_2)$,

$$\begin{aligned} |\lambda_l(t) - t| &\leq |\lambda_l(t) - \theta_2| + |\theta_2 - t| \\ &\leq |\lambda_l(\theta_2 - \varepsilon) - \theta_2| + \varepsilon \\ &\leq |\lambda_l(\theta_2 - \varepsilon) - (\theta_2 - \varepsilon)| + 2\varepsilon. \end{aligned} \tag{126}$$

By (124) and the continuity of λ_l, $\lambda_l(\theta_2 - \varepsilon) \to \theta_2 - \varepsilon$ as $l \to \infty$, which, along with (126), yields (125). This establishes (104).

We now turn to the proof of (105), which completes the proof that Statement 1 holds with $n = 2$. Let $T < \theta_2$. By (106) and (107), we can choose $1 < \hat{k} < K_m$ such that $T < \rho_m(\hat{k})$. We show that for each $l \geq l_{\hat{k}}$,

$$\Lambda_{h_l}(\psi)(\lambda_l(t)) = \Lambda_h(\psi)(t) \qquad \text{for all } t \in [0, \rho_m(\hat{k})), \tag{127}$$

which completes the proof of (105). Fix $l \geq l_{\hat{k}}$. By Lemma 12, for each $1 \leq j \leq m$ and $1 \leq k < K_j$, $\Lambda_h(\psi)$ satisfies

$$\Lambda_h(\psi)(t) = \mathscr{L}_{h(\rho_j(k))} \cdots \mathscr{L}_{h(\rho_j(1))}(\Lambda_h(\psi)(s_j-) + y_j), \tag{128}$$

for all $t \in [\rho_j(k), \rho_j(k+1))$, where $\Lambda_h(\psi)(0-) := 0$. Similarly, by Lemma 12, for each $1 \le j \le m$ and $1 \le k < K_j^l = K_j$ (by (g)), $\Lambda_{h_l}(\psi)$ satisfies

$$\begin{aligned}\Lambda_{h_l}(\psi)(t) &= \mathscr{L}_{h_l(\rho_j^l(k))} \cdots \mathscr{L}_{h_l(\rho_j^l(1))}(\Lambda_{h_l}(\psi)(s_j-) + y_j) \\ &= \mathscr{L}_{h(\rho_j(k))} \cdots \mathscr{L}_{h(\rho_j(1))}(\Lambda_{h_l}(\psi)(s_j-) + y_j),\end{aligned} \tag{129}$$

for $t \in [\rho_j^l(k), \rho_j^l(k+1))$, where $\Lambda_{h_l}(\psi)(0-) := 0$ and we have used (h) and (i) in the second equality and properties of the derivative projection matrix in Lemma 2. By (121) and (122), for $1 \le j < m$ and $1 \le k < K_j$, or for $j = m$ and $1 \le k < \hat{k}$,

$$\Lambda_{h_l}(\psi)(\lambda_l(t)) = \mathscr{L}_{h(\rho_j(k))} \cdots \mathscr{L}_{h(\rho_j(1))}(\Lambda_{h_l}(\psi)(s_j-) + y_j), \quad t \in [\rho_j(k), \rho_j(k+1)). \tag{130}$$

Then, because $\Lambda_{h_l}(\psi)(0) = \psi(0) = \Lambda_h(\psi)(0)$ by (128), (130), and a simple recursion argument, we see that (127) holds. □

We can now prove the induction step.

Lemma 15. *Let $2 \le n \le N$ and suppose that Statement 1 holds. Then Statement 1 holds with $n+1$ in place of n.*

Because the proof of Lemma 15 is rather long, we first provide a brief outline of the argument. The proof is divided into four sections (whose titles are indicated using bold italic font). In the first part, we partition the interval $[0, \theta_{n+1})$ into intervals of the form $[t_k, t_{k+1})$, $0 \le k < K$, for some $K \in \mathbb{N}_\infty$, where the time t_k is chosen such that $|\mathscr{I}(h(t_k))| = n$ for $1 \le k < K$. In the second part, we construct each time change λ_l on the intervals $[t_k, t_{k+1})$, $0 \le k < K$ and use the induction hypothesis and time-shift properties of the ESP (Lemma 3) and derivative problem (Lemma 4) to prove the time changes satisfy the desired convergence properties. In the third part, we recursively prove convergence of the time changes and the time-changed paths on intervals of the form $[0, t_k)$ for $1 \le k < K+1$. The proof relies on the induction hypothesis, the fact that $\Lambda_h(\psi)(\cdot)$ is continuous at t_k (by Proposition 4) and the Lipschitz continuity of the derivative map (Proposition 2). In the fourth part, we extend the convergence of the time changes and time-changed paths to the interval $[0, \theta_{n+1})$.

Proof of Lemma 15. Let f, (h, g), $\{f_l\}_{l\in\mathbb{N}}$, $\{(h_l, g_l)\}_{l\in\mathbb{N}}$, and ψ be as in Statement 1. If $\theta_{n+1} = \theta_n$, then (104) and (105) hold by assumption. For the remainder of the proof. we assume that $\theta_{n+1} > \theta_n$.

Partition of the Interval $[0, \theta_{n+1})$ into Subintervals $\{[t_k, t_{k+1})\}_{1\le k<K}$. Set $t_0 := 0$, $t_1 := \theta_n$ and, for $k \ge 1$ such that $t_k < \theta_{n+1}$, define ρ_k to be the first time after t_k that either ψ jumps or h hits a face that does not lie in the subset of faces that $h(t_k)$ lies in; that is,

$$\rho_k := \inf\{t > t_k : \psi(t) \ne \psi(t-)\} \wedge \inf\{t > t_k : \mathscr{I}(h(t)) \not\subseteq \mathscr{I}(h(t_k))\} \tag{131}$$

and let t_k be the first time after ρ_k that h lies at the intersection of n or more faces; that is,

$$t_{k+1} := \inf\{t \ge \rho_k : |\mathscr{I}(h(t))| \ge n\}, \tag{132}$$

where we recall that the infimum over the empty set is taken to be infinity. We claim that

$$|\mathscr{I}(h(\rho_k))| \le 1. \tag{133}$$

For a proof by contradiction, suppose $|\mathscr{I}(h(\rho_k))| \ge 2$. If $\psi(\rho_k) \ne \psi(\rho_k-)$, then the definition of $\mathbb{D}_s^h(\mathbb{R}^J)$ given in (96) implies that $h(\rho_k) \in G^\circ$, and so $|\mathscr{I}(h(\rho_k))| = 0$, which leads to a contradiction. On the other hand, if $\psi(\cdot)$ is continuous at ρ_k, then it is clear from (131) that there exists $i_k \in \mathscr{I}(h(\rho_k))$ such that $i_k \notin \mathscr{I}(h(t_k))$. By condition 3 of the boundary jitter property, h must hit the interior of each face F_i, $i \in \mathscr{I}(h(\rho_k))$, infinitely often immediately prior to ρ_k. In particular, there exists $s \in (t_k, \rho_k)$ such that $h(s) \in F_{i_k}$. However, because $i_k \notin \mathscr{I}(h(t_k))$ and $s < \rho_k$, this contradicts definition (131) for ρ_k. In either case, we obtain a contradiction. Therefore, (133) holds. It then follows from the definition of $\mathbb{D}_s^h(\mathbb{R}^J)$ given in (96), (133), the continuity of h, and the upper semicontinuity property of $\mathscr{I}(\cdot)$ shown in Lemma 14 that

$$t_k < \rho_k < t_{k+1}. \tag{134}$$

If $t_k = \theta_{n+1}$ for some $k \in \mathbb{N}$, then set $K := k$ and end the sequence. Alternatively, if $t_k < \theta_{n+1}$ for all $k \in \mathbb{N}$, define $K := \infty$. In this case, we claim that $t_k \to \theta_{n+1}$ as $k \to \infty$. The claim clearly holds if $t_k \to \infty$ as $k \to \infty$, and on the other hand, if t_k is uniformly bounded for all $k \in \mathbb{N}$, then there must be an accumulation point, and it follows from (131) and (132) and the fact that ψ has finitely many discontinuities that the accumulation point must be θ_{n+1}. In either case, the following holds:

$$\text{Given } T<\theta_{n+1}, \text{ there exists } 1\le k<K+1 \text{ such that } T<t_k. \tag{135}$$

Definition of the Time-Changes λ_l on $[0,\theta_{n+1})$. We use the induction hypothesis and time-shift properties of the ESP (Lemma 3) and derivative problem to define the time changes $\{\lambda_l\}_{l\in\mathbb{N}}$ on intervals of the form $[t_k,t_{k+1})$ for $0\le k<K$. We begin with the interval $[0,t_1)$. Because $t_1=\theta_n$ by definition and Statement 1 holds by the induction hypothesis, there is a sequence $\{\lambda_l^1\}_{l\in\mathbb{N}}$ of time changes on $[0,t_1)$ such that for all $T<\infty$,

$$\lim_{l\to\infty}\sup_{t\in[0,t_1\wedge T)}|\lambda_l^1(t)-t|=0, \tag{136}$$

and for all $T<t_1$,

$$\lim_{l\to\infty}\sup_{t\in[0,T]}\left|\Lambda_{h_l}(\psi)(\lambda_l^1(t))-\Lambda_h(\psi)(t)\right|=0. \tag{137}$$

For each $l\in\mathbb{N}$, define λ_l on the interval $[0,t_1)$ by

$$\lambda_l(t):=\lambda_l^1(t)\qquad \text{for } t\in[0,t_1). \tag{138}$$

Now suppose $1\le k<K$. Because h spends zero Lebesgue measure on the boundary ∂G because of condition 2′ of the boundary jitter property (Definition 3), it follows from (134), (133), and the upper semicontinuity of $\mathcal{I}(\cdot)$ that there exists

$$t_k<\xi_k<\rho_k \tag{139}$$

such that

$$h(\xi_k)\in G^\circ \qquad \text{and} \qquad |\mathcal{I}(h(t))|<n \text{ for all } t\in[\xi_k,\rho_k). \tag{140}$$

Define

$$h^k(\cdot):=h(\xi_k+\cdot),\qquad g^k(\cdot):=g(\xi_k+\cdot)-g(\xi_k),\qquad f^k(\cdot):=h(\xi_k)+f(\xi_k+\cdot)-f(\xi_k). \tag{141}$$

By the time-shift property of the ESP, (h^k,g^k) is a solution to the ESP for f^k. In addition, because (h,g) satisfies the boundary jitter property and $h(\xi_k)\in G^\circ$ by (140), it follows from (40) that

$$(h^k,g^k) \text{ satisfies the boundary jitter property.} \tag{142}$$

Finally, by (40), (140), and (132),

$$\begin{aligned}\theta_n^k&:=\inf\{t\ge 0:|\mathcal{I}(h^{\xi_k}(t))|\ge n\}\\&=\inf\{t\ge\xi_k:|\mathcal{I}(h(t))|\ge n\}-\xi_k\\&=t_{k+1}-\xi_k.\end{aligned} \tag{143}$$

For each $l\in\mathbb{N}$, define h_l^k, g_l^k, and f_l^k as follows:

$$h_l^k(\cdot):=h_l(\xi_k+\cdot),\qquad g_l^k(\cdot):=g_l(\xi_k+\cdot)-g_l(\xi_k),\qquad f_l^k(\cdot):=h_l(\xi_k)+f_l(\xi_k+\cdot)-f_l(\xi_k). \tag{144}$$

Again, by the time-shift property of the ESP, (h_l^k,g_l^k) is a solution to the ESP for f_l^k. Because f_l converges to f uniformly on compact intervals as $l\to\infty$, it follows that

$$f_l^k \text{ converges to } f^k \text{ uniformly on compact intervals as } l\to\infty. \tag{145}$$

Therefore, by the Lipschitz continuity of the ESM [Proposition 1(ii)],

$$(h_l^k,g_l^k) \text{ converges to } (h^k,g^k) \text{ uniformly on compact intervals as } l\to\infty. \tag{146}$$

Define

$$\psi^k(\cdot):=\Lambda_h(\psi)(\xi_k)+\psi(\xi_k+\cdot)-\psi(\xi_k). \tag{147}$$

By the time-shift property of the derivative problem (Lemma 4),

$$\Lambda_{h^k}(\psi^k)(\cdot) = \Lambda_h(\psi)(\xi_k + \cdot). \tag{148}$$

The definitions of ψ^k and h^k, the fact that $h(\xi_k) \in G^\circ$ so $\mathbb{H}_{h(\xi_k)} = \mathbb{R}^J$, the fact that $\psi \in \mathbb{D}_s^h(\mathbb{R}^J)$ and (96) imply that

$$\psi^k \in \mathbb{D}_s^{h^k}(\mathbb{R}^J). \tag{149}$$

By (142), (145), (149), and the fact that Statement 1 holds by assumption (with h^k, g^k, f^k, h_l^k, g_l^k, f_l^k, ψ^k, and θ_n^k in place of h, g, f, h_l, g_l, f_l, ψ, and θ_n, respectively), there is a sequence $\{\lambda_l^k\}_{l\in\mathbb{N}}$ of time changes on $[0,\theta_n^k)$ such that, for all $T<\infty$,

$$\lim_{l\to\infty} \sup_{t\in[0,\theta_n^k\wedge T)} |\lambda_l^k(t) - t| = 0, \tag{150}$$

and for all $T<\theta_n^k$,

$$\lim_{l\to\infty} \sup_{t\in[0,T]} \left|\Lambda_{h_l^k}(\psi^k(\lambda_l^k(t)) - \Lambda_{h^k}(\psi^k)(t)\right| = 0. \tag{151}$$

For each $l\in\mathbb{N}$, define λ_l on the interval $[t_k, t_{k+1}) = [t_k, \xi_k + \theta_n^k)$ [because of (143)] by

$$\lambda_l(t) := \begin{cases} t & \text{for } t\in[t_k,\xi_k) \\ \lambda_l^k(t-\xi_k) & \text{for } t\in[\xi_k, t_{k+1}). \end{cases} \tag{152}$$

Convergence of the Time-Changed Paths on Intervals of the Form $[0,t_k)$. We use the principle of mathematical induction to show that, for each $1\le k<K+1$ and for all $T<\infty$,

$$\lim_{l\to\infty} \sup_{t\in[0,t_k\wedge T)} |\lambda_l(t) - t| = 0, \tag{153}$$

and for all $T<t_k$,

$$\lim_{l\to\infty} \sup_{t\in[0,T]} |\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)| = 0. \tag{154}$$

The base case $k=1$ follows immediately from (136) and (137). Now suppose $1\le k<K+1$ and (153) holds for all $T<\infty$ and (154) holds for all $T<t_k$. We first prove (153) holds for all $T<\infty$ with $k+1$ in place of k. Let $T<\infty$. We have

$$\sup_{t\in[0,t_{k+1}\wedge T)} |\lambda_l(t)-t| = \sup_{t\in[0,t_k\wedge T)} |\lambda_l(t)-t| \vee \sup_{t\in[\xi_k\wedge T, t_{k+1}\wedge T)} |\lambda_l(t)-t|. \tag{155}$$

By (152) and (143), we have

$$\sup_{t\in[\xi_k\wedge T, t_{k+1}\wedge T)} |\lambda_l(t)-t| \le \sup_{t\in[0,(\theta_n^k\wedge(T-\xi_k))\vee 0)} |\lambda_l^k(t)-t|. \tag{156}$$

By (155), our assumption that (153) holds, (156), and (150), we see that (153) holds with $k+1$ in place of k.

Next, we prove (154) holds for all $T<t_{k+1}$. If $T<t_k$, then (154) follows by assumption. Let $t_k\le T<t_{k+1}$. We have

$$\sup_{t\in[0,T]} \left|\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)\right| \le \sup_{t\in[0,\xi_k\wedge T]} \left|\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)\right| + \sup_{t\in[\xi_k\wedge T, T]} |\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)|. \tag{157}$$

We first show that

$$\lim_{l\to\infty} \sup_{t\in[0,\xi_k\wedge T]} \left|\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)\right| = 0. \tag{158}$$

Because $|\mathcal{I}(h(t_k))|\ge 2$ by (132), Proposition 4 implies that $\Lambda_h(\psi)(\cdot)$ is continuous at t_k. Let $\varepsilon>0$. By (139) and the continuity of $\Lambda_h(\psi)(\cdot)$ at t_k, we can choose $\delta>0$ sufficiently small such that

$$0 < t_k - \delta < t_k + \delta < \xi_k < \rho_k - \delta \tag{159}$$

and

$$|\Lambda_h(\psi)(t)-\Lambda_h(\psi)(t_k)|<\varepsilon \qquad \text{for all } t\in[t_k-\delta,t_k+\delta]. \tag{160}$$

By the definition of $\mathbb{D}_s^h(\mathbb{R}^J)$ in (96) and the fact that $h(t_k)\in\partial G$, we can assume that $\delta>0$ is sufficiently small so that $\psi(\cdot)$ is constant on $[t_k-2\delta,t_k]$. Along with the definition of ρ_k in (131), this implies that

$$\psi(\cdot) \text{ is constant on } [t_k-2\delta,\rho_k). \tag{161}$$

Because of the upper semicontinuity of $\mathcal{I}(\cdot)$ shown in Lemma 14 and the definition of ρ_k in (131), we can choose $\delta>0$ possibly even smaller to ensure that

$$\mathcal{I}(h(t))\subseteq\mathcal{I}(h(t_k)) \qquad \text{for all } t\in[t_k-\delta,\rho_k). \tag{162}$$

It follows from (154) (which holds for $T<t_k$ by assumption) that

$$\lim_{l\to\infty}\sup_{t\in[0,t_k-\delta]}\left|\Lambda_{h_l}(\psi)(\lambda_l(t))-\Lambda_h(\psi)(t)\right|=0. \tag{163}$$

Then by (163) and (160), we can choose $l_0\in\mathbb{N}$ sufficiently large such that for all $l\ge l_0$,

$$\begin{aligned}|\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta))-\Lambda_h(\psi)(t_k)|&\le|\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta))-\Lambda_h(\psi)(t_k-\delta)|+|\Lambda_h(\psi)(t_k-\delta)-\Lambda_h(\psi)(t_k)|\\&<2\varepsilon.\end{aligned} \tag{164}$$

By (146), the fact that (153) holds with $k+1$ in place of k, and the relation $\rho_k<t_{k+1}$ stated in (134), we see that

$$\lim_{l\to\infty}\sup_{t\in[0,\rho_k-\delta]}|h_l(\lambda_l(t))-h(t)|=0.$$

Then, by (153), the last display, (162), and the upper semicontinuity of $\mathcal{I}(\cdot)$, we can choose $l_0\in\mathbb{N}$ possibly larger such that for all $l\ge l_0$,

$$\lambda_l(t_k-\delta)\ge t_k-2\delta, \tag{165}$$

and

$$\mathcal{I}(h_l(\lambda_l(t)))\subseteq\mathcal{I}(h(t_k)) \qquad \text{for all } t\in[t_k-\delta,\rho_k-\delta]. \tag{166}$$

By condition 2 of the derivative problem, (166), and (159), $\Lambda_h(\psi)(t_k)\in\mathbb{H}_{h(t_k)}\subseteq\mathbb{H}_{h_l(\lambda_l(t))}$ for all $t\in[t_k-\delta,\xi_k]$. A straightforward verification then shows that the constant pair of functions $(\widetilde{\phi}(\cdot),\widetilde{\eta}(\cdot))\equiv(\Lambda_h(\psi)(t_k),0)$ is the unique solution to the derivative problem along $\widetilde{h}_l(\cdot):=h_l(\lambda_l(t_k-\delta+\cdot))$ for the constant function $\widetilde{\psi}(\cdot)\equiv\Lambda_h(\psi)(t_k)$ on the interval $[0,\xi_k-t_k+\delta]$. By the time-shift property of the derivative problem stated in Lemma 4 (with $t_k-\delta$ in place of s), (165), and (161), it holds that, for each $l\ge l_0$, the pair of functions $(\widehat{\phi}(\cdot),\widehat{\eta}(\cdot)):=(\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta+\cdot)),\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta+\cdot))-\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta)))$ is the unique solution to the derivative problem along $h_l(\lambda_l(t_k-\delta+\cdot))$ for the constant function $\widehat{\psi}(\cdot)\equiv\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta))$ on the interval $[0,\xi_k-t_k+\delta]$. Therefore, by the Lipschitz continuity of the derivative map (Proposition 2) and (164),

$$\begin{aligned}\sup_{t\in[t_k-\delta,\xi_k]}|\Lambda_{h_l}(\psi)(\lambda_l(t))-\Lambda_h(\psi)(t)|&=\sup_{t\in[0,\xi_k-t_k+\delta]}|\widehat{\phi}(t)-\widetilde{\phi}(t)|\\&\le\kappa_\Lambda(\alpha)|\Lambda_{h_l}(\psi)(\lambda_l(t_k-\delta))-\Lambda_h(\psi)(t_k)|\\&<2\kappa_\Lambda(\alpha)\varepsilon.\end{aligned}$$

Because $\varepsilon>0$ was arbitrary, the last display and (163) together imply that (158) holds.

Next, we show that

$$\lim_{l\to\infty}\sup_{t\in[\xi_k\wedge T,T]}\left|\Lambda_{h_l}(\psi)(\lambda_l(t))-\Lambda_h(\psi)(t)\right|=0. \tag{167}$$

Define h^k, g^k, f^k and in (141), h_l^k, g_l^k, and f_l^k as in (144) and ψ^k as in (147). For each $l\in\mathbb{N}$, define $\psi_l^k\in\mathbb{D}_s(\mathbb{R}^J)$ by

$$\psi_l^k(\cdot):=\Lambda_{h_l}(\psi)(\xi_k)+\psi(\xi_k+\cdot)-\psi(\xi_k). \tag{168}$$

Then by the time-shift property of the derivative problem (Lemma 4),

$$\Lambda_{h_l^k}(\psi_l^k)(\cdot) = \Lambda_{h_l}(\psi)(\xi_k + \cdot). \tag{169}$$

Thus, by (169), (148), and the triangle inequality,

$$\begin{aligned}\sup_{t\in[\xi_k\wedge T,T]}\left|\Lambda_{h_l}(\psi)(\lambda_l(t)) - \Lambda_h(\psi)(t)\right| &= \sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi_l^k)(\lambda_l^k(t)) - \Lambda_{h^k}(\psi^k)(t)\right| \\ &\le \sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi_l^k)(\lambda_l^k(t)) - \Lambda_{h_l^k}(\psi^k)(\lambda_l^k(t))\right| \\ &\quad + \sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi^k)(\lambda_l^k(t)) - \Lambda_{h^k}(\psi^k)(t)\right|.\end{aligned} \tag{170}$$

By the Lipschitz continuity of the derivative map (Proposition 2), the definition of ψ_l^k in (168), and the definition of ψ^k in (147), we have

$$\begin{aligned}\sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi_l^k)(\lambda_l^k(t)) - \Lambda_{h_l^k}(\psi^k)(\lambda_l^k(t))\right| &\le \kappa_\Lambda(\alpha)\sup_{t\in[0,\lambda_l^k((T-\xi_k)\vee 0)]}\left|\psi_l^k(t) - \psi^k(t)\right| \\ &\le \kappa_\Lambda(\alpha)\left|\Lambda_{h_l}(\psi)(\xi_k) - \Lambda_h(\psi)(\xi_k)\right|.\end{aligned}$$

Then, by (158) and the fact that $\lambda_l(\xi_k) = \xi_k$ because of (152), letting $l\to\infty$ yields

$$\lim_{l\to\infty}\sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi_l^k)(\lambda_l^k(t)) - \Lambda_{h_l^k}(\psi^k)(\lambda_l^k(t))\right| = 0. \tag{171}$$

Because $(T-\xi_k)\vee 0 < \theta_n^k$, it follows from (151) that

$$\lim_{l\to\infty}\sup_{t\in[0,(T-\xi_k)\vee 0]}\left|\Lambda_{h_l^k}(\psi^k)(\lambda_l^k(t)) - \Lambda_{h^k}(\psi^k)(t)\right| = 0. \tag{172}$$

Together, (170)–(172) imply that (167) holds. Because (158) and (167) both hold, we see that (154) holds with $k+1$ in place of k. With the induction step established, the principle of mathematical induction implies that, for each $1\le k<K+1$, (153) holds for all $T<\infty$ and (154) holds for all $T<t_k$.

Convergence of the Time-Changed Paths on $[0,\theta_{n+1})$**.** We first prove that (104) holds for all $T<\infty$ with θ_{n+1} in place of θ_n. Let $T<\infty$. Suppose $T<\theta_{n+1}$. Then there exists $1\le k<K+1$ such that $T<t_k$, so by (153),

$$\lim_{l\to\infty}\sup_{t\in[0,T)}|\lambda_l(t)-t| \le \lim_{l\to\infty}\sup_{t\in[0,t_k)}|\lambda_l(t)-t| = 0. \tag{173}$$

Now suppose $T\ge\theta_{n+1}$. Let $\varepsilon\in(0,\theta_n)$ be arbitrary. Because (173) holds for all $T<\theta_{n+1}$, we have

$$\lim_{l\to\infty}\sup_{t\in[0,\theta_{n+1}-\varepsilon)}|\lambda_l(t)-t| = 0. \tag{174}$$

Thus, we are left to show that

$$\lim_{\varepsilon\downarrow 0}\lim_{l\to\infty}\sup_{t\in[\theta_{n+1}-\varepsilon,\theta_{n+1})}|\lambda_l(t)-t| = 0. \tag{175}$$

By the triangle inequality and the fact that λ_l is nondecreasing, for all $\varepsilon\in(0,\theta_{n+1})$ and $t\in[\theta_{n+1}-\varepsilon,\theta_{n+1})$,

$$|\lambda_l(t)-t| \le |\lambda_l(t)-\theta_{n+1}| + |\theta_{n+1}-t| \le |\lambda_l(\theta_{n+1}-\varepsilon)-\theta_{n+1}| + \varepsilon. \tag{176}$$

By (174) and the continuity of λ_l, we have $\lambda_l(\theta_{n+1}-\varepsilon)\to\theta_{n+1}-\varepsilon$ as $l\to\infty$. Thus, (175) holds, which implies that (104) holds with θ_{n+1} in place of θ_n. We are left with the final step of proving that (105) holds for all $T<\theta_{n+1}$. Let $T<\theta_{n+1}$. By (135), there exists $1\le k<K+1$ such that $T<t_k$. It then follows from (154) that (105) holds. □

Proof of Lemma 11. By Lemmas 13 and 15 and the principle of mathematical induction, Statement 1 holds for $n=2,\dots,N+1$. Because $\theta_{N+1}=\infty$ by definition, the lemma follows. □

Acknowledgments
The authors thank Mike Giles for useful feedback on the numerical examples in a previous version of this paper. They also thank the referees for a careful reading of this paper and for their helpful comments and corrections. In particular, their comments led to the addition of the appendix.

Appendix. Explicit Expression for the Projection Maps and Derivative Projection Matrices
In this section, we obtain explicit expressions under Assumptions 1–3, for the projection map π^α (see Assumption 3), the map ξ^α (see Lemma 1), and the derivative projection matrices $\mathcal{L}_x^\alpha$ (see Lemma 2) associated with the data $\{(d_i(\alpha), n_i, c_i), i \in \mathcal{I}\}$.

We fix $\alpha \in U$ and drop the "α" notation. Let $\mathbb{I} := \{\mathcal{I}(x) : x \in \partial G\}$. For $I \in \mathbb{I}$ let

$$F_I := \{x \in \partial G : \mathcal{I}(x) = I\}, \tag{A.1}$$
$$d_I := \text{cone}\{d_i, i \in I\}, \tag{A.2}$$
$$\mathbb{A}_I := \pi^{-1}(F_I) := \{x \in \mathbb{R}^J : \pi(x) \in F_I\}. \tag{A.3}$$

Lemma A.1. *The subsets $\mathbb{A}_I$, $I \in \mathbb{I}$ are disjoint, and $\mathbb{R}^J \setminus G^\circ = \cup_{I\in\mathbb{I}} \mathbb{A}_I$. In addition, for $I \in \mathbb{I}$,*

$$\mathbb{A}_I = \{z - u : z \in F_I, u \in d_I\}. \tag{A.4}$$

Proof. From the definition of $\mathbb{I}$ we see that $\{F_I, I \in \mathbb{I}\}$ partitions ∂G into disjoint subsets. By Assumption 3 and the fact that $d(x) = \{0\}$ for all $x \in G^\circ$, it follows that $\pi(x) \in \partial G$ for all $x \in \mathbb{R}^J \setminus G^\circ$, and so

$$\pi^{-1}(\partial G) := \{x \in \mathbb{R}^J : \pi(x) \in \partial G\} = \mathbb{R}^J \setminus G^\circ.$$

Along with (A.3), this implies that $\{\mathbb{A}_I, I \in \mathbb{I}\}$ partitions $\mathbb{R}^J \setminus G^\circ$ into disjoint subsets. Let $I \in \mathbb{I}$. Suppose $x \in \mathbb{A}_I$. By (A.3), Assumption 3, and (A.2), $z := \pi(x) \in F_I$, $u := \pi(x) - x \in d_I$, and $x = z - u$. Alternatively, suppose $z \in F_I$, $u \in d_I$, and $x := z - u$. It follows from the uniqueness of π (see Remark 3) that $\pi(z - u) = z \in F_I$. This proves that (A.4) holds. □

For $I \in \mathbb{I}$, define

$$\mathbb{H}_I := \{y \in \mathbb{R}^J : \langle y, n_i \rangle = 0 \ \forall \ i \in I\}.$$

The next lemma follows immediately from lemmas 8.1 and 8.2 of Lipshutz and Ramanan (2018) and Assumption 1.

Lemma A.2. *For each $I \in \mathbb{I}$, $\mathbb{H}_I \cap span(d_I) = \{0\}$, and for each $y \in \mathbb{R}^J$, there exists a unique pair of vectors $v \in \mathbb{H}_I$ and $w \in span(d_I)$ such that $y = v + w$.*

Let $Q := (n_1 \quad \cdots \quad n_N)$ denote the $J \times N$ matrix with column vectors $\{n_i, i \in \mathcal{I}\}$. For $I \in \mathbb{I}$, let S_I denote the $N \times |I|$ matrix whose (i, j)th entry equals one if i is the jth element of I and zero otherwise for $i \in \mathcal{I}$ and $j \in I$. Define $R_I := RS_I$ and $Q_I := QS_I$ to be the $J \times |I|$ matrices with column vectors $\{d_i, i \in I\}$ and $\{n_i, i \in I\}$, respectively. By Assumption 1 and Lemma A.2, the null space of $Q_I^T R_I$ is equal to zero, and so $Q_I^T R_I$ is invertible. Define

$$V_I := -S_I(Q_I^T R_I)^{-1} Q_I^T \in \mathbb{R}^{N\times J}, \tag{A.5}$$

and

$$P_I := E_J + RV_I \in \mathbb{R}^{J\times J}, \tag{A.6}$$

where we recall that E_J denotes the $J \times J$ identity matrix. Note that, if $N = J$ and $\mathcal{I} \in \mathbb{I}$, then $V_\mathcal{I} = -R^{-1}$ and $P_\mathcal{I} = 0$.

We derive some properties of the matrices V_I and P_I. By the definitions of $\mathbb{H}_I$ and Q_I, we have $Q_I^T v = 0$ for all $v \in \mathbb{H}_I$. Therefore, (A.5) and (A.6) imply that

$$V_I v = 0, \qquad v \in \mathbb{H}_I, \tag{A.7}$$
$$P_I v = v, \qquad v \in \mathbb{H}_I. \tag{A.8}$$

In addition, if $w \in \text{span}(d_I)$, then there exists $r \in \mathbb{R}^{|I|}$ such that $w = R_I r$. Thus, by (A.5) and the fact that $R_I = RS_I$,

$$RV_I w = -R_I(Q_I^T R_I)^{-1} Q_I^T R_I r = -w.$$

It follows from (A.6) that

$$P_I w = 0, \qquad w \in \text{span}(d_I). \tag{A.9}$$

Let $c = (c_1 \quad \cdots \quad c_N)^T \in \mathbb{R}^N$ denote the column vector whose ith component is equal to c_i for $i \in \mathcal{I}$, and define

$$q_I := S_I(Q_I^T R_I)^{-1} S_I^T c \in \mathbb{R}^N. \tag{A.10}$$

Given $x \in F_I$, it follows from the fact that $\langle x, n_i \rangle = c_i$ for all $i \in I$ that $Q_I x = S_I c$. Thus, by (A.5),

$$V_I x = -q_I, \qquad x \in F_I. \tag{A.11}$$

Lemma A.3. *The projection map $\pi : \mathbb{R}^J \mapsto G$ satisfies*

$$\pi(x) = \begin{cases} x & \text{if } x \in G^\circ, \\ P_I x + R q_I & \text{if } x \in \mathbb{A}_I,\ I \in \mathbb{I}, \end{cases} \tag{A.12}$$

and the map $\xi : \mathbb{R}^J \to \mathbb{R}^N_+$ satisfies

$$\xi(x) = \begin{cases} 0 & \text{if } x \in G^\circ, \\ V_I x + q_I & \text{if } x \in \mathbb{A}_I,\ I \in \mathbb{I}. \end{cases} \tag{A.13}$$

In addition, for each $x \in \partial G$, the derivative projection matrix $\mathcal{L}_x \in \mathbb{R}^{J \times J}$ satisfies $\mathcal{L}_x = P_{\mathcal{I}(x)}$. Furthermore, for each $I \in \mathbb{I}$,

$$\mathbb{A}_I = \{x \in \mathbb{R}^J : P_I x + R q_I \in F_I \text{ and } V_I x + q_I \in \mathbb{R}^N_+\}. \tag{A.14}$$

Remark A.1. Lemma A.3 provides explicit expressions for the projection map and derivative projection matrices. However, implementation of the method requires checking if $x \in \mathbb{A}_I$ for each $I \in \mathbb{I}$, and $|\mathbb{I}|$ grows exponentially with the number of faces. For certain examples of reflected diffusions in domains with a large number of faces, more efficient expressions may be available (e.g., by exploiting symmetry properties of the directions of reflection); however, such an approach is specific to the problem of interest. In addition, for many reflected diffusions of interest, a.s. the reflected diffusion does not hit the intersection of of two or more faces—for example, the reflected diffusion studied in Section 5.2 (see, e.g., Ichiba and Karatzas 2010)—or the constraining process L^α is constant when the reflected diffusion Z^α is at the intersection of two or more faces (see, e.g., theorem 1 of Reiman and Williams 1988). That being said, with positive probability, the discrete approximation $L^{\alpha,\Delta}$ increases when $Z^{\alpha,\Delta}$ lies in the intersection of two or more faces; however, this probability converges to zero with the discretization parameter. Thus, when implementing the expression for the projection map and derivative projection matrices, an efficient algorithm will first check if $x \in \mathbb{A}_I$ for $I \in \{\emptyset, 1, \ldots, N\}$ before checking if $x \in \mathbb{A}_I$ for $|I| \geq 2$.

Proof. Fix $I \in \mathbb{I}$. Let $x \in \mathbb{A}_I$. We first prove $\pi(x) = P_I x + R q_I$. By (A.4) there exists $z_x \in F_I$ and $u_x \in d_I$ such that $x = z_x - u_x$. Then, by (A.6), (A.9), and (A.11),

$$P_I x + R q_I = (E_J + R V_I) z_x + R q_I = z_x \in F_I,$$

and so

$$P_I x + R q_I - x = u_x \in d_I. \tag{A.15}$$

By the uniqueness of the projection map π, this proves that $\pi(x) = P_I x + R q_I$. Next, we show that $\xi(x) = V_I x + q_I$. By Lemma 1, (A.12), and (A.15),

$$R\xi(x) = \pi(x) - x = P_I x + R q_I - x = u_x.$$

Because $u_x \in d_I$, this implies that $V_I x + q_I \in \mathbb{R}^N_+$ and $[V_I x + q_I]^i > 0$ only if $i \in I$. Therefore, by the uniqueness of ξ (see Lemma 1), $\xi(x) = V_I x + q_I$ holds. To show that $\mathcal{L}_x = P_{\mathcal{I}(x)}$, suppose $y \in \mathbb{R}^J$. By Lemma A.2, we can choose $v_y \in \mathbb{H}_I = \mathbb{H}_x$ and $w_y \in \text{span}(d_I) = \text{span}(d(x))$ such that $y = v_y + w_y$. It follows from (A.6)–(A.9) that $P_I y = P_I v_y + P_I w_y = v_y$ and $P_I y - y = -w_y \in \text{span}(d_I)$. By the uniqueness of the derivative projection matrix, this proves $\mathcal{L}_x y = P_{\mathcal{I}(x)} y$.

We are left to prove (A.14). Suppose $x \in \mathbb{A}_I$. By (A.12) and (A.3), $P_I x + R q_I = \pi(x) \in F_I$. In addition, by (A.6), (A.12), and Assumption 3, $R(V_I x + q_I) = \pi(x) - x \in d_I$, and so, in view of Assumption 1 and the fact that $[V_I x + q_I]^i = 0$ for all $i \notin I$ by (A.5) and (A.10), it follows that $V_I x + q_I \in \mathbb{R}^N_+$. On the other hand, suppose $x \in \mathbb{R}^J$ is such that $P_I x + R q_I \in F_I$ and $V_I x + q_I \in \mathbb{R}^N_+$. Then $u_x := R(V_I x + q_I) \in d_I$ and $z_x := x + u_x = P_I x + R q_I \in F_I$. Therefore, $x = z_x - u_x \in \mathbb{A}_I$ by (A.3). □

References

Asmussen S, Glynn PW (2007) *Stochastic Simulation: Algorithms and Analysis* (Springer, New York).

Banner AD, Fernholz R, Karatzas I (2005) Atlas models of equity markets. *Ann. Appl. Probab.* 15(4):2296–2330.

Blanchet J, Murthy KRA (2018) Exact simulation of multidimensional reflected Brownian motion. *J. Appl. Probab.* 55(1):137–156.

Bossy M, Gobet E, Talay D (2004) Symmetrized Euler scheme for an efficient approximation of reflected diffusions. *J. Appl. Probab.* 4(3):877–889.

Chen H, Mandelbaum A (1991) Leontief systems, RBVs and RBMs. Davis MHA, Elliott RJ, eds. *Applied Stochastic Analysis*, Stochastics Monographs, vol. 5 (Gordon and Breach, New York), 1–43.

Costantini C, Pacciarotti B, Sartoretto F (1998) Numerical approximation for functionals of reflecting diffusion processes. *SIAM J. Appl. Math.* 58(1):73–102.

Dupuis P, Ramanan K (1999) Convex duality and the Skorokhod problem. I. *Probab. Theory Related Fields* 115(2):153–195.

Fernholz ER (2002) *Stochastic Portfolio Theory* (Springer, New York).

Glasserman P (2003) *Monte Carlo Methods in Financial Engineering* (Springer, New York).

Gobet E (2001) Efficient schemes for the weak approximation of reflected diffusions. *Monte Carlo Methods Appl.* 7(1–2):193–202.

Harrison JM (2013) *Brownian Models of Performance and Control* (Cambridge University Press, Cambridge, UK).

Ichiba T, Karatzas I (2010) On collisions of Brownian particles. *Ann. Appl. Probab.* 20(3):951–977.

Ichiba T, Karatzas I, Shkolnikov M (2013) Strong solutions of stochastic equations with rank-based coefficients. *Probab. Theory Related Fields* 156(1–2):229–248.

Ichiba T, Papathanakos V, Banner A, Karatzas I, Fernholz R (2011) Hybrid atlas models. *Ann. Appl. Probab.* 21(2):609–644.

Kang W, Williams RJ (2007) An invariance principle for semimartingale reflecting Brownian motions in domains with piecewise smooth boundaries. *Ann. Appl. Probab.* 17(2):741–779.

Lépingle D (1995) Euler scheme for reflected stochastic differential equations. *Math. Comput. Simulation* 38(1–3):119–126.

Lipshutz D, Ramanan K (2018) On directional derivatives of Skorokhod maps in convex polyhedral domains. *Ann. Appl. Probab.* 28(2):688–750.

Lipshutz D, Ramanan K (2019a) Pathwise differentiability of reflected diffusions in convex polyhedral domains. *Annales de l'Institut Henri Poincare Probab. Statist.* Forthcoming.

Lipshutz D, Ramanan K (2019b) Sensitivity analysis of the stationary distribution of reflected Brownian motion in a convex polyhedral cone. Working paper, Technion, Haifa, Israel.

Liu Y (1993) Numerical approches to stochastic differential equations with boundary conditions. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Mandelbaum A, Pats G (1998) State-dependent stochastic networks. Part I. Approximations and applications with continuous diffusion limits. *Ann. Appl. Probab.* 8(2):569–646.

Milshtein GN (1995) The solving of boundary value problems by numerical integration of stochastic equations. *Math. Comput. Simulation* 38(1–3):77–85.

Peterson WP (1991) A heavy traffic limit theorem for networks of queues with multiple customer types. *Math. Oper. Res.* 16(1):90–118.

Pettersson R (1995) Approximations for stochastic differential equations with reflecting convex boundaries. *Stochastic Processes Appl.* 59(2): 295–308.

Pettersson R (1997) Penalization schemes for reflecting stochastic differential equations. *Bernoulli* 3(4):403–414.

Ramanan K (2006) Reflected diffusions defined via the extended Skorokhod map. *Electronic J. Probab.* 11(36):934–992.

Ramanan K, Reiman MI (2003) Fluid and heavy traffic diffusion limits for a generalized processor sharing model. *Ann. Appl. Probab.* 13(1): 100–139.

Ramanan K, Reiman MI (2008) The heavy traffic limit of an unbalanced generalized processor sharing model. *Ann. Appl. Probab.* 18(1):22–58.

Reiman MI (1984) Open queueing networks in heavy traffic. *Math. Oper. Res.* 9(3):441–458.

Reiman MI, Williams RJ (1988) A boundary property of semimartingale reflecting Brownian motions. *Probab. Theory Related Fields* 77(1):87–97.

Rogers LCG, Williams D (2000) *Diffusions, Markov Processes and Martingales. Volume 2: Itô Calculus*, 2nd ed. (Cambridge University Press, Cambridge, UK).

Słomínski L (1994) On approximation of solutions of multidimensional SDE's with reflecting boundary conditions. *Stochastic Processes Appl.* 50(2):197–219.

Słomínski L (2001) Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Processes Appl.* 94(2):317–337.

Ward A, Glynn PW (2003) A diffusion approximation for a Markovian queue with reneging. *Queueing Systems* 43(1–2):103–128.

Whitt W (2002) *An Introduction to Stochastic-Process Limits and Their Applications to Queues*. Internet Supplement. Accessed April 30, 2019, http://www.columbia.edu/~ww2040/supplement.html.

Yan L (2002) The Euler scheme with irregular coefficients. *Ann. Probab.* 30(3):1172–1194.

Yang J, Kushner HJ (1991) A Monte Carlo method for sensitivity analysis and parametric optimization of nonlinear stochastic systems. *SIAM J. Control Optim.* 29(5):1216–1249.