Check for updates

J. Appl. Probab. **63,** 623–647 (2026)
*doi:*10.1017/jpr.2025.10050

SPACE-GRID APPROXIMATIONS OF HYBRID STOCHASTIC DIFFERENTIAL EQUATIONS AND FIRST-PASSAGE PROPERTIES

HANSJOERG ALBRECHER,* *University of Lausanne*
OSCAR PERALTA,** *Cornell University*

Abstract

Hybrid stochastic differential equations (SDEs) are a useful tool for modeling continuously varying stochastic systems modulated by a random environment, which may depend on the system state itself. In this paper we establish the pathwise convergence of solutions to hybrid SDEs using space-grid discretizations. Though time-grid discretizations are a classical approach for simulation purposes, our space-grid discretization provides a link with multi-regime Markov-modulated Brownian motions. This connection allows us to explore aspects that have been largely unexplored in the hybrid SDE literature. Specifically, we exploit our convergence result to obtain efficient and computationally tractable approximations for first-passage probabilities and expected occupation times of the solutions to hybrid SDEs. Lastly, we illustrate the effectiveness of the resulting approximations through numerical examples.

Keywords: Pathwise convergence; state-dependent switching; first-passage probability; expected occupation time; Markov-modulated Brownian motion

2020 Mathematics Subject Classification: Primary 60K37
Secondary 60G17; 60H35

1. Introduction

Stochastic systems with modulation have attracted considerable attention in probability, both from applied and theoretical perspectives. Examples of these are Markov-modulated Poisson processes [17], Markov additive processes [4, Chapter XI], and regime-switching stochastic differential equations [33, Section II.2], which have found applications in risk theory [5, Chapter VII], queueing [32], and finance [16], to name a few. The main appeal of this framework comes from the flexibility when modeling phenomena whose parameters depend on random environmental factors. For instance, modulation may be used to model correlated catastrophic events in insurance, different workload regimes in a queue, or seasonal sentiment changes in financial markets. In cases where the modulation is Markovian, a robust toolbox of matrix-analytic methods has been developed to compute relevant quantities; see, e.g., [9] and references therein.

Received 8 September 2025; accepted 8 September 2025.
* Postal address: Faculty of Business and Economics, University of Lausanne, and Swiss Finance Institute, University of Lausanne, Quartier de Chambronne, 1015 Lausanne, Switzerland. Email: hansjoerg.albrecher@unil.ch
** Postal address: School of Operations Research and Information Engineering, Cornell University, Rhodes Hall, Ithaca 14850, New York, United States. Email: op65@cornell.edu

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

In this manuscript we are interested in a class of processes whose dynamics are described by a continuously varying component arising as the solution of a stochastic differential equation, and an environmental finite-state component modeled after a jump process that may or may not be Markovian. Such a class of processes is referred to in the literature as *hybrid stochastic differential equations* (SDEs), which generally take the form

$$\mathrm{d}X(t) = \mu(J(t), X(t))\,\mathrm{d}t + \sigma(J(t), X(t))\,\mathrm{d}B(t),$$

where B is a Brownian motion, J is the environmental process, and μ and σ are real functions satisfying certain regularity conditions.

In particular, we will focus on hybrid SDEs exhibiting the following characteristics:

- The process X is almost surely (a.s.) continuous and one-dimensional.
- At a given instant, the environmental process switches state with an intensity dependent on the main component.

In simpler terms, the class of hybrid SDEs we are interested in has components whose evolution is interlaced and dependent on each other. This contrasts with the Markov-modulated case, where we can draw a path of the environmental component without any knowledge of the main one (see [30]). To highlight this interlacing feature, such models are referred to in the literature as hybrid SDEs with state-dependent switching; here, we will simply call them hybrid SDEs for brevity.

Existence, uniqueness, and stability properties for hybrid SDEs have been extensively studied in recent years [29, 39, 40]. Another stream of research lies in investigating efficient simulation methods for hybrid SDEs, most of which rely on adapting well-known convergence results to this more challenging scenario [30, 38]. However, computing explicit probabilistic descriptors for such a class of processes has proven challenging, even in simple cases. For instance, several descriptors for Markov-modulated Brownian motion have been explicitly obtained in the literature (see, e.g., [3, 11, 14, 24, 31]), but virtually none of these results have been extended to hybrid SDEs, even in the Markovian case. Such a task seems intractable simply because the toolbox for general diffusions is considerably more limited than the one available for Brownian motion.

Our contributions to the literature of hybrid SDEs are the following. First, we analyze a pathwise approximation technique using a space-grid discretization, resulting in approximations that belong to the class of multi-regime Markov-modulated Brownian motions. These processes, when restricted to each band of the space grid, behave like a Markov-modulated Brownian motion. We emphasize that our proposed approximation scheme is novel and does not align with either the Wong–Zakai methods [30] or the classical Euler–Maruyama methods [39, (5.4)], both of which are the most common approximation methods for the simulation of hybrid SDEs in the literature. Instead, our proposed approximation is designed to exploit aspects that extend beyond the practicalities of simulation, focusing on the computation of descriptors that are currently only known for simple cases. As an example, we employ our pathwise approximation result to provide approximations for the descriptors for first-exit times of hybrid SDEs over a band $[0, a]$, $a > 0$, using recent results on the stationary measures of multi-regime Markov-modulated Brownian motion in queueing theory [1, 21]. We remark that, to the best of our knowledge, this is the first attempt to compute first-passage probabilities and expected occupation times for the solutions to hybrid SDEs, even when reduced to the case in which the environmental process is Markovian.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

The structure of this paper is as follows. In Section 2 we establish an appropriate framework for constructing strong solutions to hybrid SDEs. By making slight modifications to the classic construction found in [33, Section II.2.1], we are able to present the construction within a uniformization framework. Later, in Section 3, we examine the proposed multi-regime Markov-modulated Brownian motion approximation and prove its pathwise convergence to the solution of the original hybrid SDE uniformly over increasing compact intervals. In particular, our main result in Theorem 3.2 establishes the weak convergence of both its continuous and discrete components. This approximation result is achieved in two steps: first, by measuring the distance between the continuous components up to the first time the discrete components differ, and then by providing tail estimates until such discrete decoupling occurs. In Section 4, we illustrate how our convergence result can be employed to approximate first-passage probabilities and expected occupation times of hybrid SDEs, and we also explore some numerical examples of these approximations. Lastly, in Section 5, we offer a concise summary of our findings and discuss potential avenues for future research.

2. Hybrid stochastic differential equations

Let us provide a precise description of a hybrid SDE and its solution $(J, X) = \{(J(t), X(t))\}_{t\geq 0}$, initially adhering to the traditional framework as seen in sources such as [39]. For such a construction we require a complete probability space $(\Omega, \mathbb{P}, \mathcal{F})$ that supports the following independent components:

- A standard Brownian motion $B = \{B(t)\}_{t\geq 0}$; this will dictate the continuously varying nature of X.
- A Poisson random measure $\mathfrak{p}(\mathrm{d}t, \mathrm{d}z)$ on $\mathbb{R}_+ \times \mathbb{R}_+$ of intensity $\mathrm{d}t \times \mathrm{d}z$; this will be used to describe the jump dynamics of J.

Then, (J, X) is the solution to the SDE

$$\mathrm{d}X(t) = \mu(J(t), X(t))\,\mathrm{d}t + \sigma(J(t), X(t))\,\mathrm{d}B(t), \qquad X(0) = x_0 \in \mathbb{R}, \tag{2.1}$$

$$\mathrm{d}J(t) = \int_{\mathbb{R}_+} h(J(t-), X(t-), z)\,\mathfrak{p}(\mathrm{d}t, \mathrm{d}z), \qquad J(0) = i_0 \in \mathcal{E}, \tag{2.2}$$

where X is an a.s. continuous real process, J is a càdlàg jump process with finite state space $\mathcal{E} = \{1, \ldots, p\}$, $\mu\colon \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$, $\sigma\colon \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$, and $h\colon \mathcal{E} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{Z}$. Note that the value of h in $\mathbb{Z}$ dictates how many steps J moves (backward or forward) in $\mathcal{E}$ at each time epoch of $\mathfrak{p}$. Furthermore, we require that (J, X) be adapted to the $\mathbb{P}$-completed filtration $\mathcal{F}_t$ generated by $\{B(s), \mathfrak{p}(s, \cdot);\ s \leq t\}$. A pair (J, X) that satisfies the aforementioned characteristics is referred to as a *strong solution* of (2.1)–(2.2). For pedagogical purposes, using a uniformization perspective [35], in this section we will demonstrate the existence of such a solution under simple regularity assumptions in due course.

Here, we will focus on hybrid SDEs with an environmental process J that switches at a state-dependent rate. Specifically, we aim to let J evolve according to

$$\mathbb{P}(J(t + \mathrm{d}t) = j \mid X(t), J(t) = i) = \delta_{ij} + \Lambda_{ij}(X(t))\,\mathrm{d}t \tag{2.3}$$

for a certain family of intensity matrices $\{\mathbf{\Lambda}(x)\}_{x\in\mathbb{R}}$ that are assumed to be càdlàg, and thus measurable. Though (2.2) and (2.3) may initially seem like unrelated definitions for J, we only need to relate them by aligning the jump structure of h in (2.2) with the intensities in (2.3).

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

Our approach to achieving this alignment differs slightly from the method presented in [33]. However, this new construction enables us to translate (2.2) into a uniformization framework, thereby allowing us to utilize the existing toolbox for such methods. In our case, we need the following boundedness condition.

Assumption 2.1. *The family* $\{\mathbf{\Lambda}(x)\}_{x\in\mathbb{R}}$ *is uniformly bounded, i.e.*

$$\gamma := \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\Lambda_{ii}(x)| < \infty.$$

For our construction, we define h by

$$h(i, x, z) = \sum_{j=1}^{p} (j-i) 1_{\{z\in\Delta_{ij}(x)\}}. \tag{2.4}$$

For each fixed i and $x\in\mathbb{R}$, $\{\Delta_{ij}(x) : j\in\mathcal{E}\}$ is a measurable partition of the interval $[0, \gamma)$, where the Lebesgue measure of $\Delta_{ij}(x)$ is given by $\mathrm{Leb}(\Delta_{ij}(x)) = \gamma\delta_{ij} + \Lambda_{ij}(x)$. Note that we ought to ensure that h as specified in (2.4) is a measurable function of its three entries, so that we need to choose the families $\{\Delta_{ij}(x) : j\in\mathcal{E}\}$ carefully. One straightforward way to guarantee this measurability is by defining $\Delta_{ij}(x) = [\ell_{ij}(x), r_{ij}(x))$, where $\ell_{i1}(x) = 0$, $r_{i,j}(x) = \ell_{ij}(x) + (\gamma\delta_{ij} + \Lambda_{ij}(x))$, $\ell_{i,j+1}(x) = r_{ij}(x)$, and $r_{ip}(x) = \gamma$. Unless otherwise stated, this is the standard choice used throughout the text.

With the above choice, all the points of $\mathfrak{p}$ that land outside of $\mathbb{R}_+ \times [0, \gamma)$ do not contribute to jumps in (2.2). In other words, we can rewrite (2.2) as

$$\mathrm{d}J(t) = \int_{[0,\gamma)} h(J(t-), X(t-), z)\, \mathfrak{p}^*(\mathrm{d}t, \mathrm{d}z), \quad J(0) = i_0 \in \mathcal{E}, \tag{2.5}$$

where $\mathfrak{p}^*$ is the Poisson random measure $\mathfrak{p}$ restricted to $\mathbb{R}_+ \times [0, \gamma)$. The advantage of considering $\mathfrak{p}^*$ instead of $\mathfrak{p}$ is that the placement of the atoms of $\mathfrak{p}^*$ may be regarded as a Poisson process $\{\theta_k\}_{k\geq 1}$ of intensity γ over $\mathbb{R}_+$ (the t-axis), with each arrival θ_k carrying an independent mark U_k that is uniformly distributed over the interval $[0, \gamma)$ (the z-axis).

With the aforementioned choice of h and the characterization of $\mathfrak{p}^*$ in terms of $\{(\theta_k, U_k)\}_{k\geq 1}$, we are now ready to explicitly construct the strong solution to (2.1) and (2.2). To facilitate our construction, from now on we assume the following.

Assumption 2.2. *For all* $i\in\mathcal{E}$, $\mu(i, \cdot)$ *and* $\sigma(i, \cdot)$ *are Lipschitz-continuous, i.e. there exists some* $K > 0$ *such that* $|\mu(i, x) - \mu(i, y)| \vee |\sigma(i, x) - \sigma(i, y)| \leq K|x - y|$ *for all* $x, y\in\mathbb{R}$. *Moreover,* $\mu(i, \cdot)$ *and* $\sigma(i, \cdot)$ *have left- and right-derivatives everywhere.*

Under Assumption 2.2, it is standard that a strong solution to the (ordinary) SDE

$$X^{[i,x,v]}(t) = x + \int_0^t \mu\big(i, X^{[i,x,v]}(r)\big)\, \mathrm{d}r + \int_0^t \sigma\big(i, X^{[i,x,v]}(r)\big)\, \mathrm{d}B^{[v]}(r) \tag{2.6}$$

exists for all $i\in\mathcal{E}$, $x\in\mathbb{R}$, and $v, t\geq 0$, where $B^{[v]}(t) := B(v + t) - B(v)$ is a time-shifted version of B. In other words, $X^{[i,x,v]}$ corresponds to the solution of the SDE driven by $B^{[v]}$, with coefficients $\mu(i, \cdot)$ and $\sigma(i, \cdot)$, and starting point x.

Remark 2.1. Although there are more general settings that allow for a strong solution to (2.6), explosions are guaranteed not to occur under Assumption 2.2. This allows us to provide a clearer understanding of the construction of the hybrid SDE. Moreover, for the purpose

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

of investigating first-passage properties out of a space-band, we can exploit the fact that any locally Lipschitz continuous coefficient (a common assumption in the literature) can be localized, and this localized version satisfies Assumption 2.2.

The existence of a pair (J, X) that solves (2.1) and (2.2) (alternatively, (2.1) and (2.5)) is directly derived from [22, Chapter IV.9]. Since its construction is crucial to our convergence developments, we will now offer a concise review for the sake of completeness. Define $X(0) = x_0$ and $J(0) = i_0$, and for $t \in (0, \theta_1)$, let $J(t) = i_0$ and $X(t) = X^{[i_0,x_0,0]}(t)$: we have defined the processes X and J in $[0, \theta_1)$. Since we want X to be continuous, we ought to define $X(\theta_1) = X(\theta_1-)$. According to (2.5), J jumps at time θ_1 to the state $i_0 + h(X(\theta_1), i_0, U_1)$, which is equal to state k if and only if

$$U_1 \in \Delta_{i_0,k}(X(\theta_1)). \tag{2.7}$$

In simple terms, we are using U_1 and $X(\theta_1)$ in such a way that J jumps to state k at time θ_1 with a mass given by $\mathrm{Leb}(\Delta_{i_0,k}(X(\theta_1)))/\gamma$, corresponding to the (i_0, k)th entry of the probability matrix $I + \mathbf{\Lambda}(X(\theta_1))/\gamma$, the uniformized version of $\mathbf{\Lambda}(X(\theta_1))$. After having defined (J, X) in $[0, \theta_1]$, we construct X in subsequent intervals $(\theta_\ell, \theta_{\ell+1}]$ by *concatenating* strong solutions of the type (2.6) with appropriate chosen values for (i, x, v), as well as using a decision rule similar to (2.7) to establish which states J visits. More specifically, in a recursive manner, for $\ell = 1, 2, 3, \ldots$ and $t \in (0, \theta_{\ell+1} - \theta_\ell)$ let

$$\begin{aligned} X(\theta_\ell + t) &= X^{[i_\ell,x_\ell,\theta_\ell]}(t), \\ X(\theta_{\ell+1}) &= X(\theta_{\ell+1}-), \\ J(\theta_\ell + t) &= J(\theta_\ell), \\ J(\theta_{\ell+1}) &= k \quad \text{if and only if } U_{\ell+1} \in \Delta_{i_\ell,k}(X(\theta_{\ell+1})). \end{aligned} \tag{2.8}$$

A few aspects that are straightforward to verify about this construction:

- X and J are $\mathcal{F}_t$-adapted;
- X is continuous at $\theta_1, \theta_2, \ldots$, and thus has a.s. continuous paths;
- X uniquely solves the hybrid SDE (2.1) on every interval $(\theta_\ell, \theta_{\ell+1})$, $\ell = 0, 1, 2, \ldots$;
- J is càdlàg and uniquely solves (2.5).

Thus, (J, X) is indeed the strong solution to (2.2).

We still need to verify that (2.3) holds, which we prove in a slightly more general scenario next. Recall that (2.3) only makes sense if $\mathbf{\Lambda}(x)$ is càdlàg with respect to x. A generalization of this property for discontinuous $\mathbf{\Lambda}(\cdot)$ is given by

$$\begin{aligned} &\mathbb{P}\big(J(t+v) = J(t) \text{ for all } v \in [0, s] \\ &\qquad \mid \mathcal{F}_t \otimes \mathcal{F}^B_{t+s}, X(t) = x, J(t) = i\big) = \exp\left(\int_0^s \Lambda_{ii}\big(X^{[i,x,t]}(v)\big)\,\mathrm{d}v\right), \end{aligned} \tag{2.9}$$

$$\mathbb{P}(J(\theta_\ell) = j \mid J(\theta_\ell-) = i, J(\theta_\ell-) \neq J(\theta_\ell), X(\theta_\ell)) = \frac{\Lambda_{ij}(X(\theta_\ell))}{|\Lambda_{ii}(X(\theta_\ell))|}, \quad i \neq j, \tag{2.10}$$

where $\mathcal{F}^B_t$ is the $\mathbb{P}$-completion of the σ-algebra generated by $\{B(s)\colon 0 \le s \le t\}$, and $\otimes$ denotes the product σ-algebra. In essence, (2.9) and (2.10) correspond to how inhomogeneous Markov jump processes are classically constructed through their integrated jump intensities/hazard

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

rates (see, e.g., [34, Chapter 13]). That (2.9) and (2.10) imply (2.3) for continuous $\mathbf{\Lambda}(\cdot)$ is readily obtained by noting that $\exp\big(\int_0^{\mathrm{d}t}\Lambda_{ii}\big(X^{[i,x,t]}(v)\big)\,\mathrm{d}v\big) = 1+\Lambda_{ii}(x)\,\mathrm{d}t$, where we employed the continuity of $X^{[i,x,t]}$ too. For the sake of completeness, we now show that (2.9) and (2.10) indeed hold.

Lemma 2.1. *Let (J,X) be constructed via (2.8). Then (2.9) and (2.10) hold.*

Proof. Integrating with respect to the number and position of Poisson arrivals of $\mathfrak{p}^*$ in $[t, t+s]$, say $\{\theta'_\ell\}_\ell$,

$$\begin{aligned}
&\mathbb{P}\big(J(t+v)=J(t) \text{ for all } v\in[0,s] \mid \mathcal{F}_t\otimes\mathcal{F}^B_{t+s}, X(t)=x, J(t)=i\big)\\
&=\sum_{k=0}^{\infty}\frac{(\gamma s)^k}{k!}\mathrm{e}^{-\gamma s}\\
&\qquad\times\int_0^s\cdots\int_0^s\frac{\prod_{j=1}^k\mathbb{P}\big(J(\theta'_j)=J(\theta'_j-)\mid\theta'_j=t+v_j, X^{[i,x,t]}(v_j), J(\theta'_j)=i\big)}{s^k}\,\mathrm{d}v_1\cdots\mathrm{d}v_k\\
&=\sum_{k=0}^{\infty}\frac{(\gamma s)^k}{k!}\mathrm{e}^{-\gamma s}\int_0^s\cdots\int_0^s\frac{\prod_{j=1}^k\big(1+\Lambda_{ii}(X^{[i,x,t]}(v_j))/\gamma\big)}{s^k}\,\mathrm{d}v_1\cdots\mathrm{d}v_k\\
&=\sum_{k=0}^{\infty}\frac{(\gamma s)^k}{k!}\mathrm{e}^{-\gamma s}\left(1+\frac{\int_0^s\Lambda_{ii}\big(X^{[i,x,t]}(v)\big)\,\mathrm{d}v}{\gamma s}\right)^k=\exp\left(\int_0^s\Lambda_{ii}\big(X^{[i,x,t]}(v)\big)\,\mathrm{d}v\right),
\end{aligned}$$

where in the first equality we used that on the event $\{J(t+v)=J(t)\text{for all } v\in[0,s],\ J(t)=i,\ X(t)=t\}$, $X(t+v)=X^{[i,x,t]}(v)$ for all $v\in[0,s]$ by construction. Finally, for $i\neq j\in\mathcal{E}$,

$$\begin{aligned}
&\mathbb{P}\big(J(\theta_\ell)=j\mid J(\theta_\ell-)=i, J(\theta_\ell-)\neq J(\theta_\ell), X(\theta_\ell)\big)\\
&\qquad=\frac{\mathbb{P}\big(J(\theta_\ell)=j\mid J(\theta_\ell-)=i, X(\theta_\ell)\big)}{\mathbb{P}\big(J(\theta_\ell-)\neq J(\theta_\ell)\mid J(\theta_\ell-)=i, X(\theta_\ell)\big)}=\frac{\Lambda_{ij}(X(\theta_\ell))}{|\Lambda_{ii}(X(\theta_\ell))|},
\end{aligned}$$

completing the proof. □

We note that the construction of strong solutions for hybrid SDEs can be carried out in more general scenarios. For instance, Brownian-driven multidimensional hybrid SDEs with unbounded jump intensities are considered in [29].

3. Space-grid approximation of hybrid SDEs

Although solutions to state-dependent hybrid SDEs are easy to simulate following the steps in (2.8), studying their distributional properties is a challenging task (see [39] and references therein). Here we advocate studying hybrid SDEs using an approximation method via discretizing the space grid. This approximation differs from the well-known methods based on Wong–Zakai [30] and Euler–Maruyama [39, (5.4)] methods, and thus requires an analysis of its own, which is carried out next.

For each $i\in\mathcal{E}$, let $\widehat{\mu}(i,\cdot)$ and $\widehat{\sigma}(i,\cdot)$ be piecewise constant càdlàg approximations of $\mu(i,\cdot)$ and $\sigma(i,\cdot)$ on a space grid $\{\zeta_m\}$ of $\mathbb{R}$, with $|\widehat{\mu}(i,\cdot)|\le|\mu(i,\cdot)|$ and $|\widehat{\sigma}(i,\cdot)|\le|\sigma(i,\cdot)|$. Furthermore, let $\widehat{\mathbf{\Lambda}}(\cdot)$ be a piecewise constant càdlàg approximation of $\mathbf{\Lambda}(\cdot)$ over the space grid $\{\zeta_m\}$, with $\widehat{\mathbf{\Lambda}}(x)$ being an intensity matrix such that $\sup_i|\widehat{\Lambda}_{ii}(x)|\le\gamma$ for all $x\in\mathbb{R}$. Additionally,

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

for each $i \in \mathcal{E}$ and $x \in \mathbb{R}$, let $\{\widehat{\Delta}_{ij}(x)\}_j$ be a partition of $[0, \gamma)$ with $\mathrm{Leb}(\widehat{\Delta}_{ij}(x)) = \gamma\delta_{ij} + \widehat{\Lambda}_{ij}(x)$, where each aforementioned set is such that the following function is measurable:

$$\widehat{h}(i, x, z) = \sum_{j=1}^{p} (j - i) 1_{\{z \in \widehat{\Delta}_{ij}(x)\}}.$$

For such (fixed) functions and partitions, we assume the following.

Assumption 3.1. *There exists some filtered probability space* $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{G}_t\}_{t\geq 0})$ *that supports an independent Brownian motion* B *and a Poisson random measure* $\mathfrak{p}^*(\mathrm{d}t, \mathrm{d}x)$ *of intensity* $\mathrm{d}t \times \mathrm{d}x$ *on* $\mathbb{R}_+ \times [0, \gamma)$ *such that:*

- $\{\mathcal{G}_t\}_{t\geq 0}$ is *admissible* with respect to B and $\mathfrak{p}^*$, i.e. $B(t + s) - B(t)$ and $\mathfrak{p}^*((t, t + s], \cdot)$ are independent of $\mathcal{G}_t$ for all $s, t \geq 0$;
- $(\widehat{J}, \widehat{X})$ is the $\mathcal{G}_t$-adapted solution of the hybrid SDE

$$\begin{aligned} \mathrm{d}\widehat{X}(t) &= \widehat{\mu}(\widehat{J}(t), \widehat{X}(t))\,\mathrm{d}t + \widehat{\sigma}(\widehat{J}(t), \widehat{X}(t))\,\mathrm{d}B(t), \quad \widehat{X}(0) = x_0 \in \mathbb{R}, \\ \mathrm{d}\widehat{J}(t) &= \int_{[0,\gamma)} \widehat{h}(\widehat{J}(t-), \widehat{X}(t-), z)\,\mathfrak{p}^*(\mathrm{d}t, \mathrm{d}z), \qquad \widehat{J}(0) = i_0 \in \mathcal{E}. \end{aligned} \tag{3.1}$$

In Assumption 3.1, we essentially seek a *weak solution* for the discretized hybrid SDE (3.1). Namely, in the presence of discontinuous coefficients we need to assume the existence of some probability space (endowed with a filtration that may be larger than the one generated by $\{B(s), \mathfrak{p}(s, \cdot); s \leq t\}$) such that $(\widehat{J}, \widehat{X})$ solves (3.1). This requirement is necessary because SDEs with discontinuous coefficients generally do not have strong solutions. For example, in [25], the existence of a strong solution for SDEs with discontinuous locally integrable drift is established only under a unit diffusion coefficient, while in [27], the same conclusion is reached for piecewise Lipschitz-continuous drift coefficients and Lipschitz-continuous diffusion coefficients. Despite this limitation, we discuss two approximating schemes below for which Assumption 3.1 is guaranteed to hold.

Example 3.1. Suppose that, for all $i \in \mathcal{E}$, $\sigma(i, \cdot)$ and $\widehat{\sigma}(i, \cdot)$ are equal to some $\sigma_i \geq 0$. By [25], for each $i \in \mathcal{E}$, $x \in \mathbb{R}$, and $v \geq 0$, there exists a strong solution $\widehat{X}^{[i,x,v]}$ to the SDE

$$\widehat{X}^{[i,x,v]}(t) = x + \int_0^t \widehat{\mu}\big(i, \widehat{X}^{[i,x,v]}(t)\big)\,\mathrm{d}r + \int_0^t \widehat{\sigma}\big(i, \widehat{X}^{[i,x,v]}(t)\big)\,\mathrm{d}B^{[v]}(r). \tag{3.2}$$

Under these conditions, using the same concatenating procedure as in Section 2 with μ and σ replaced by $\widehat{\mu}$ and $\widehat{\sigma}$, we can readily construct the strong solution to (3.1), ensuring that Assumption 3.1 holds.

Example 3.2. Suppose that, for all $i \in \mathcal{E}$, $\sigma(i, \cdot) = \sigma(\cdot)$, where σ is a positive function bounded away from 0 and ∞. Then the solution to the hybrid SDE (2.1) and (2.2) can be injectively mapped into a unit diffusion hybrid SDE, denoted by $(J, X^\dagger)$, where $X^\dagger(t) := g(X(t))$ and

$$g(x) = \int_{x_0}^{x} \frac{1}{\sigma(y)}\,\mathrm{d}y.$$

This mapping, which is a homeomorphism in space, corresponds to the well-known Lamperti transform (see [26] and [30] for the hybrid framework generalization). Consequently, we can

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

consider an approximation $(\widehat{J}^{\dagger}, \widehat{X}^{\dagger})$ of $(J, X^{\dagger})$ that falls under the category of Example 3.1 and is thus guaranteed to have a strong solution. It is worth noting that while the original hybrid SDE is not the one being approximated, its first-passage properties can be readily studied through those of $(\widehat{J}^{\dagger}, \widehat{X}^{\dagger})$, given the homeomorphic nature of the Lamperti transform.

Remark 3.1. We emphasize that Examples 3.1 and 3.2 are merely instances where a strong solution can be constructed, but weak solutions permit more general scenarios. Indeed, in [8], it is proven that (3.2) admits a unique weak solution under relatively mild conditions for piecewise constant $\widehat{\mu}$ and $\widehat{\sigma}$. Within the hybrid SDE framework, a weak solution to (3.1) could potentially be constructed by appropriately concatenating paths of weak solutions of (3.2). Providing a thorough analysis of weak solutions for (3.1) is a challenging task, with directions that deviate from our main objective. For this reason, we refrain from pursuing this in the present paper.

Due to the discretization characteristics of $\widehat{\mu}$ and $\widehat{\sigma}$, the process $\widehat{X}$ falls within the class of *multi-regime Markov-modulated Brownian motions* studied in [21] (see [28] for an earlier reference in the case $\widehat{\sigma} \equiv 0$). Indeed, the piecewise-constant nature of $\widehat{\mu}$, $\widehat{\sigma}$, and $\widehat{\mathbf{\Lambda}}$ implies that, when restricted to a space interval (ζ_m, ζ_{m+1}) and to the time intervals for which $\widehat{J}$ is equal to i, $\widehat{X}$ behaves like a Brownian motion with drift $\widehat{\mu}(i, \zeta_m)$ and noise coefficient $\widehat{\sigma}(i, \zeta_m)$. Moreover, since the functions $\widehat{\mu}$, $\widehat{\sigma}$, and $\widehat{\mathbf{\Lambda}}$ are space-grid approximations of μ, σ, and $\mathbf{\Lambda}$, we can expect that $\widehat{X}$ approximates the original solution X. A main purpose of this paper is to formalize this statement.

To prove pathwise convergence of $(\widehat{J}, \widehat{X})$ to (J, X) as one of its parameters goes to ∞, we proceed in two steps. First, we investigate the convergence of $\widehat{X}$ to X on the event that $\widehat{J}$ coincides with J in some time interval $[0, \xi)$. Then, we let ξ go to infinity and obtain the rate at which the probabilities of $\widehat{J}$ being different from J on $[0, \xi)$ converge to 0.

3.1. Measuring $|X - \widehat{X}|$

Here we focus on measuring $|X - \widehat{X}|$ over a compact time interval where J and $\widehat{J}$ are identical, where both (J, X) and $(\widehat{J}, \widehat{X})$ are defined in the probability space $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{G}_t\}_{t\geq 0})$ considered in Assumption 3.1. Generally, measuring the distance between two weak solutions is not feasible, as they may not be defined on the same probability space. However, since (J, X) is guaranteed to possess a strong solution in $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{G}_t\}_{t\geq 0})$, it is reasonable to measure $|X - \widehat{X}|$ in that space. We emphasize that, due to Assumption 3.1, the probability space $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{G}_t\}_{t\geq 0})$ is not necessarily fixed. In particular, two different approximations (along with (J, X)) may exist in distinct probability spaces.

We begin our analysis by introducing a fundamental condition that will be beneficial for our subsequent developments. We offer a simpler proof compared to that in [39, Proposition∼2.3], aiming to demonstrate that the proof is valid even with discontinuous coefficients.

Lemma 3.1. *Let (J,X) and $(\widehat{J}, \widehat{X})$ satisfy Assumptions 2.1 and 3.1. Then*

$$\mathbb{E}\big(\sup_{s\leq t} |X(s)|^2\big) < \infty, \qquad \mathbb{E}\big(\sup_{s\leq t} |\widehat{X}(s)|^2\big) < \infty.$$

Proof. Let

$$\tau_m = \inf\{s \geq 0 \colon X(s) \vee \widehat{X}(s) > m\}. \tag{3.3}$$

Then $\int_0^{t\wedge\tau_m} |\sigma(J(s), X(s))|^2 \, \mathrm{d}s < \infty$, so that $V(t) := \int_0^{t\wedge\tau_m} \sigma(J(s), X(s)) \, \mathrm{d}B(s)$ is a local martingale. Employing Doob's inequaliy and Itô isometry on $V(t)$, along with Tonelli's theorem,

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

we get

$$\mathbb{E}\Big(\sup_{r\le t\wedge\tau_m}\Big|\int_0^{r\wedge\tau_m}\sigma(J(s),X(s))\,\mathrm{d}B(s)\Big|^2\Big)\le 4\mathbb{E}\Big(\int_0^{t\wedge\tau_m}|\sigma(J(s),X(s))|^2\,\mathrm{d}s\Big)$$
$$\le 4\int_0^t\mathbb{E}\big(\sup_{r\le s\wedge\tau_m}|\sigma(J(r),X(r))|^2\big)\,\mathrm{d}s.$$

Now, the Lipschitz condition of $\sigma(i,\cdot)$ implies a linear growth that takes the form

$$\sup_{r\le s\wedge\tau_m}|\sigma(J(r),X(r))|\le\sup_{i\in\mathcal{S}}|\sigma(i,0)|+K\sup_{r\le s\wedge\tau_m}|X(r)|,$$

which, by the inequality $(a+b)^2\le 2a^2+2b^2$, means

$$\mathbb{E}\Big(\sup_{r\le t\wedge\tau_m}\Big|\int_0^{r\wedge\tau_m}\sigma(J(s),X(s))\,\mathrm{d}B(s)\Big|^2\Big)$$
$$\le 8t\sup_{i\in\mathcal{S}}|\sigma(i,0)|+8K\int_0^t\mathbb{E}\big(\sup_{r\le s\wedge\tau_m}|X(r)|^2\big)\,\mathrm{d}s. \quad (3.4)$$

Employing the Cauchy–Schwartz inequality and analogous steps leads to the following inequalities:

$$\mathbb{E}\Big(\sup_{r\le t\wedge\tau_m}\Big|\int_0^{r\wedge\tau_m}\mu(J(s),X(s))\,\mathrm{d}s\Big|^2\Big)\le t\int_0^t\mathbb{E}\big(\sup_{r\le s\wedge\tau_m}|\mu(J(r),X(r))|^2\big)\,\mathrm{d}s$$
$$\le 2t^2\sup_{i\in\mathcal{S}}|\mu(i,0)|$$
$$+2tK\int_0^t\mathbb{E}\big(\sup_{r\le s\wedge\tau_m}|X(r)|^2\big)\,\mathrm{d}s. \quad (3.5)$$

Using the inequality $(a+b+c)^2\le 3a^2+3b^2+3c^2$ in (2.1), along with (3.4) and (3.5), we have

$$\mathbb{E}\big(\sup_{r\le t\wedge\tau_m}|X(r)|^2\big)\le 3|x_0|^2+3\mathbb{E}\Big(\sup_{r\le t\wedge\tau_m}\Big|\int_0^{r\wedge\tau_m}\mu(J(s),X(s))\,\mathrm{d}s\Big|^2\Big)$$
$$+3\mathbb{E}\Big(\sup_{r\le t\wedge\tau_m}\Big|\int_0^{r\wedge\tau_m}\sigma(J(s),X(s))\,\mathrm{d}B(s)\Big|^2\Big)$$
$$\le K_1+K_2\int_0^t\mathbb{E}\big(\sup_{r\le t\wedge\tau_m}|X(r)|^2\big)\,\mathrm{d}s$$

for some constants $K_1, K_2>0$ that are not dependent on m. Gronwall's lemma then implies that $\mathbb{E}\big(\sup_{r\le t\wedge\tau_m}|X(r)|^2\big)\le K_1\mathrm{e}^{K_2t}$; taking $m\to\infty$ yields $\mathbb{E}\big(\sup_{s\le t}|X(s)|^2\big)\le K_1\mathrm{e}^{K_2t}<\infty$. The steps to prove

$$\mathbb{E}\big(\sup_{s\le t}|X(s)|^2\big)\le K_1\mathrm{e}^{K_2t}<\infty$$

can be repeated verbatim by replacing μ,σ,J, and X with $\widehat{\mu},\widehat{\sigma},\widehat{J}$, and $\widehat{X}$, employing along the way the inequalities

$$\sup_{r\le s\wedge\tau_m}|\widehat{\sigma}(\widehat{J}(r),\widehat{X}(r))|\le\sup_{r\le s\wedge\tau_m}|\sigma(\widehat{J}(r),\widehat{X}(r))|$$
$$\le\sup_{i\in\mathcal{S}}|\sigma(i,0)|+K\sup_{r\le s\wedge\tau_m}|\widehat{X}(r)|,$$
$$\sup_{r\le s\wedge\tau_m}|\widehat{\mu}(\widehat{J}(r),\widehat{X}(r))|\le\sup_{r\le s\wedge\tau_m}|\mu(\widehat{J}(r),\widehat{X}(r))|$$
$$\le\sup_{i\in\mathcal{S}}|\mu(i,0)|+K\sup_{r\le s\wedge\tau_m}|\widehat{X}(r)|.$$

□

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

Now, we quantify the distance between X and $\widehat{X}$ up to the *decoupling* time ι defined by $\iota := \inf\{s \geq 0 \colon J(s) \neq \widehat{J}(s)\}$.

Theorem 3.1. *For $t \geq 0$,*

$$\mathbb{E}\big(\sup_{s\leq t\wedge\iota} |X(s) - \widehat{X}(s)|^2\big) \leq C(t)\big(\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2\big), \quad (3.6)$$

where $C(t) = 16(t^2 \vee 1)\mathrm{e}^{4K^2(t+4)^2}$.

Proof. Let τ_m be as in (3.3), and for $t \geq 0$ define $Z_m(t) = \mathbb{E}\big(\sup_{s\leq t\wedge\tau_m\wedge\iota} |X(s) - \widehat{X}(s)|^2\big)$. By Lemma 3.1, $Z_m(t)$ is finite for all $t \geq 0$. Using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, together with the fact that $\widehat{J}$ agrees with J on the time interval $[0, \iota)$, we get

$$Z_m(t) \leq 2\mathbb{E}\left(\sup_{s\leq t\wedge\tau_m\wedge\iota} \left| \int_0^s \mu(J(r), X(r)) - \widehat{\mu}(J(r), \widehat{X}(r))\, \mathrm{d}r \right|^2\right) + 2\mathbb{E}\left(\sup_{s\leq t\wedge\tau_m\wedge\iota} \left| \int_0^s \sigma(J(r), X(r)) - \widehat{\sigma}(J(r), \widehat{X}(r))\, \mathrm{d}B(r) \right|^2\right). \quad (3.7)$$

Employing similar techniques to those in the proof of Lemma 3.1, we get

$$\mathbb{E}\left(\sup_{s\leq t\wedge\tau_m\wedge\iota} \left| \int_0^s \mu(J(r), X(r)) - \widehat{\mu}(J(r), \widehat{X}(r))\, \mathrm{d}r \right|^2\right) \leq t \int_0^t \mathbb{E}\big(\sup_{r\leq s\wedge\tau_m\wedge\iota} |\mu(J(r), X(r)) - \widehat{\mu}(J(r), \widehat{X}(r))|^2\big)\, \mathrm{d}s, \quad (3.8)$$

$$\mathbb{E}\left(\sup_{s\leq t\wedge\tau_m\wedge\iota} \left| \int_0^s \sigma(J(r), X(r)) - \widehat{\sigma}(J(r), \widehat{X}(r))\, \mathrm{d}r \right|^2\right) \leq 4 \int_0^t \mathbb{E}\big(\sup_{r\leq s\wedge\tau_m\wedge\iota} |\sigma(J(r), X(r)) - \widehat{\sigma}(J(r), \widehat{X}(r))|^2\big)\, \mathrm{d}s. \quad (3.9)$$

Furthermore, the Lipschitz continuity of μ implies that

$$\begin{aligned} &|\mu(J(r), X(r)) - \widehat{\mu}(J(r), \widehat{X}(r))|^2 \\ &\quad \leq 2|\mu(J(r), X(r)) - \mu(J(r), \widehat{X}(r))|^2 + 2|\mu(J(r), \widehat{X}(r)) - \widehat{\mu}(J(r), \widehat{X}(r))|^2 \\ &\quad \leq 2K^2|X(r) - \widehat{X}(r)|^2 + 2\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2; \end{aligned} \quad (3.10)$$

similarly,

$$|\sigma(J(r), X(r)) - \widehat{\sigma}(J(r), \widehat{X}(r))|^2 \leq 2K^2|X(r) - \widehat{X}(r)|^2 + 2\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2. \quad (3.11)$$

By virtue of (3.7)–(3.11), we get

$$\begin{aligned} Z_m(t) &\leq 4t^2 \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + 16t \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2 \\ &\quad + 4K^2(t+4) \int_0^t Z_m(r)\, \mathrm{d}r \\ &\leq 16(t^2 \vee 1)\big(\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2\big) \\ &\quad + 4K^2(t+4) \int_0^t Z_m(r)\, \mathrm{d}r. \end{aligned}$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

Finally, Gronwall's lemma implies that

$$\begin{aligned} Z_m(t) &\le 16(t^2 \vee 1)\mathrm{e}^{4K^2(t+4)t}\big(\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2\big) \\ &\le C(t)\big(\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2\big), \end{aligned}$$

so that (3.6) holds by taking $m \to \infty$. □

As usual, mean-square distance provides bounds on pointwise distance, which in our case takes the following form.

Corollary 3.1. *Suppose that there exists some $\alpha > 0$ such that*

$$\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2 \le \alpha^2.$$

Then, for $\delta > 0$ and for all sufficiently large $n \ge 1$,

$$\mathbb{P}\big(\sup_{s\le t\wedge\iota} |X(s) - \widehat{X}(s)| \ge \alpha\sqrt{n^\delta C(t)}\big) \le n^{-\delta}.$$

Proof. Markov's inequality and Theorem 3.1 imply that, for any $\beta_{n,t,\alpha} > 0$,

$$\begin{aligned} &\mathbb{P}\big(\sup_{s\le t\wedge\iota} |X(s) - \widehat{X}(s)| \ge \beta_{n,t,\alpha}\big) \\ &\quad\le \frac{\mathbb{E}\big(\sup_{s\le t\wedge\iota} |X(s) - \widehat{X}(s)|^2\big)}{\beta_{n,t,\alpha}^2} \\ &\quad\le \frac{C(t)\big(\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\mu(i,x) - \widehat{\mu}(i,x)|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} |\sigma(i,x) - \widehat{\sigma}(i,x)|^2\big)}{\beta_{n,t,\alpha}^2} \le \frac{C(t)\alpha^2}{\beta_{n,t,\alpha}^2}. \end{aligned}$$

The result follows by choosing $\beta_{n,t,\alpha} = \alpha\sqrt{n^\delta C(t)}$. □

Corollary 3.1 implies that we can choose appropriate values of t and α that are dependent on n, say t_n and α_n, such that, for $\delta > 0$, $t_n \to \infty$ and $\alpha_n\sqrt{n^\delta C(t_n)} \to 0$ as $n \to \infty$, and then conclude that $|X - \widehat{X}|$ converges in probability to 0 uniformly over increasing compact intervals (up to time ι). Here we choose $t_n = \sqrt{\log\log n}$ and $\alpha_n = n^{-\delta_* - \delta/2}$ for some $\delta_* > 0$. In such a case, for all $n \ge 1$ with $\sqrt{\log n \log n} > 4$ and $\log n > \log\log n$,

$$\alpha_n\sqrt{n^\delta C(t_n)} = 16(t_n^2 \vee 1)\mathrm{e}^{4K^2(t_n+4)^2} n^{-\delta_*} \le 16(\log n)\mathrm{e}^{8K^2\log\log n} n^{-\delta_*} = 16(\log n)^{\beta_*} n^{-\delta_*},$$

where $\beta_* = 1 + 8K^2$. In light of this, we assume the following.

Assumption 3.2. *Consider the functions $\widehat{\mu}$ and $\widehat{\sigma}$ associated with the process $(\widehat{J}, \widehat{X})$, all of which now exhibit an explicit dependence on $n \ge 1$ via their superscript. Then, there exist δ_*, $\delta > 0$ such that*

$$\sup_{i\in\mathcal{E},\, x\in\mathbb{R}} \big|\mu(i,x) - \widehat{\mu}^{(n)}(i,x)\big|^2 + \sup_{i\in\mathcal{E},\, x\in\mathbb{R}} \big|\sigma(i,x) - \widehat{\sigma}^{(n)}(i,x)\big|^2 \le n^{-\delta_* - \delta/2}.$$

To summarize, under Assumptions 2.1, 3.1, and 3.2, for all sufficiently large $n \ge 1$,

$$\mathbb{P}\big(\sup_{s\le t_n\wedge\iota} \big|X(s) - \widehat{X}^{(n)}(s)\big| \ge 16(\log n)^{\beta_*} n^{-\delta_*}\big) \le n^{-\delta}, \quad \text{where } \beta_* = 1 + 8K^2. \tag{3.12}$$

From (3.12), we can determine various types of convergence in different situations. If the existence of a solution under Assumption 3.1 can be established in a strong sense with a fixed

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

probability space for all approximations, then (3.12) demonstrates the convergence in probability of $\widehat{X}^{(n)}$ to X up to time ι. Furthermore, if $\delta > 1$, then, using Borel–Cantelli arguments, the convergence holds in a strong sense. In the general case where the solution considered in Assumption 3 is weak, (3.12) establishes the weak convergence of $\widehat{X}^{(n)}$ to X up to time ι. Refer to [12, 13] for instances where measuring pointwise distances in distinct probability spaces leads to weakly convergent approximations; an additional example using the analogous formulation of triangular arrays can be found in [15].

Now that we have assessed the convergence of $\widehat{X}^{(n)}$ to X within the interval $[0, \iota)$, in the following section we establish some probabilistic properties of the decoupling time ι.

3.2. Measuring the decoupling event

In order to guarantee that the decoupling time ι occurs at a large epoch, we need to ensure that the jumps of $\widehat{J}$ remain 'similar' to those of J. The way to guarantee this is through the sets $\{\widehat{\Delta}_{ij}(x)\}_{i,j\mathcal{E}}$ associated to the discretized hybrid SDE (3.1). The idea is to make each set $\widehat{\Delta}_{ij}(x)$ as close as possible to $\Delta_{ij}(x)$, with the difference that the $\mathrm{Leb}(\widehat{\Delta}_{ij}(x)) = \gamma\delta_{ij} + \widehat{\Lambda}_{ij}(x)$ while $\mathrm{Leb}(\Delta_{ij}(x)) = \gamma\delta_{ij} + \Lambda_{ij}(x)$.

Recall that, in Section 2, for a fixed $x \in \mathbb{R}$ and $i \in \mathcal{E}$, we let $\Delta_{ij}(x) = [\ell_{ij}(x), r_{ij}(x))$, where $\ell_{i1}(x) = 0$, $r_{ij}(x) = \ell_{ij}(x) + (\gamma\delta_{ij} + \Lambda_{ij}(x))$, $\ell_{i,j+1}(x) = r_{ij}(x)$, and $r_{ip}(x) = \gamma$. Next, we introduce $\Gamma_{ij}(x) = [\ell_{ij}(x), r_{ij}(x) \wedge \widehat{r}_{ij}(x))$, where $\widehat{r}_{ij}(x) = \ell_{ij}(x) + (\gamma\delta_{ij} + \widehat{\Lambda}_{ij}(x))$. Note that

$$\mathrm{Leb}\big(\Gamma_{ij}(x)\big) = (\gamma\delta_{ij} + \Lambda_{ij}(x)) \wedge (\gamma\delta_{ij} + \widehat{\Lambda}_{ij}(x)) \leq \mathrm{Leb}\big(\widehat{\Delta}_{ij}(x)\big);$$

in simple words, since $\widehat{\Lambda}_{ij}(x)$ is thought of as being close to $\Lambda_{ij}(x)$, the length of $\Gamma_{ij}(x)$ approximates the required Lebesgue measure of $\widehat{\Delta}_{ij}(x)$ from below. For the last step, for a fixed $i \in \mathcal{E}$ and for all $x \in \mathbb{R}$, we let $\{\Gamma^*_{ij}(x)\}_j$ be a partition of the set $[0, \gamma) \setminus (\cup_j \Gamma_{ij}(x))$ that keeps the mapping $(x, z) \mapsto 1_{\{z \in \Gamma_{ij}(x)\}}$ measurable, with $\mathrm{Leb}\big(\Gamma^*_{ij}(x)\big) = \mathrm{Leb}\big(\widehat{\Delta}_{ij}(x)\big) - \mathrm{Leb}\big(\Gamma_{ij}(x)\big)$. (Employing the measurability properties of $\Lambda_{ij}(\cdot)$ and $\widehat{\Lambda}_{ij}(\cdot)$, constructing a partition $\{\Gamma^*_{ij}(x)\}_j$ with the desired characteristics is always achievable. Presenting an explicit construction is tedious and, more importantly, it is not relevant to subsequent developments; for this reason, we refrain from detailing it.) We then define $\widehat{\Delta}_{ij}(x) := \Gamma_{ij}(x) \cup \Gamma^*_{ij}(x)$, from which we can readily verify that $\widehat{\Delta}_{ij}(x)$ indeed has a Lebesgue measure equal to $\gamma\delta_{ij} + \widehat{\Lambda}_{ij}(x)$, $\{\widehat{\Delta}_{ij}(x)\}_j$ is a partition of $[0, \gamma)$, and the mapping $(x, z) \mapsto 1_{\{z \in \widehat{\Delta}_{ij}(x)\}}$ is measurable. This implies the characteristics needed for $\widehat{\Delta}_{ij}(x)$ as discussed in Section 3. We now relate the sets $\Gamma_{ij}(x)$ to the decoupling time ι.

Lemma 3.2. *Define* $\kappa = \inf\big\{m \geq 1 \colon U_m \notin \bigcup_{j=1}^{p} \Gamma_{J(\theta_{m-1}),j}(X(\theta_m)) \cap \Gamma_{\widehat{J}(\theta_{m-1}),j}(\widehat{X}(\theta_m))\big\}$. *Then* $\theta_\kappa \leq \iota$.

Proof. It is enough to prove that $\{\kappa > m\} \subseteq \{\iota > \theta_m\}$ for all $m \geq 1$. Assuming that this holds for some fixed $m \geq 1$, We proceed by induction. Then,

$$\begin{aligned}\{\kappa > m+1\} &= \{\kappa > m\} \cap \left\{U_{m+1} \in \bigcup_{j=1}^{p} \Gamma_{J(\theta_m),j}(X(\theta_{m+1})) \cap \Gamma_{\widehat{J}(\theta_m),j}(\widehat{X}(\theta_{m+1}))\right\} \\ &\subseteq \{\iota > \theta_m\} \cap \left\{U_{m+1} \in \bigcup_{j=1}^{p} \Gamma_{J(\theta_m),j}(X(\theta_{m+1})) \cap \Gamma_{\widehat{J}(\theta_m),j}(\widehat{X}(\theta_{m+1}))\right\}\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

$$
\begin{aligned}
&= \bigcup_{j=1}^{p} \{\iota > \theta_m\} \cap \left\{ U_{m+1} \in \Gamma_{J(\theta_m),j}(X(\theta_{m+1})) \cap \Gamma_{\widehat{J}(\theta_m),j}(\widehat{X}(\theta_{m+1})) \right\} \\
&\subseteq \bigcup_{j=1}^{p} \{\iota > \theta_m\} \cap \left\{ U_{m+1} \in \Delta_{J(\theta_m),j}(X(\theta_{m+1})) \cap \widehat{\Delta}_{\widehat{J}(\theta_m),j}(\widehat{X}(\theta_{m+1})) \right\} \\
&= \bigcup_{j=1}^{p} \{\iota > \theta_m\} \cap \{J(\theta_{m+1}) = \widehat{J}(\theta_{m+1}) = j\} = \{\iota > \theta_{m+1}\},
\end{aligned}
$$

where in the fourth line we employed that $\Gamma_{ij}(x)$ is contained in $\Delta_{ij}(x)$ and $\widehat{\Delta}_{ij}(x)$ for all $i, j \in \mathcal{E}$, $x \in \mathbb{R}$. This completes the induction and the proof. □

As a preliminary step, let us provide bounds for the probability of a possible decoupling at a given epoch θ_m for fixed values of $X(\theta_m)$, $\widehat{X}^{(n)}(\theta_m)$, and $J(\theta_{m-1}) = \widehat{J}^{(n)}(\theta_{m-1})$.

Lemma 3.3. *For a square matrix $A = \{a_{ij}\}$, let $\|\cdot\|$ denote the norm defined by $\|A\| = \sup_{i,j}\{|a_{ij}|\}$. Then, for $i \in \mathcal{E}$ and $x, y \in \mathbb{R}$,*

$$
\begin{aligned}
\mathbb{P}\left(U_m \notin \bigcup_{j=1}^{p} \Gamma_{ij}(x) \cap \Gamma_{ij}(y) \right) &\le \frac{p}{\gamma} \left(p\|\mathbf{\Lambda}(x) - \mathbf{\Lambda}(y)\| + 2p \sup_{z\in\mathbb{R}} \left\| \widehat{\mathbf{\Lambda}}^{(n)}(z) - \mathbf{\Lambda}(z) \right\| \right) \\
&\le \frac{2p^2}{\gamma} \left(\|\mathbf{\Lambda}(x) - \mathbf{\Lambda}(y)\| + \sup_{z\in\mathbb{R}} \left\| \widehat{\mathbf{\Lambda}}^{(n)}(z) - \mathbf{\Lambda}(z) \right\| \right).
\end{aligned}
$$

Proof. Note that $\Gamma_{ij}(x) \cap \Gamma_{ij}(y) = [\ell_{ij}(x, y), r_{ij}(x, y))$, where $\ell_{ij}(x, y) = \ell_{ij}(x) \vee \ell_{ij}(y)$ and $r_{ij}(x, y) = (r_{ij}(x) \wedge \widehat{r}_{ij}(x)) \wedge (r_{ij}(y) \wedge \widehat{r}_{ij}(y))$, with the understanding that $[a, b)$ is an empty interval if $a \ge b$. Moreover, since $\ell_{i1}(x, y) = 0$ and, for all $1 \le j < p$ we have $\ell_{ij}(x, y) \le \ell_{i,j+1}(x, y)$ and $r_{ij}(x, y) \le r_{i,j+1}(x, y)$, the set where U_k may fall, leading to a potential decoupling, can be written as

$$
[0, \gamma) \setminus \left(\bigcup_{j=1}^{p} \Gamma_{ij}(x) \cap \Gamma_{ij}(y) \right) = [0, \gamma) \setminus \left(\bigcup_{j=1}^{p} [\ell_{ij}(x, y), r_{ij}(x, y)) \right) = \bigcup_{j=1}^{p} [r_{ij}(x, y), \ell_{i,j+1}(x, y)),
$$

where $\ell_{i,p+1}(x, y) := \gamma$. Thus, the probability of decoupling is bounded as follows:

$$
\begin{aligned}
&\mathbb{P}\left(U_k \notin \bigcup_{j=1}^{p} \Gamma_{ij}(x) \cap \Gamma_{ij}(y) \right) \\
&\quad = \frac{1}{\gamma} \operatorname{Leb}\left([0, \gamma) \setminus \bigcup_{j=1}^{p} \Gamma_{ij}(x) \cap \Gamma_{ij}(y) \right) \\
&\quad = \frac{1}{\gamma} \operatorname{Leb}\left(\bigcup_{j=1}^{p} (r_{ij}(x, y), \ell_{i,j+1}(x, y)] \right) \\
&\quad \le \frac{1}{\gamma} \sum_{j=1}^{p} \operatorname{Leb}\left((r_{ij}(x, y), \ell_{i,j+1}(x, y)] \right)
\end{aligned}
$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

$$
\begin{aligned}
&\le \frac{1}{\gamma}\sum_{j=1}^{p} |\ell_{i,j+1}(x,y) - r_{ij}(x,y)| \\
&= \frac{1}{\gamma}\sum_{j=1}^{p} |(\ell_{i,j+1}(x) \vee \ell_{i,j+1}(y)) - ((r_{ij}(x) \wedge \widehat{r}_{ij}(x)) \wedge (r_{ij}(y) \wedge \widehat{r}_{ij}(y)))| \\
&= \frac{1}{\gamma}\sum_{j=1}^{p} |(r_{ij}(x) \vee r_{ij}(y)) - ((r_{ij}(x) \wedge \widehat{r}_{ij}(x)) \wedge (r_{ij}(y) \wedge \widehat{r}_{ij}(y)))| \\
&\le \frac{1}{\gamma}\sum_{j=1}^{p} (|r_{ij}(x) - r_{ij}(y)| + |r_{ij}(y) - \widehat{r}_{ij}(y)| + |r_{ij}(x) - \widehat{r}_{ij}(x)|).
\end{aligned}
$$

Note that from the recursive definition of $r_{ij}(z) = \sum_{k=1}^{j} (\gamma\delta_{ik} + \Lambda_{ik}(z))$, it holds that, for all $x, y, z \in \mathbb{R}$,

$$
|r_{ij}(x) - r_{ij}(y)| \le \sum_{k=1}^{j} |\Lambda_{ik}(x) - \Lambda_{ik}(y)| \le j\|\mathbf{\Lambda}(x) - \mathbf{\Lambda}(y)\| \le p\|\mathbf{\Lambda}(x) - \mathbf{\Lambda}(y)\|,
$$

and since $r_{ij}(z) - \widehat{r}_{ij}(z) = (\ell_{ij}(z) + \gamma\delta_{ij} + \Lambda_{ij}(z)) - (\ell_{ij}(z) + \gamma\delta_{ij} + \widehat{\Lambda}_{ij}(z)) = \Lambda_{ij}(z) - \widehat{\Lambda}_{ij}(z)$,

$$
|r_{ij}(z) - \widehat{r}_{ij}(z)| = |\Lambda_{ij}(z) - \widehat{\Lambda}_{ij}(z)| \le \|\mathbf{\Lambda}(z) - \widehat{\mathbf{\Lambda}}(z)\| \le p\|\mathbf{\Lambda}(z) - \widehat{\mathbf{\Lambda}}(z)\|.
$$

Substituting these bounds into the sum yields the result. □

Next, we study the arrivals $\{\theta_k\}_{k\ge1}$ that occur in the interval $[0, t_n]$. To do so, let us introduce $\{N(t)\}_{t\ge0}$, the Poisson counting process of intensity γ associated to $\{\theta_k\}_{k\ge1}$. Each arrival represents an opportunity for J and $\widehat{J}$ to get decoupled, so that providing a tight tail estimate for $N(t_n)$ is relevant to our analysis.

Lemma 3.4. $\mathbb{P}(N(t_n) > \lfloor \log\log n \rfloor) = o(1)$ *as* $n \to \infty$.

Proof. Using [19, Proposition 1], for sufficiently large n,

$$
\mathbb{P}(N(t_n) > \lfloor \log\log n \rfloor) \le \left(1 - \frac{\gamma t_n}{\lfloor \log\log n \rfloor + 1}\right)^{-1} \left(\frac{(\gamma t_n)^{\lfloor \log\log n \rfloor}}{\lfloor \log\log n \rfloor!} \mathrm{e}^{-\gamma t_n}\right);
$$

the first element in the product on the right-hand side clearly converges to 1, while the second element is bounded by

$$
\begin{aligned}
&\frac{(\gamma t_n)^{\lfloor \log\log n \rfloor}}{\lfloor \log\log n \rfloor!} \\
&\quad \le \left((\gamma t_n)^{\lfloor \log\log n \rfloor}\right)\left(\sqrt{2\pi \lfloor \log\log n \rfloor}\left(\frac{\lfloor \log\log n \rfloor}{\mathrm{e}}\right)^{\lfloor \log\log n \rfloor} \exp\left\{\frac{1}{12\lfloor \log\log n \rfloor + 1}\right\}\right)^{-1} \\
&\quad = O\left(\left(\frac{\mathrm{e}\gamma t_n}{\lfloor \log\log n \rfloor}\right)^{\lfloor \log\log n \rfloor}\right) = o(1).
\end{aligned}
$$

□

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

From Lemma 3.3, it is clear that we need some regularity assumptions on $\mathbf{\Lambda}$ and $\widehat{\mathbf{\Lambda}}$ in order to bound the probability of a decoupling occuring. Similarly to Assumption 3.2, let us now make explicit the dependence of $\widehat{\mathbf{\Lambda}}$, $\widehat{J}$, $\widehat{X}$, ι, and κ on n via a superscript.

Assumption 3.3. *The intensity matrix $\mathbf{\Lambda}(\cdot)$ is* log*-Hölder continuous, in the sense that there exists a constant $G > 0$ such that, for any $x, y \in \mathbb{R}$,*

$$\|\mathbf{\Lambda}(x) - \mathbf{\Lambda}(y)\| \leq \frac{G}{-\log|x-y|}.$$

Furthermore, $\widehat{\mathbf{\Lambda}}^{(n)}$ is right-continuous with left limits and converges uniformly to $\mathbf{\Lambda}$ at a log *rate; more specifically,*

$$\sup_{z\in\mathbb{R}} \left\|\widehat{\mathbf{\Lambda}}^{(n)}(z) - \mathbf{\Lambda}(z)\right\| \leq \frac{G}{\log n}.$$

Remark 3.2. Under Assumption 3.3, $\mathbf{\Lambda}$ does not admit any discontinuities. However, recall that log-Hölder continuity is considerably less restrictive than Hölder or Lipschitz continuity. This means that we can *mimic* discontinuous behaviour for $\mathbf{\Lambda}$ at a point z_0 by considering instead some steep function that behaves like $(\log(z - z_0))^{-1}$ for sufficiently close $z > z_0$.

Lemma 3.5. *Fix some small $\varepsilon_1 > 0$. Then, as $n \to \infty$,*

$$\mathbb{P}\big(\kappa^{(n)} \leq N(t_n),\ N(t_n) \leq \lfloor \log\log n \rfloor,\ \sup_{s\leq t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| \leq n^{-\delta_*+\varepsilon_1}\big) = o(1).$$

Proof. For brevity, in this proof let us write $\widehat{\mathbf{\Lambda}}$, $\widehat{J}$, $\widehat{X}$, ι, and κ instead of $\widehat{\mathbf{\Lambda}}^{(n)}$, $\widehat{J}^{(n)}$, $\widehat{X}^{(n)}$, $\iota^{(n)}$, and $\kappa^{(n)}$, respectively. Note that

$$\begin{aligned}
&\big\{\kappa \leq N(t_n),\ N(t_n) \leq \lfloor \log\log n \rfloor,\ \sup_{s\leq t_n\wedge\iota} |X(s) - \widehat{X}(s)| \leq n^{-\delta_*+\varepsilon_1}\big\} \\
&\quad = \bigcup_{\ell=1}^{\lfloor \log\log n \rfloor} \bigcup_{k=1}^{\ell} \big\{\kappa = k,\ N(t_n) = \ell,\ \sup_{s\leq t_n\wedge\iota} |X(s) - \widehat{X}(s)| \leq n^{-\delta_*+\varepsilon_1}\big\} \\
&\quad \subseteq \bigcup_{\ell=1}^{\lfloor \log\log n \rfloor} \bigcup_{k=1}^{\ell} \big\{\kappa = k,\ N(t_n) = \ell,\ \sup_{s\leq \theta_k} |X(s) - \widehat{X}(s)| \leq n^{-\delta_*+\varepsilon_1}\big\}. \qquad (3.13)
\end{aligned}$$

Letting $F = \big\{\kappa > k-1,\ N(t_n) = \ell,\ \sup_{s\leq\theta_k} |X(s) - \widehat{X}(s)| \leq n^{-\delta_*+\varepsilon_1}\big\}$ and

$$\mathcal{H} = \sigma\big(N(t_n), \{J(s)\}_{s\leq\theta_{k-1}}, \{\widehat{J}(s)\}_{s\leq\theta_{k-1}}, \{X(s)\}_{s\leq\theta_k}, \{\widehat{X}(s)\}_{s\leq\theta_k}\big),$$

we get

$$\begin{aligned}
&\mathbb{P}\big(\kappa = k,\ N(t_n) = \ell,\ \sup_{s\leq\theta_k} |X(s) - \widehat{X}(s)| \leq n^{-\delta_*+\varepsilon_1} \mid \mathcal{H}\big) \\
&\quad = \mathbb{P}(\iota = \theta_k \mid \mathcal{H}) \times 1_F = \mathbb{P}\left(U_k \notin \bigcup_{j=1}^{p} \Gamma_{J(\theta_{k-1}),j}(X(\theta_k)) \cap \Gamma_{\widehat{J}(\theta_{k-1}),j}(\widehat{X}(\theta_k))\right) \times 1_F \\
&\quad \leq \frac{2p^2}{\gamma}\big(\|\mathbf{\Lambda}(X(\theta_k)) - \mathbf{\Lambda}(\widehat{X}(\theta_k))\| + \sup_{z\in\mathbb{R}} \|\widehat{\mathbf{\Lambda}}(z) - \mathbf{\Lambda}(z)\|\big) \times 1_F \\
&\quad \leq \frac{2p^2}{\gamma}\left(\frac{G}{-\log|X(\theta_k) - \widehat{X}(\theta_k)|} + \frac{G}{\log n}\right) \times 1_F \leq \frac{2p^2 G}{\gamma}\left[\frac{1}{\delta_* - \varepsilon_1} + 1\right](\log n)^{-1}.
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

From (3.13), we get

$$\begin{aligned}
&\mathbb{P}\big(\kappa \le N(t_n),\ N(t_n) \le \lfloor \log\log n \rfloor,\ \sup_{s\le \theta_k\wedge \iota} |X(s) - \widehat{X}(s)| \le n^{-\delta_*+\varepsilon_1}\big) \\
&\quad \le \sum_{\ell=1}^{\lfloor \log\log n\rfloor} \sum_{k=1}^{\ell} \mathbb{P}\big(\kappa = k,\ N(t_n) = \ell,\ \sup_{s\le\theta_k} |X(s) - \widehat{X}(s)| \le n^{-\delta_*+\varepsilon_1}\big) \\
&\quad \le \sum_{\ell=1}^{\lfloor \log\log n\rfloor} \sum_{k=1}^{\ell} \frac{G}{\gamma}\left[\frac{1}{\delta_* - \varepsilon_1} + 1\right](\log n)^{-1} \le \frac{G}{\gamma}\left[\frac{1}{\delta_* - \varepsilon_1} + 1\right]\frac{\lfloor \log\log n\rfloor^2}{\log n} = o(1),
\end{aligned}$$

which concludes the proof. □

Now we are ready to prove the main result of the paper.

Theorem 3.2. *Let Assumptions 2.1, 2.2, 3.1, 3.2, and 3.3 hold. Then, for any fixed $\varepsilon_1 > 0$,*

$$\mathbb{P}\big(\big\{ \sup_{s\le t_n} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\} \cup \{\iota^{(n)} \le t_n\}\big) = o(1) \quad \text{as } n \to \infty. \tag{3.14}$$

Proof. By simple set manipulations,

$$\begin{aligned}
&\big\{\sup_{s\le t_n} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\} \cup \{\iota^{(n)} \le t_n\} \\
&\quad = \big\{\iota^{(n)} > t_n,\ \sup_{s\le t_n} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\} \cup \{\iota^{(n)} \le t_n\} \\
&\quad = \big\{\iota^{(n)} > t_n,\ \sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\} \cup \{\iota^{(n)} \le t_n\} \\
&\quad \subseteq \big\{ \sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\} \cup \{\iota^{(n)} \le t_n\}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\{\iota^{(n)} \le t_n\} &\subseteq \{\kappa^{(n)} \le N(t_n)\} \\
&\subseteq \big\{\kappa^{(n)} \le N(t_n),\ N(t_n) \le \lfloor \log\log n\rfloor,\ \sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| \le n^{-\delta_*+\varepsilon_1}\big\} \\
&\qquad \cup \{N(t_n) > \lfloor \log\log n\rfloor\} \cup \big\{ \sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big\}.
\end{aligned}$$

Thus, the left-hand side of (3.14) is bounded by

$$\begin{aligned}
&\mathbb{P}\big(\kappa^{(n)} \le N(t_n),\ N(t_n) \le \lfloor \log\log n\rfloor,\ \sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| \le n^{-\delta_*+\varepsilon_1}\big) \\
&\qquad + \mathbb{P}(N(t_n) > \lfloor \log\log n\rfloor) + \mathbb{P}\big(\sup_{s\le t_n\wedge\iota^{(n)}} \big|X(s) - \widehat{X}^{(n)}(s)\big| > n^{-\delta_*+\varepsilon_1}\big);
\end{aligned}$$

since each one of the aforementioned summands is shown to be $o(1)$ as $n \to \infty$ in Lemma 3.5, Lemma 3.4, and (3.12), respectively, the result follows. □

Theorem 3.2 essentially indicates that under Assumptions 2.1, 2.2, 3.1, 3.2, and 3.3, there exists a sequence of multi-regime Markov-modulated Brownian motions $\widehat{X}^{(n)}$ that weakly converge to the solution X of a hybrid stochastic differential equation, as well as their respective underlying processes $\widehat{J}^{(n)}$ and J. This convergence occurs uniformly over increasing compact intervals. If, additionally, the solution for each approximation in Assumption 3.1 is strong, the probability space $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{G}_t\}_{t\ge 0})$ may remain fixed for all $n \ge 1$, allowing the result to be upgraded to convergence in probability. In any case, this implies that the first-passage probabilities of X into the set $(-\infty, a) \cup (a, +\infty)$ can be approximated by the corresponding first-passage probabilities of $\widehat{X}^{(n)}$. In the following section we leverage this to offer efficient approximations for certain first-passage probabilities for hybrid stochastic differential equations.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

4. First-passage probabilities and expected occupation times for solutions of hybrid SDEs

Define the first-passage times $\tau_0^- = \inf\{t > 0 \colon X(t) \le 0\}$, $\tau_a^+ = \inf\{t > 0 \colon X(t) \ge a\}$, where $a > 0$. For $q \ge 0$, we are interested in providing approximations for the Laplace transform of the first-passage times,

$$m_{ij}^-(u, q, a) = \mathbb{E}\big(\mathrm{e}^{-q(\tau_0^- \wedge \tau_a^+)} 1_{\{J(\tau_0^- \wedge \tau_a^+)=j, X(\tau_0^- \wedge \tau_a^+)=0\}} \mid J(0) = i, X(0) = u\big),$$
$$m_{ij}^+(u, q, a) = \mathbb{E}\big(\mathrm{e}^{-q(\tau_0^- \wedge \tau_a^+)} 1_{\{J(\tau_0^- \wedge \tau_a^+)=j, X(\tau_0^- \wedge \tau_a^+)=a\}} \mid J(0) = i, X(0) = u\big),$$

for $0 < u < a \le \infty$ and $i, j \in \mathcal{E}$. In essence, the functions m_{ij}^+ and m_{ij}^- characterize the distributional law of the exit times of X from the band $(0, a)$. Employing the law of total probability, it is straightforward to reinterpret m_{ij}^- and m_{ij}^- as

$$m_{ij}^-(u, q, a) = \mathbb{P}\big(\tau_0^- \wedge \tau_a^+ < e_q, J(\tau_0^- \wedge \tau_a^+) = j, X(\tau_0^- \wedge \tau_a^+) = 0 \mid J(0) = i, X(0) = u\big),$$
$$m_{ij}^+(u, q, a) = \mathbb{P}\big(\tau_0^- \wedge \tau_a^+ < e_q, J(\tau_0^- \wedge \tau_a^+) = j, X(\tau_0^- \wedge \tau_a^+) = a \mid J(0) = i, X(0) = u\big),$$

where e_q is an $\mathrm{Exp}(q)$ random variable independent of everything else; from now on, we adopt this reinterpretation.

In the existing literature, explicit solutions to m_{ij}^+ and m_{ij}^- only exist in very simple scenarios: either $q = 0$ and $\mathcal{E} = \{1\}$ (see, e.g., [23, Chapter 4]), or μ_i and σ_i are constant for all $i \in \mathcal{E}$ (see, e.g., [24]). Moreover, we are also interested in the present value of the expected ocupation times in state $j \in \mathcal{E}$ while below level $b \in (0, a)$, defined as

$$O_{ij}(u, q, a, b) = \mathbb{E}\bigg(\int_0^{\tau_0^- \wedge \tau_a^+} \mathrm{e}^{-qs} 1_{\{J(s)=j,\, 0<X(s)\le b\}} \,\mathrm{d}s \mid J(0) = i, X(0) = u \bigg).$$

Similarly to m_{ij}^- and m_{ij}^+, here we reinterpret O_{ij} as

$$O_{ij}(u, q, a, b) = \mathbb{E}\bigg(\int_0^{\tau_0^- \wedge \tau_a^+ \wedge e_q} 1_{\{J(s)=j,\, 0<X(s)\le b\}} \,\mathrm{d}s \mid J(0) = i, X(0) = u \bigg).$$

In this section we propose to exploit the approximations laid out in Theorem 3.2, along with the existing theory for multi-regime Markov-modulated Brownian motions, to provide an efficient approximation scheme based on computing

$$\widehat{m}_{ij}^-(u, q, a) = \mathbb{P}\big(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+ < e_q, \widehat{J}(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+) = j, \widehat{X}(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+) = 0 \mid \widehat{J}(0) = i, \widehat{X}(0) = u\big),$$
$$\widehat{m}_{ij}^+(u, q, a) = \mathbb{P}\big(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+ < e_q, \widehat{J}(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+) = j, \widehat{X}(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+) = a \mid \widehat{J}(0) = i, \widehat{X}(0) = u\big),$$
$$\widehat{O}_{ij}(u, q, a, b) = \mathbb{E}\bigg(\int_0^{\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+ \wedge e_q} 1_{\{\widehat{J}(s)=j,\, 0<\widehat{X}(s)\le b\}} \,\mathrm{d}s \mid \widehat{J}(0) = i, \widehat{X}(0) = u \bigg),$$

where $\widehat{\tau}_0^- = \inf\{t > 0 \colon \widehat{X}(t) \le 0\}$ and $\widehat{\tau}_a^+ = \inf\{t > 0 \colon \widehat{X}(t) \ge a\}$.

In the case that $\widehat{\tau}_0^-$ converges in probability to τ_0^- and that $\widehat{\tau}_a^+$ converges in probability to τ_a^+, and using the continuity of the paths of $\widehat{X}$ and X, it follows that $\widehat{m}_{ij}^-$ and $\widehat{m}_{ij}^+$ converge pointwise to m_{ij}^- and m_{ij}^+, respectively. Furthermore, if $q > 0$ then, by the dominated convergence theorem, $\widehat{O}_{ij}$ converges to O_{ij}.

Remark 4.1. In general, $\widehat{\tau}_0^-$ and $\widehat{\tau}_a^+$ may not converge to τ_0^- and τ_a^+ as $n \to \infty$, since hitting times are not continuous functionals of the paths. For example, consider the deterministic process $f(t) = -(t-1)1_{[1,\infty)}(t)$ and its approximation $f_\varepsilon(t) = f(t) + \varepsilon$, for $\varepsilon > 0$. The first-passage

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

time to $(-\infty, 0]$ is 0 for f, while it is greater than 1 for f_ε, thus demonstrating that convergence of first-passage times need not hold in general. This discrepancy can potentially be resolved by assuming that the process X immediately enters $(-\infty, 0)$ after τ_0^-, and immediately enters (a, ∞) after τ_a^+. However, verifying the validity of such assumptions can be nontrivial and may require sophisticated tools beyond the scope of our present framework. A more direct and practical alternative is to replace $\widehat{\tau}_0^-$ and $\widehat{\tau}_a^+$ with the first-passage times to $(-\infty, n^{-\delta_*+\varepsilon_1}]$ and $[a - n^{-\delta_*+\varepsilon_1}, \infty)$, respectively. In this case, convergence in probability of the approximating first-passage times to τ_0^- and τ_a^+ follows from Theorem 3.2. Henceforth, and for simplicity, we assume that $\widehat{\tau}_0^-$ and $\widehat{\tau}_a^+$ – defined as the first-passage times to $(-\infty, 0]$ and $[a, \infty)$, respectively – converge in probability to τ_0^- and τ_a^+. We emphasize, however, that replacing these with the first-passage times to $(-\infty, n^{-\delta_*+\varepsilon_1}]$ and $[a - n^{-\delta_*+\varepsilon_1}, \infty)$ guarantees convergence of the associated descriptors and does not compromise computational tractability.

While the functions $\widehat{m}_{ij}^-$ and $\widehat{m}_{ij}^+$ are not readily available in the existing literature, we can compute them by embedding independent copies of the excursion $(\widehat{J}_*, \widehat{X}_*) := \{(\widehat{J}(t), \widehat{X}(t)) : t \leq \widehat{\tau}_0^- \wedge \widehat{\tau}_a^+ \wedge e_q\}$ into a certain recurrent *queue* on the strip $[0, a]$ (that is, a process reflected on the level boundaries 0 and a). This procedure is similar to those exploited in [1, 36, 37] to provide finite-time ruin probabilities for multi-regime Markov-modulated risk processes or its subclasses. Below we spell out the details of our construction.

Let $\{(\widehat{J}_*^{\{\ell\}}, \widehat{X}_*^{\{\ell\}})\}_{\ell \geq 1}$ be a sequence of independent copies of $(\widehat{J}_*, \widehat{X}_*)$. Here we assume that the total length of each excursion $\widehat{X}_*^{\{\ell\}}$, say $T^{\{\ell\}}$, has finite mean; this trivially holds if $q > 0$, or if either $\widehat{\tau}_0^-$ or $\widehat{\tau}_a^+$ have a finite mean. Additionally, let our probability space support three independent sequences of Exp(1)-distributed independent random variables $\{\rho_{-M}^{\{\ell\}}\}$, $\{\rho_0^{\{\ell\}}\}$, and $\{\rho_M^{\{\ell\}}\}$. Based on the construction of multi-regime queues in [1, 21], here we define Y on a space grid $\{\zeta_m\}_{-M \leq m \leq M}$ (with $\zeta_{-M} = 0$, $\zeta_0 = u$, and $\zeta_M = a$) modulated by the jump process L that has a state space $\mathcal{E} \cup \{\partial_0\}$, both of which evolve as follows:

(i) On the time interval $[0, \nu_1^{\{1\}})$ with $\nu_1^{\{1\}} = T^{\{1\}}$, let (L, Y) coincide with $(\widehat{J}_*^{\{\ell\}}, \widehat{X}_*^{\{1\}})$.

(ii) At time $\nu_1^{\{1\}}$, one of three events happens:

- If $\widehat{X}^{\{1\}}(\nu_1^{\{1\}}) = 0$, let $\nu_2^{\{1\}} = \nu_1^{\{1\}} + \rho_{-M}^{\{1\}}$ and $(Y(t) = 0, L(t) = L(\nu_1^{\{1\}} -))$ for all $t \in [\nu_1^{\{1\}}, \nu_2^{\{1\}})$.
- If $\widehat{X}^{\{1\}}(\nu_1^{\{1\}}) = a$, let $\nu_2^{\{1\}} = \nu_1^{\{1\}} + \rho_M^{\{1\}}$ and $(Y(t) = a, L(t) = L(\nu_1^{\{1\}} -))$ for all $t \in [\nu_1^{\{1\}}, \nu_2^{\{1\}})$.
- If $\widehat{X}^{\{1\}}(\nu_1^{\{1\}}) \in (0, a)$, let $\nu_2^{\{1\}} = \nu_1^{\{1\}}$.

(iii) At time $\nu_2^{\{1\}}$, one of two events happens:

- If $Y(\nu_2^{\{1\}}) \leq u$, let Y increase at unit rate up to reaching level u, say, at time $\nu_3^{\{1\}}$. Let $L(t) = \partial_0$ for all $t \in [\nu_2^{\{1\}}, \nu_3^{\{1\}})$.
- If $Y(\nu_2^{\{1\}}) > u$, let Y decrease at unit rate up to reaching level u, say, at time $\nu_3^{\{1\}}$. Let $L(t) = \partial_0$ for all $t \in [\nu_2^{\{1\}}, \nu_3^{\{1\}})$.

(iv) Let $P_*^{\{1\}} = \nu_3^{\{1\}} + \rho_0\{1\}_u$ and $(Y(t) = u, L(t) = \partial_0)$ for all $t \in [\nu_3^{\{1\}}, \nu_4^{\{1\}})$.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

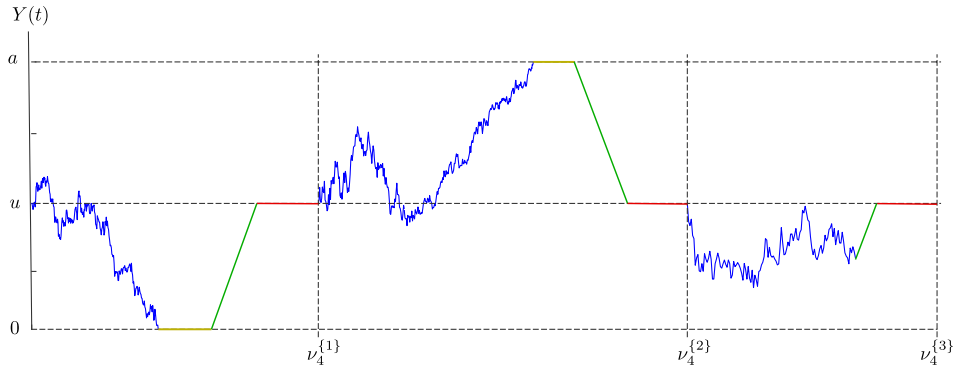


FIGURE 1. Queue model resulting from concatenating $\{\widehat{X}_*^{\{\ell\}}\}_{\ell\geq 1}$ using steps (i)–(iv), which are shown in the colors blue, mustard, green, and red, respectively.

(v) Repeat steps (i)–(iv) for $\ell = 2, 3, \ldots$, shifting the time accordingly in order to concatenate the excursions along with the increasing sequences $\{\nu_i^{\{\ell\}}\}_\ell$, $i \in \{1, 2, 3, 4\}$. This produces a process (L, Y) that regenerates at the epochs $\{\nu_4^{\{\ell\}}\}_{\ell\geq 1}$.

See Figure 1 for a graphical description of this construction.

As pointed out earlier, the process (L, Y) falls within the class of multi-regime Markov-modulated Brownian motion queueing models (see [1, 21] for details), which has a law characterized by the level-dependent matrices $(Q(x), R(x), S(x))_{0\leq x\leq a}$ with the following interpretation:

- $Q_{ij}(x)$ is the jump intensity of L from i to j while Y is in level x. We further specify the dependence on the level x by employing intensity matrices $\{Q^{(m)}\}_{-M+1\leq m\leq M}$ and $\{\tilde{Q}^{(m)}\}_{-M\leq m\leq M}$, where
$$Q(x) = \begin{cases} Q^{(m)} & \text{for} \quad \zeta_{m-1} < x < \zeta_m, \\ \tilde{Q}^{(m)} & \text{for} \quad x = \zeta_m. \end{cases}$$
- $R_{ii}(x)$ is the drift of Y at level x while L is in state i (by convention we let $R_{ij}(x) = 0$ for all $i \neq j$). We further specify this dependence on the level x by employing diagonal matrices $\{R^{(m)}\}_{-M+1\leq m\leq M}$ and $\{\tilde{R}^{(m)}\}_{-M\leq m\leq M}$, where
$$R(x) = \begin{cases} R^{(m)} & \text{for} \quad \zeta_{m-1} < x < \zeta_m, \\ \tilde{R}^{(m)} & \text{for} \quad x = \zeta_m. \end{cases}$$
- $S_{ii}(x)$ is the diffusion coefficient of Y at level x while L is in state i (by convention we let $S_{ij}(x) = 0$ for all $i \neq j$). We further specify the dependence on the level x by employing nonnegative diagonal matrices $\{S^{(m)}\}_{-M+1\leq m\leq M}$, where $S(x) = S^{(m)}$ for $\zeta_{m-1} \leq x < \zeta_m$.

Under the previous considerations, the corresponding matrices for our particular model with state space $\mathcal{E} \cup \{\partial_0\}$ take the form

$$Q^{(m)} = \begin{pmatrix} \widehat{\Lambda}(\zeta_{m-1}) - qI & q\mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix}, \qquad m \in \{-M+1, -M+2, \ldots, M\},$$
$$\tilde{Q}^{(m)} = \begin{pmatrix} \widehat{\Lambda}(\zeta_m) - qI & q\mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix}, \qquad m \in \{-M+1, -M+2, \ldots, M-1\} \setminus \{0\},$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

$$\tilde{Q}^{(-M)} = \tilde{Q}^{(M)} = \begin{pmatrix} -I & \mathbf{1} \\ \mathbf{0} & 0 \end{pmatrix},$$

$$\tilde{Q}^{(0)} = \begin{pmatrix} \widehat{\Lambda}(\zeta_0) - qI & q\mathbf{1} \\ e_i^\top & -1 \end{pmatrix},$$

$$R^{(m)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\mu}_i(\zeta_{m-1})\} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \qquad m \in \{-M+1, -M+2, \ldots, 0\},$$

$$R^{(m)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\mu}_i(\zeta_{m-1})\} & \mathbf{0} \\ \mathbf{0} & -1 \end{pmatrix}, \qquad m \in \{1, 2, \ldots, M\},$$

$$\tilde{R}^{(m)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\mu}_i(\zeta_m)\} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \qquad m \in \{-M+1, -M+2, \ldots, -1\},$$

$$\tilde{R}^{(m)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\mu}_i(\zeta_m)\} & \mathbf{0} \\ \mathbf{0} & -1 \end{pmatrix}, \qquad m \in \{1, -M+2, \ldots, M-1\},$$

$$\tilde{R}^{(-M)} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix},$$

$$\tilde{R}^{(M)} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -1 \end{pmatrix},$$

$$\tilde{R}^{(0)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\mu}_i(\zeta_0)\} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix},$$

$$S^{(m)} = \begin{pmatrix} \mathrm{diag}\{\widehat{\sigma}_i(\zeta_{m-1})\} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \qquad m \in \{-M+1, -M+2, \ldots, M\},$$

where $e_i^\top$ denotes the ith canonical row vector, and $\mathrm{diag}\{a_i\}$ denotes the diagonal matrix with the elements $\{a_1, \ldots, a_p\}$ filling the diagonal. In [21], the authors provide an efficient algorithm to compute the steady-state distribution of (L, Y). In particular, using their method we are able obtain:

- the steady-state probability atoms for levels ζ_{-M} and ζ_M while in the states $\mathcal{E}$ in row vector form, say p_{-M} and p_M, where
$$(p_{-M})_j = \lim_{t\to\infty} \mathbb{P}(L(t) = j, Y(t) = 0), \qquad (p_M)_j = \lim_{t\to\infty} \mathbb{P}(L(t) = j, Y(t) = a);$$
- the steady-state probability atoms for level ζ_0 while in the state ∂_0, say p_0, where
$$p_0 = \lim_{t\to\infty} \mathbb{P}(L(t) = \partial_0, Y(t) = u);$$
- the steady-state probability distribution function while in state $j \in \mathcal{E}$ over (0,a), say F_j, where $F_j(b) = \lim_{t\to\infty} \mathbb{P}(L(t) = j, 0 < Y(t) \le b)$, $b \in (0, a)$.

Below we link these quantities with the first-passage probabilities $\widehat{m}_{ij}^+$ and $\widehat{m}_{ij}^-$, as well as with the expected occupation times $\widehat{O}_{ij}$.

Theorem 4.1. *Suppose that* $\mathbb{E}(\widehat{\tau}_0^- \wedge \widehat{\tau}_a^+ \wedge e_q \mid \widehat{J}(0) = i, \widehat{X}(0) = u) < \infty$. *Then*

$$\widehat{m}_{ij}^-(u, q, a) = \frac{(p_{-M})_j}{p_0}, \qquad \widehat{m}_{ij}^+(u, q, a) = \frac{(p_M)_j}{p_0}. \tag{4.1}$$

Moreover,

$$\widehat{O}_{ij}(u, q, a, b) = \frac{F_j(b)}{p_0}, \quad b \in (0, a). \tag{4.2}$$

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

Proof. The process (L, Y) regenerates at the epochs $\{\nu_4^{\{\ell\}}\}_{\ell\geq 1}$, all of which have a (common) finite first moment. This in turn implies that (L, Y) is positive recurrent, so by [20, Theorem 1],

$$p_M\mathbf{1} = \frac{\mathbb{E}\big(\int_0^{\nu_4^{\{1\}}} 1_{\{Y(s)=u, L(s)\in\mathcal{E}\}}\,\mathrm{d}s\big)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)} = \frac{\mathbb{P}\big(Y\big(\nu_1^{\{1\}}\big)=u,\, L\big(\nu_1^{\{1\}}\big)\in\mathcal{E}\big)\mathbb{E}\big(\rho_0^{\{1\}}\big)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)} = \frac{\widehat{m}^{+}(u,q,a)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)},$$

where we used that $\mathbb{E}\big(\rho_0^{\{1\}}\big)=1$. Employing similar arguments, we get

$$p_{-M}\mathbf{1} = \frac{\widehat{m}^{-}(u,q,a)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)}, \qquad p_0 = \frac{1}{\mathbb{E}\big(\nu_4^{\{1\}}\big)},$$

so that (4.1) follows. Additionally,

$$F_j(b) = \frac{\mathbb{E}\big(\int_0^{\nu_4^{\{1\}}} 1_{\{L(s)=j,\,0<Y(s)\leq b\}}\,\mathrm{d}s\big)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)} = \frac{\mathbb{E}\big(\int_0^{\widehat{\tau}_0^{-}\wedge\widehat{\tau}_a^{+}\wedge e_q} 1_{\{\widehat{J}(s)=j,\,0<\widehat{X}(s)\leq b\}}\,\mathrm{d}s\big)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)} = \frac{\widehat{O}_{j,b}(u,q,a)}{\mathbb{E}\big(\nu_4^{\{1\}}\big)},$$

which in turn implies (4.2). □

In short, with the help of the algorithms developed in [21], we are able to efficiently compute first-passage probabilities and their expected occupation times for $(\widehat{J}, \widehat{X})$ using Theorem 4.1, which in turn approximates the first-passage probabilities and their expected occupation times for (J, X) by virtue of Theorem 3.2. Below we explore a couple of synthetic examples.

Example 4.1. Consider a three-state hybrid SDE (X, J) evolving in the space band (0,1) with parameters

$$\mu_1(x)=0.5, \qquad \mu_2(x)=0.5(1-x), \qquad \mu_3(x)=0.5(1-x)^2, \qquad \sigma_1=\sigma_2=\sigma_3=1.$$

In essence, the process X has a force pushing it upwards while J is in any of the three states. The force is weakest while in state 3 and strongest while in 1; note that the difference between regimes is accentuated as X gets closer to level 1. In the context of actuarial science, this may, e.g., describe the case of the surplus process of an insurance company with three possible occurring regimes, all with positive drift, but some being more responsive with premium reductions when the surplus approaches 1. This can be described, for instance, by an intensity matrix function of the form

$$\mathbf{\Lambda}(x) = 10\times\begin{pmatrix} -x & x & 0 \\ (1-x) & -1 & x \\ 0 & 1-x & -(1-x) \end{pmatrix}. \tag{4.2}$$

Employing the MATLAB coded provided at https://www.hit.bme.hu/~ghorvath based on [21] together with Theorem 4.1, we are able to compute the first-passage probabilities and expected

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

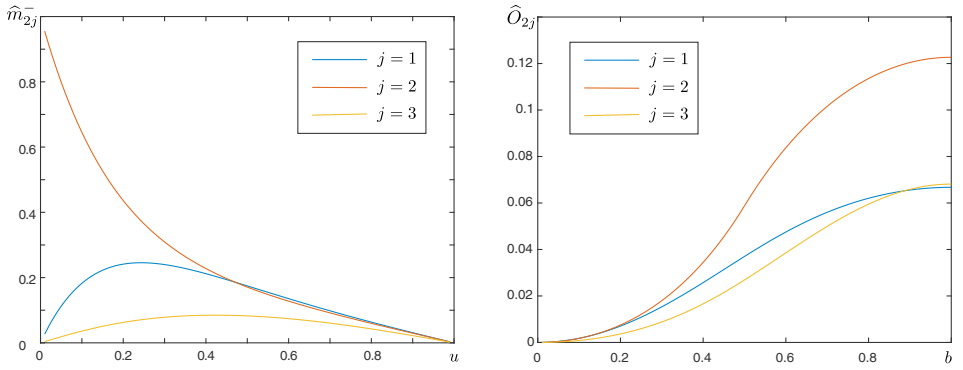


FIGURE 2. Left: Plots of $\widehat{m}^{-}_{2j}$ as a function of u. Right: Plots of $\widehat{O}_{2j}$ as a function of b with $u = 0.5$. Each plot contains the cases $j = 1, 2, 3$.

occupation times of the multi-regime Markov-modulated Brownian motion $(\widehat{J}, \widehat{X})$ that approximates (J, X). We perform such computations for $q = 0$, $a = 1$, $i = 2$, and $M = 50$, with the results shown in Figure 2.

For small values of u, downcrossings of level 0 are most likely while in state 2; this can be explained by noting that the random oscillations while in state 2 (the initial state of J) can produce a downcrossing before a switching even occurs. For medium values of u, the probabilities of downcrossing level 0 while in states 1 and 2 are comparable, both of which are considerably larger than that corresponding to state 3; this is due to the state-dependent switching behaviour of Λ, which favours states 1 and 2 while the level is low, and 2 and 3 while the level is high. For high values of u, the probability of downcrossing 0 is uniformly low for all states. On the other hand, we can observe that the expected occupation time while in state 2 is larger than that in states 1 or 2 uniformly over all u; this is a consequence of state 2 being the initial state of the system.

While Theorem 3.2 guarantees that $\widehat{m}^{-}_{ij}$ converges to m^{-}_{ij} as the space grid becomes denser, in Figure 3 we (empirically) show how fast this convergence happens. As we can appreciate from the plot, convergence is quickly achieved, with differences between the cases being essentially negligible for $M \geq 40$.

Example 4.2. Consider now the same parameters as in Example 4.1, but take $\mu_3(x) = -0.5x^2$ and $\sigma_3 = 0$. While this alternative scenario does not necessarily have an immediate practical interpretation, it does help to investigate some of the consequences of having a state that lacks any random noise behaviour. From Figure 4, we note that the downcrossing probabilities for states 1 and 3 are larger than those in Example 4.1; this is because we now have one of the states pushing the level process downwards. However, note that the probability of downcrossing 0 while in state 3 is null; this is because μ_3 approaches 0 as the level gets closer to 0, meaning that the process simply cannot cross level 0 while in state 3. Moreover, we note that for large values of b, the expected occupation time while in state 3 is larger than in states 1 or 2; this suggests that before exiting the band (0,1), the process switches to and stays in state 3 for a larger amount of time than in Example 4.1.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

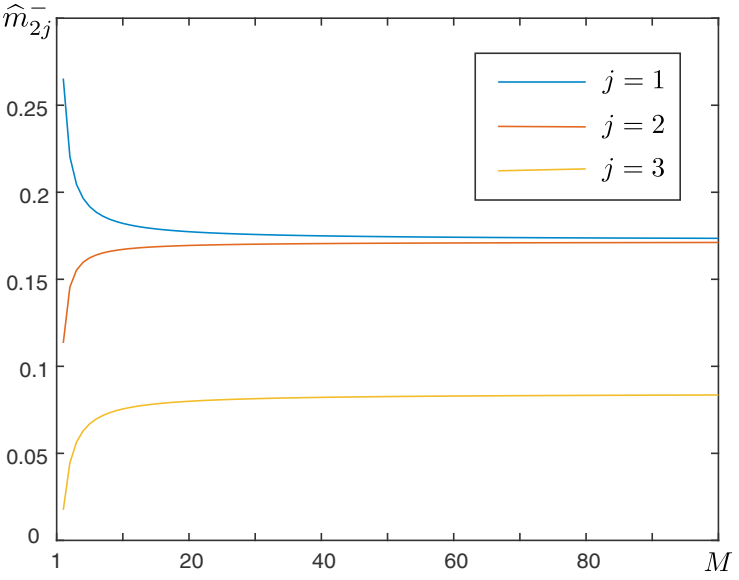


FIGURE 3. Plots of $\widehat{m}_{2j}^{-}$ as a function of M for the cases $j = 1, 2, 3$.

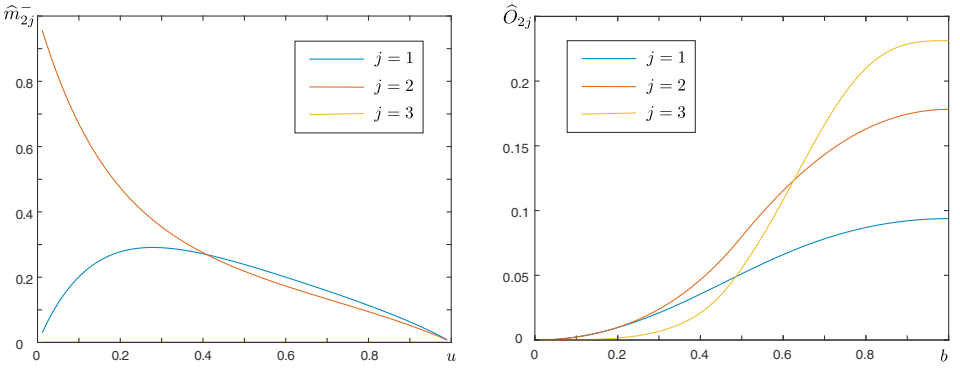


FIGURE 4. Left: Plots of $\widehat{m}_{2j}^{-}$ as a function of u. Right: Plots of $\widehat{O}_{2j}$ as a function of b with $u = 0.5$. Each plot contains the cases $j = 1, 2, 3$.

5. Summary and extensions

Under Assumptions 2.1, 2.2, 3.1 3.2, and 3.3, we considered the class of hybrid SDEs and their space-grid approximations, the latter belonging to a family of multi-regime Markov-modulated Brownian motions. We provided a rigorous proof of their pathwise convergence, which holds in a weak sense uniformly over increasing compact intervals. As an application of our convergence result, we considered the challenging problem of identifying first-passage probabilities and expected occupation times for solutions of hybrid SDEs, for which we suggest an approximate and computationally efficient answer that employs developments in the area of multi-regime Markov-modulated queues.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

We point out that applications of our convergence result may also be used to approximate other (and possibly more complex) descriptors of hybrid SDEs. For instance, we can easily build hybrid SDEs with phase-type jumps by means of the fluidization method [7]. Other modifications and extensions that are straightforward to implement in our framework are the Omega model [2, 18], the Erlangian approximations for finite-time probabilities of ruin [6], and Parisian ruin problems with phase-type clocks [10].

Furthermore, while it is possible to adapt the methods presented here to a higher-dimensional setting by appropriately modifying the norms and proofs, we chose to focus on the one-dimensional case for two main reasons. First, higher-dimensional arguments introduce significant notational complexity, which may obscure the key ideas and results. Second, the existing literature on algorithmic results for multi-regime Markov-modulated Brownian motions is limited to the one-dimensional setting, which means that extending to higher dimensions would have limited immediate applicability. Future work may explore these extensions in contexts where they become practically relevant.

Acknowledgements

The authors are grateful to the anonymous referees for their valuable comments and suggestions, which helped to improve the quality of the paper.

Funding Information

The authors would like to acknowledge financial support from the Swiss National Science Foundation Project 200021_191984.

Competing Interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AKAR, N., GURSOY, O., HORVATH, G. AND TELEK, M. (2021). Transient and first passage time distributions of first- and second-order multi-regime Markov fluid queues via ME-fication. *Methodol. Comput. Appl. Prob.* **23**, 1257–1283.

[2] ALBRECHER, H., GERBER, H. U. AND SHIU, E. S. (2011). The optimal dividend barrier in the Gamma–Omega model. *Europ. Actuarial J.* **1**, 43–55.

[3] ASMUSSEN, S. (1995). Stationary distributions for fluid flow models with or without Brownian noise. *Commun. Statist. Stoch. Models* **11**, 21–49.

[4] ASMUSSEN, S. (2003). *Applied Probability and Queues*, Vol. 2. Springer, New York.

[5] ASMUSSEN, S. AND ALBRECHER, H. (2010). *Ruin Probabilities*. World Scientific, Singapore.

[6] ASMUSSEN, S., AVRAM, F. AND USABEL, M. (2002). Erlangian approximations for finite-horizon ruin probabilities. *ASTIN Bull.* **32**, 267–281.

[7] BADESCU, A., BREUER, L., DA SILVA SOARES, A., LATOUCHE, G., REMICHE, M. A. AND STANFORD, D. (2005). Risk processes analyzed as fluid queues. *Scand. Actuarial J.* **2005**, 127–141.

[8] BASS, R. F. AND PARDOUX, E. (1987). Uniqueness for diffusions with piecewise constant coefficients. *Prob. Theory Relat. Fields* **76**, 557–572.

[9] BLADT, M. AND NIELSEN, B. F. (2017). *Matrix-Exponential Distributions in Applied Probability*. Springer, New York.

[10] BLADT, M., NIELSEN, B. F. AND PERALTA, O. Parisian types of ruin probabilities for a class of dependent risk-reserve processes. *Scand. Actuarial J.* 2019, 32–61.

[11] BREUER, L. (2012). Occupation times for Markov-modulated Brownian motion. *J. Appl. Prob.* **49**, 549–565.

[12] CSÖRGÖ, M. AND HORVÁTH L. (1988). Rate of convergence of transport processes with an application to stochastic differential equations. *Prob. Theory Relat. Fields* **78**, 379–387.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press

[13] CSÖRGO, M. AND RÉVÉSZ, P. (2014). *Strong Approximations in Probability and Statistics*. Academic Press, Cambridge, MA.
[14] D'AURIA, B., IVANOVS, J., KELLA, O. AND MANDJES, M. (2012). Two-sided reflection of Markov-modulated Brownian motion. *Stoch. Models* **28**, 316–332.
[15] DE ACOSTA, A. (1982). Invariance principles in probability for triangular arrays of *B*-valued random vectors and some applications. *Ann. Prob.* **10**, 346–373.
[16] ELLIOTT, R. J., SIU, T. K., CHAN, L. AND LAU, J. W. (2007). Pricing options under a generalized Markov-modulated jump-diffusion model. *Stoch. Anal. Appl.* **25**, 821–843.
[17] FISCHER, W. AND MEIER-HELLSTERN, K. (1993). The Markov-modulated poisson process (MMPP) cookbook. *Performance Evaluation* **18**, 149–171.
[18] GERBER, H. U., SHIU, E. S. AND YANG, H. (2012). The Omega model: From bankruptcy to occupation times in the red. *Europ. Actuarial J.* **2**, 259–272.
[19] GLYNN, P. W. (1987). Upper bounds on Poisson tail probabilities. *Operat. Res. Lett.* **6**, 9.
[20] GLYNN, P. W. (1994). Some topics in regenerative steady-state simulation. *Acta Appl. Math.* **34**, 225–236.
[21] HORVÁTH, G. AND TELEK, M. (2017). Matrix-analytic solution of infinite, finite and level-dependent second-order fluid models. *Queueing Systems* **87**, 325–343.
[22] IKEDA, N. AND WATANABE, S. (2014). *Stochastic Differential Equations and Diffusion Processes*. Elsevier, Amsterdam.
[23] ITÔ, K. AND MCKEAN, H. P. (2012). *Diffusion Processes and their Sample Paths*. Springer, New York.
[24] IVANOVS, J. (2010). Markov-modulated Brownian motion with two reflecting barriers. *J. Appl. Prob.* **47**, 1034–1047.
[25] KRYLOV, N. V. AND RÖCKNER, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Prob. Theory Relat. Fields* **131**, 154–196.
[26] LAMPERTI, J. (1964). A simple construction of certain diffusion processes. *J. Math. Kyoto Univ.* **4**, 161–170.
[27] LEOBACHER, G. AND SZÖLGYENYI, M. (2017). A strong order $1/2$ method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Prob.* **27**, 2383–2418.
[28] MANDJES, M., MITRA, D. AND SCHEINHARDT, W. (2003). Models of network access using feedback fluid queues. *Queueing Systems* **44**, 365–398.
[29] NGUYEN, D. H. AND YIN, G. (2016). Modeling and analysis of switching diffusion systems: Past-dependent switching with a countable state space. *SIAM J. Control Optim.* **54**, 2450–2477.
[30] NGUYEN, G. T. AND PERALTA, O. (2021). Wong–Zakai approximations with convergence rate for stochastic differential equations with regime switching. Preprint, arXiv:2101.03250.
[31] NGUYEN, G. T. AND PERALTA, O. (2022). Rate of strong convergence to Markov-modulated Brownian motion. *J. Appl. Prob.* **59**, 1–16.
[32] PRABHU, N. AND ZHU, Y. (1989). Markov-modulated queueing systems. *Queueing Systems* **5**, 215–245.
[33] SKOROKHOD, A. V. (2009). *Asymptotic Methods in the Theory of Stochastic Differential Equations*. American Mathematical Society, Providence, RI.
[34] TRIVEDI, K. S. AND BOBBIO, A. (2017). *Reliability and Availability Engineering: Modeling, Analysis, and Applications*. Cambridge University Press.
[35] VAN DIJK, N. M., VAN BRUMMELEN, S. P. J. AND BOUCHERIE, R. J. (2018). Uniformization: Basics, extensions and applications. *Performance Evaluation* **118**, 8–32.
[36] VAN HOUDT, B. AND BLONDIA, C. (2005). Approximated transient queue length and waiting time distributions via steady state analysis. *Stoch. Models* **21**, 725–744.
[37] YAZICI, M. A. AND AKAR, N. (2017). The finite/infinite horizon ruin problem with multi-threshold premiums: A Markov fluid queue approach. *Ann. Operat. Res.* **252**, 85–99.
[38] YIN, G., MAO, X., YUAN, C. AND CAO, D. (2010). Approximation methods for hybrid diffusion systems with state-dependent switching processes: Numerical algorithms and existence and uniqueness of solutions. *SIAM J. Math. Anal.* **41**, 2335–2352.
[39] YIN, G. G. AND ZHU, C. (2009). *Hybrid Switching Diffusions: Properties and Applications*. Springer, New York.
[40] ZHANG, S. Q. (2020). Regime-switching diffusion processes: Strong solutions and strong Feller property. *Stoch. Anal. Appl.* **38**, 97–123.

https://doi.org/10.1017/jpr.2025.10050 Published online by Cambridge University Press