Optimal Dividend Problem: Asymptotic Analysis[*]

Asaf Cohen[†] and Virginia R. Young[†]

Abstract. We revisit the classical problem of optimal payment of dividends and determine the degree to which the diffusion approximation serves as a valid approximation of the classical risk model for this problem. Our results parallel some of those in Bäuerle [*Math. Finance*, 14 (2004), pp. 99–113], but we obtain sharper results because we use a different technique for obtaining them. Specifically, Bäuerle uses probabilistic techniques and relies on convergence in distribution of the underlying processes. By contrast, we use comparison results from the theory of differential equations, and these methods allow us to determine the rate of convergence of the value functions in question.

Key words. optimal dividend strategy, Cramer–Lundberg risk process, diffusion approximation, approximation error

AMS subject classifications. 91B30, 91G80, 93E20, 45J05, 60G99, 90B20

DOI. 10.1137/20M1354738

1. Introduction. A long-standing problem in insurance mathematics is optimal payment of dividends; see, for example, the survey by Avanzi [1]. In this paper, we are concerned about the degree to which the diffusion approximation serves as a valid approximation of the classical risk model when optimizing dividend payments. Gerber, Shiu, and Smith [16] address approximations to the dividend problem. Also, Bäuerle [3] considers the scaled dividend problem and proves that, as the scaling factor increases without bound, the value function converges to the one under the diffusion approximation.

We, now, compare the model and techniques used and the results obtained by Bäuerle [3] with the corresponding items in our work. Bäuerle [3] considers the case for which the rate of dividend payments is bounded, which corresponds to a classical continuous-time control problem. While we use the same diffusion scaling, we, on the other hand, do not restrict the dividend rate to be bounded, which leads to a singular control problem. Moreover, Bäuerle [3] uses probabilistic techniques and relies on convergence in distribution of the underlying processes under Skorokhod's J1 topology. This procedure has two main components: (1) showing that any limit point on any arbitrary sequence of controls in the prelimit problem does not perform better than the value function, and (2) showing that a candidate control for the prelimit problem attains the value function. For the first component, the proof uses tightness arguments, heavily utilizing the continuity of the underlying processes and the uniform

[†]Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 USA (shloshim@gmail.com, https://sites.google.com/site/asafcohentau/; vryoung@umich.edu).

boundedness of the control. Due to the singular control considered here, the compactness and tightness arguments used by Bäuerle [3] are not valid for our work.[1] By contrast, we rely on the closed-form expression of the limiting value function and use comparison results from the theory of differential equations; these methods allow us to determine the rate of convergence of the value functions in question; see Theorem 3.4. Another issue we address is the rate of convergence of the difference between the value function in the prelimit problem and the payoff function in the prelimit problem using the optimal threshold from the limiting problem; see Theorem 3.8. The importance of this result stems from the fact that the latter threshold admits a closed form, unlike for the prelimit problem. Finally, Bäuerle [3] also includes proportional reinsurance, but we omit reinsurance in the interest of simplicity.

The background for the comparison principle we are using is introduced in Cohen and Young [10]. In that paper, the authors provide the rate of convergence of the probability of ruin in the Cramér–Lundberg model to its diffusion approximation. The present paper shows that this method can be elevated from the uncontrolled problem to an optimal control problem, which on top of this is a *singular* control problem. Additionally, the comparison principles enables us to compare the value functions for different policies.

The remainder of the paper is organized as follows. In section 2, we present the Cramér-Lundberg (CL) model and state results from Azcue and Muler [2] that we use to bound our value function. Then, in section 3, we scale the CL model and show that, as the scaling factor increases without bound, the resulting value function converges to the one under the diffusion approximation, and we determine the rate of that convergence. In that section, we also show that, if the insurer uses the optimal strategy under the diffusion approximation but for the scaled CL risk model, then doing so is ε-optimal, and we specify the order of ε relative to the scaling factor.

2. Classical risk model and Azcue and Muler [2].

2.1. CL model with dividends. Consider an insurer whose surplus process $U = \{U_t\}_{t \geq 0}$ before paying dividends is described by a CL model, that is, the insurer receives premium income at a constant rate c and pays claims according to a compound Poisson process. Specifically,

$$U_t = x + ct - \sum_{i=1}^{N_t} Y_i \tag{2.1}$$

in which $U_0 = x \geq 0$ is the initial surplus, $N = \{N_t\}_{t \geq 0}$ is a homogeneous Poisson process with intensity $\lambda > 0$, and the claim sizes $Y_1, Y_2, \ldots$ are independent and identically distributed,

---

[1]To bypass this issue, one may use the *time-stretching method*, introduced by Meyer and Zheng in [23] and extensively used by Kurtz [20, 18, 19], Kushner and Martins [22, 21], Budhiraja and Ghosh [4] and Budhiraja and Ross [5, 6], Costantini and Kurtz [11, 12], and Cohen [7, 8]. The basic idea of this method is that time is stretched in accordance with the singular controls, so that the stretched underlying processes are uniformly Lipschitz continuous. Hence, tightness is attained under the J1 topology. Finally, time is shrunk in accordance with the limiting stretched control. The payoff/cost of the shrunken processes are, then, compared with the value function. Recently, Cohen [9] showed that by working with the weak-M1 topology instead of the J1 topology, the time transformations are unnecessary since they are embedded in the definition of the parametric representation, which define the weak-M1 topology. Hence, one may pursue a probabilistic proof under this topology to get convergence. However, this result does not provide a rate of convergence.

positive random variables, independent of N. All random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the natural filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ induced by the random variables.

Let F_Y denote the common cumulative distribution function of $\{Y_i\}_{i \in \mathbb{N}}$, and assume that Y has finite moment generating function $M_Y(s) = \mathbb{E}(e^{Ys})$ for s in a neighborhood of 0; thus, $\mathbb{E}(Y^k) < \infty$ for $k = 1, 2, \ldots$. Finally, assume that the premium rate c satisfies $c > \lambda \mathbb{E} Y$, and write $c = (1 + \theta)\lambda \mathbb{E} Y$ with positive relative risk loading $\theta > 0$.

The insurer pays dividends to its shareholders according to a process $D = \{D_t\}_{t \geq 0}$ in which D_t equals the cumulative dividends paid on or before time t, with $D_{0-} = 0$. A dividend strategy D is *admissible* if D is nondecreasing and is predictable with respect to the filtration $\mathbb{F}$.

The surplus process $X = \{X_t\}_{t \geq 0}$ after paying dividends is given by

$$(2.2) \qquad X_t = x + ct - \sum_{i=1}^{N_t} Y_i - D_t$$

in which $X_0 = x \geq 0$ is the initial surplus. Define the time of ruin τ by

$$(2.3) \qquad \tau = \inf\{t \geq 0 : X_t < 0\}.$$

The insurer seeks to maximize the expected payoff of discounted dividends between now and the time of ruin, with corresponding value function V defined by

$$(2.4) \qquad V(x) = \sup_D \mathbb{E}\left[\int_0^\tau e^{-\delta t} dD_t \,\middle|\, X_0 = x\right]$$

in which $\delta > 0$ is the discount rate, and the supremum is taken over admissible dividend strategies.

Gerber [14] shows that the optimal dividend strategy for the problem in (2.4) is a *band* strategy. A band strategy reduces to a *barrier* strategy if the initial surplus is less than the lowest band or if claim sizes are exponentially distributed. From Theorem 2.45 of Schmidli [24], the value function V is the minimal nonnegative solution of the following integro-differential variational inequality on $\mathbb{R}^+$:

$$(2.5) \qquad \min\left[(\lambda + \delta)v(x) - cv_x(x) - \lambda \int_0^x v(x-y) dF_Y(y),\; v_x(x) - 1\right] = 0.$$

Furthermore, Theorem 2.39 of Schmidli [24] states that V is differentiable from the left and from the right on $(0, \infty)$, and (2.5) holds separately for both left and right derivatives.

2.2. Results from Azcue and Muler [2].

We look for bounds for the value function V as sub- and super-solutions of (2.5), after we scale the CL model in section 3.1. To that end, define the operator F, acting on $u \in \mathcal{C}^1(\mathbb{R}^+)$, by the variational inequality in (2.5), that is,

$$(2.6) \qquad \begin{aligned} &F\big(x, u(x), u_x(x), u(\cdot)\big) \\ &= \min\left[(\lambda + \delta)u(x) - cu_x(x) - \lambda \int_0^x u(x-y) dF_Y(y),\; u_x(x) - 1\right]. \end{aligned}$$

We say that a function $u \in \mathcal{C}^1(\mathbb{R}^+)$ is a *subsolution* of $F = 0$ if

$$F\big(x, u(x), u_x(x), u(\cdot)\big) \le 0$$

for all $x \ge 0$. Similarly, we say that a function $v \in \mathcal{C}^1(\mathbb{R}^+)$ is a *supersolution* of $F = 0$ if

$$F\big(x, v(x), v_x(x), v(\cdot)\big) \ge 0$$

for all $x \ge 0$.

We state results from sections 4 and 5 of Azcue and Muler [2] as they apply to the model in this paper. They state their results for viscosity sub- and super-solutions because they control surplus via reinsurance; however, their results also apply to our no-reinsurance model with classical sub- and super-solutions. First, Azcue and Muler [2] prove a comparison result for functions that satisfy the following conditions:

1. $u : \mathbb{R}^+ \to \mathbb{R}$ is locally Lipschitz.
2. If $0 \le x < y$, then $u(y) - u(x) \ge y - x$.
3. There exists a constant $k > 0$ such that $u(x) \le x + k$ for all $x \ge 0$.

They note that the value function in (2.4) satisfies these three conditions.

Proposition 4.2 of Azcue and Muler [2] shows that if u is a subsolution and if v is a supersolution of $F = 0$, both satisfying Conditions 1, 2, and 3 with $u(0) \le v(0)$, then $u \le v$ on $\mathbb{R}^+$. Because the value function is a solution of $F = 0$ and, hence, a supersolution of $F = 0$, we will use this result in section 3.2 to find a lower bound of the value function.

Proposition 5.1 of Azcue and Muler [2] shows that if v is an absolutely continuous super-solution of $F = 0$ satisfying Condition 3, then $V \le v$ on $\mathbb{R}^+$. We will use this result in section 3.2 to find an upper bound of the value function.

3. Asymptotic analysis.

3.1. Scaled model and diffusion approximation.
Next, we scale the CL model by $n > 0$, as in Cohen and Young [10]. In the scaled system, define $\lambda_n = n\lambda$, so n large is essentially equivalent to λ large. Scale the claim severity by defining $Y_n = Y/\sqrt{n}$; thus, the variance of total claims during $[0, t]$ is invariant under the scaling, that is, $\lambda_n \mathbb{E}(Y_n^2) = \lambda \mathbb{E}(Y^2)$ for all $n > 0$. Finally, define the premium rate by $c_n = c + (\sqrt{n} - 1)\lambda \mathbb{E}Y$; thus, $c_n - \lambda_n \mathbb{E}Y_n = c - \lambda \mathbb{E}Y$ is also invariant under the scaling. We can also write $c_n = (\sqrt{n} + \theta)\lambda \mathbb{E}Y$ in which $c = (1 + \theta)\lambda \mathbb{E}Y$; moreover, we can write $c_n = (1 + \theta_n)\lambda_n \mathbb{E}Y_n$ in which $\theta_n = \theta/\sqrt{n}$. The diffusion approximation of the scaled surplus process before dividends is, therefore,

$$(3.1) \qquad \big(c_n - \lambda_n \mathbb{E}Y_n\big)dt + \sqrt{\lambda_n \mathbb{E}(Y_n^2)}\, dB_t = (c - \lambda \mathbb{E}Y)dt + \sqrt{\lambda \mathbb{E}(Y^2)}\, dB_t$$

$$= \theta \lambda \mathbb{E}Y\, dt + \sqrt{\lambda \mathbb{E}(Y^2)}\, dB_t$$

for some standard Brownian motion $B = \{B_t\}_{t \ge 0}$. Note that the diffusion approximation of the scaled CL model is *independent* of n. See Iglehart [17], Bäuerle [3], Gerber, Shiu, and Smith [16], and Schmidli [25] for more information about this scaling.

Let V_n denote the value function under the scaled CL model. We wish to bound V_n by modifying V_D via functions of order $\mathcal{O}(n^{-1/2})$ in which V_D is the value function when

uncontrolled surplus follows the diffusion approximation in (3.1). Thus, first, we digress to compute V_D, which uniquely solves the following free-boundary problem:

$$(3.2) \qquad \begin{cases} \delta v(x) = \theta \lambda \mathbb{E} Y v_x(x) + \dfrac{1}{2} \lambda \mathbb{E}(Y^2) v_{xx}(x), & 0 \leq x \leq b_D, \\ v(0) = 0, \quad v_x(b_D) = 1, \quad v_{xx}(b_D) = 0, \end{cases}$$

with $V_D(x) = V_D(b_D) + (x - b_D)$ for $x > b_D$. See Gerber and Shiu [15] for a thorough analysis of V_D's problem. Via a straightforward application of techniques from ordinary differential equations, we obtain

$$(3.3) \qquad V_D(x) = \begin{cases} \dfrac{e^{\gamma_1 x} - e^{-\gamma_2 x}}{\gamma_1 e^{\gamma_1 b_D} + \gamma_2 e^{-\gamma_2 b_D}}, & 0 \leq x \leq b_D, \\ V_D(b_D) + (x - b_D), & x > b_D, \end{cases}$$

in which $0 < \gamma_1 < \gamma_2$ are given by

$$(3.4) \qquad \gamma_1 = \frac{1}{\lambda \mathbb{E}(Y^2)} \left[-\theta \lambda \mathbb{E} Y + \sqrt{(\theta \lambda \mathbb{E} Y)^2 + 2 \delta \lambda \mathbb{E}(Y^2)} \right]$$

and

$$(3.5) \qquad \gamma_2 = \frac{1}{\lambda \mathbb{E}(Y^2)} \left[\theta \lambda \mathbb{E} Y + \sqrt{(\theta \lambda \mathbb{E} Y)^2 + 2 \delta \lambda \mathbb{E}(Y^2)} \right],$$

and the free boundary b_D equals

$$(3.6) \qquad b_D = \frac{2}{\gamma_1 + \gamma_2} \ln \left(\frac{\gamma_2}{\gamma_1} \right).$$

By using the expression for b_D in (3.6), we rewrite V_D as follows:

$$(3.7) \qquad V_D(x) = \begin{cases} \dfrac{1}{\gamma_1 + \gamma_2} \left(\dfrac{\gamma_1}{\gamma_2} \right)^{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}} \left(e^{\gamma_1 x} - e^{-\gamma_2 x} \right), & 0 \leq x \leq b_D, \\ \dfrac{\theta \lambda \mathbb{E} Y}{\delta} + (x - b_D), & x > b_D. \end{cases}$$

From the second line in (3.7), we observe

$$(3.8) \qquad V_D(b_D) = \frac{\theta \lambda \mathbb{E} Y}{\delta},$$

the present value of a continuous perpetuity, discounted at rate δ, paying at the rate $\theta \lambda \mathbb{E} Y$, the risk loading in the premium. Also, $V_D(x) > x$ for all $x > 0$, from which it follows that

$$(3.9) \qquad \frac{\theta \lambda \mathbb{E} Y}{\delta} > b_D,$$

an inequality that will be useful later.

Remark 3.1. Because the diffusion in (3.1) approximates the CL risk process in (2.1) with λ, Y, and c replaced by λ_n, Y_n, and c_n, respectively, researchers often say that V_D approximates V_n. In Theorem 3.4 in the next section, we quantify the degree to which V_D approximates V_n.

3.2. Approximating V_n by V_D to order $\mathcal{O}(n^{-1/2})$. In this section, we bound V_n by modifying V_D via functions of order $\mathcal{O}(n^{-1/2})$ and by using Propositions 4.2 and 5.1 of Azcue and Muler [2], as they apply to the scaled problem. Note that V_D plus or minus a constant satisfies the three conditions of Azcue and Muler [2] that we list in section 2.2.

Throughout this section, let F_n denote the operator in (2.6), with c, λ, and Y replaced by c_n, λ_n, and Y_n, respectively, and write F_n as follows:

$$F_n\big(x, u(x), u_x(x), u(\cdot)\big) = \min\Big\{G_n\big(x, u(x), u_x(x), u(\cdot)\big),\, u_x(x) - 1\Big\}$$

in which the operator G_n is as defined by

(3.10)
$$\begin{aligned} G_n&\big(x, u(x), u_x(x), u(\cdot)\big) \\ &= (n\lambda + \delta)u(x) - \lambda\big(\sqrt{n} + \theta\big)\mathbb{E}Y\, u_x(x) - n\lambda \int_0^{\sqrt{n}x} u\Big(x - \frac{t}{\sqrt{n}}\Big) dF_Y(t). \end{aligned}$$

Recall that F_n evaluated at the value function V_n is identically 0.

In the next proposition, we modify V_D by a constant of order $\mathcal{O}(n^{-1/2})$ to obtain a lower bound of V_n. In Appendix A, we present the background calculation that inspired this bound.

Proposition 3.2. Assume there exists $\varepsilon > 0$ such that

(3.11)
$$\mathbb{E}\big(e^{\varepsilon Y}\big) < \infty.$$

Then, there exists $q = q(\varepsilon) > 0$ and $N = N(\varepsilon) > 0$, such that, for all $n > N$ and all $x \geq 0$,

(3.12)
$$V_D(x) - \frac{q}{\sqrt{n}} \leq V_n(x).$$

Proof. First, note that

(3.13)
$$V_D(0) - \frac{q}{\sqrt{n}} = -\frac{q}{\sqrt{n}} < 0 \leq V_n(0).$$

Thus, by Proposition 4.2 in Azcue and Muler [2], to prove inequality (3.12), it suffices to show that there exists $q > 0$ such that $V_D - q/\sqrt{n}$ is a subsolution of $F_n = 0$.

G_n is linear with respect to u, u_x, and $u(\cdot)$; thus,

$$\begin{aligned} G_n&\big(x, V_D(x) - q/\sqrt{n}, V_D'(x), V_D(\cdot) - q/\sqrt{n}\,\big) \\ &= G_n\big(x, V_D(x), V_D'(x), V_D(\cdot)\big) - \frac{q}{\sqrt{n}}\, G_n(x, 1, 0, 1) \\ &= G_n\big(x, V_D(x), V_D'(x), V_D(\cdot)\big) - \frac{q}{\sqrt{n}}\big(\delta + n\lambda S_Y(d)\big), \end{aligned}$$

in which $d = \sqrt{n}x$ and $S_Y(d) := \mathbb{P}(Y > d)$. Note that there is $N = N(\varepsilon) > 0$ such that

(3.14)
$$\mathbb{E}\Big(Y^3\, e^{\frac{\gamma_2}{\sqrt{N}}Y}\Big) < \infty.$$

Now, from (A.5) in the appendix, for $n \geq N$, we can bound G_n evaluated at $V_D - q/\sqrt{n}$ on $[0, b_D]$ as follows:

$$(3.15) \quad G_n\big(x, V_D(x) - q/\sqrt{n}, V_D'(x), V_D(\cdot) - q/\sqrt{n}\big) \leq \frac{\lambda A}{\sqrt{n}} - \frac{q}{\sqrt{n}}\big(\delta + n\lambda S_Y(d)\big) < \frac{\lambda A - q\delta}{\sqrt{n}}$$

in which the positive constant A is defined in (A.4). Choose $q = q(N)$ so that

$$(3.16) \quad q \geq \frac{\lambda A}{\delta}.$$

(Because A in (A.4) depends on N, q's lower bound in (3.16) depends on N.) Then, $\lambda A - q\delta$ is nonpositive, and inequality (3.15) implies that G_n evaluated at $V_D - q/\sqrt{n}$ is negative on $[0, b_D]$.

Because $V_D'(x) = 1$ for all $x > b_D$, it follows that, for all $x \geq 0$ and for all $n > \max\big(N, q^2\big)$,

$$F_n\big(x, V_D(x) - q/\sqrt{n}, V_D'(x), V_D(\cdot) - q/\sqrt{n}\big) \leq 0 = F_n\big(x, V_n(x), V_n'(x), V_n(\cdot)\big).$$

We have shown that $V_D - q/\sqrt{n}$ is a subsolution of $F_n = 0$, and Proposition 4.2 in Azcue and Muler [2] implies the bound of V_n in (3.12). ∎

In the next proposition, we provide an upper bound of V_n, and we use Proposition 5.1 in Azcue and Muler [2] to prove the proposition.

Proposition 3.3. *Assume there exists $\varepsilon > 0$ such that*

$$(3.17) \quad \sup_{d \geq 0} \mathbb{E}\big(e^{\varepsilon(Y-d)} \,\big|\, Y > d\big) < \infty.$$

Then, there exists $p = p(\varepsilon) > 0$ and $N' = N'(\varepsilon) > 0$ such that, for all $n \geq N'$ and all $x \geq 0$,

$$(3.18) \quad V_n(x) \leq V_D(x) + \frac{p}{\sqrt{n}}.$$

Proof. By Proposition 5.1 of Azcue and Muler [2], because $V_D + p/\sqrt{n}$ is absolutely continuous and because $V_D(x) + p/\sqrt{n} \leq x + k$ for some $k > 0$ and for all $x \geq 0$, to prove the bound in (3.18), it is enough to show that $V_D + p/\sqrt{n}$ is a supersolution of $F_n = 0$.

First, notice that by the condition (3.17), there is $\tilde{N} > 0$, such that

$$(3.19) \quad \sup_{d \geq 0} \mathbb{E}\left((Y-d)^2 \, e^{\frac{\gamma_2}{\sqrt{\tilde{N}}}(Y-d)} \,\Big|\, Y > d\right) < \infty$$

and

$$(3.20) \quad \mathbb{E}\left(Y^3 \, e^{\frac{\gamma_2}{\sqrt{\tilde{N}}}Y}\right) < \infty.$$

Second, evaluate G_n at $V_D + p/\sqrt{n}$ on $[0, b_D]$. Let $d = \sqrt{n}x$ and $C = \big(\gamma_1 + \gamma_2\big)\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}}$;

then, via a calculation similar to the one in Appendix A, we have

$$(3.21) \quad G_n\Big(x, V_D(x) + \frac{p}{\sqrt{n}}, V_D'(x), V_D(\cdot) + \frac{p}{\sqrt{n}}\Big)$$

$$= G_n\big(x, V_D(x), V_D'(x), V_D(\cdot)\big) + \frac{p}{\sqrt{n}}\, G_n(x, 1, 0, 1)$$

$$= \frac{\lambda}{2C\sqrt{n}} \int_0^1 (1-\omega)^2 \left\{ \gamma_1^3 e^{\gamma_1 x} \mathbb{E}\Big(Y^3 e^{\frac{-\gamma_1 \omega}{\sqrt{n}} Y}\Big) + \gamma_2^3 e^{-\gamma_2 x} \mathbb{E}\Big(Y^3 e^{\frac{\gamma_2 \omega}{\sqrt{n}} Y}\Big) \right\} d\omega$$

$$+ \frac{\lambda S_Y(d)}{C} \left\{ - \sqrt{n}(\gamma_1 + \gamma_2)\, \mathbb{E}(Y - d | Y > d) \right.$$

$$+ \int_0^1 (1-\omega)\, \mathbb{E}\Big[\gamma_1^2 (Y-d)^2\, e^{-\frac{\gamma_1 \omega}{\sqrt{n}}(Y-d)} - \gamma_2^2 (Y-d)^2\, e^{\frac{\gamma_2 \omega}{\sqrt{n}}(Y-d)} \,\Big|\, Y > d\Big] d\omega \right\}$$

$$+ \frac{p}{\sqrt{n}}\big(\delta + n\lambda S_Y(d)\big)$$

$$> \frac{\lambda S_Y(d)}{C} \left\{ \sqrt{n}\big(pC - (\gamma_1 + \gamma_2)\mathbb{E}(Y - d | Y > d)\big) \right.$$

$$\left. - \int_0^1 (1-\omega)\, \mathbb{E}\Big[\gamma_2^2 (Y-d)^2\, e^{\frac{\gamma_2 \omega}{\sqrt{n}}(Y-d)} - \gamma_1^2 (Y-d)^2\, e^{-\frac{\gamma_1 \omega}{\sqrt{n}}(Y-d)} \,\Big|\, Y > d\Big] d\omega \right\}.$$

Choose p so that $pC > (\gamma_1 + \gamma_2)\sup_{d \geq 0} \mathbb{E}(Y - d | Y > d)$, this supremum is finite because of the bound in (3.19). Furthermore, the bound in (3.20) implies that there exists $N' \geq \tilde{N}$ such that, if $n \geq N'$, then the expression in (3.21) is nonnegative. Also, $V_D' \geq 1$ on $[0, b_D]$, so F_n evaluated at $V_D + p/\sqrt{n}$ is nonnegative on $[0, b_D]$.

Next, evaluate G_n at $V_D + p/\sqrt{n}$ on (b_D, ∞). Again, let $d = \sqrt{n}x$; then, after simplifying,

$$G_n\Big(x, V_D(x) + \frac{p}{\sqrt{n}}, V_D'(x), V_D(\cdot) + \frac{p}{\sqrt{n}}\Big)$$

$$= G_n\big(x, x - b_D, 1, \cdot - b_D\big) + \left(\frac{\theta \lambda \mathbb{E}Y}{\delta} + \frac{p}{\sqrt{n}}\right) G_n(x, 1, 0, 1)$$

$$= \delta\left\{(x - b_D) + \frac{p}{\sqrt{n}}\right\} + n\lambda S_Y(d) \left\{\frac{\theta \lambda \mathbb{E}Y}{\delta} - b_D + \frac{p - \mathbb{E}(Y - d | Y > d)}{\sqrt{n}}\right\} > 0$$

in which the inequality follows from (3.9) and from choosing $p > \sup_{d \geq 0} \mathbb{E}(Y - d | Y > d)$. Also, $V_D' = 1$ on (b_D, ∞); thus, F_n evaluated at $V_D + p/\sqrt{n}$ equals zero on (b_D, ∞).

We have shown that $V_D + p/\sqrt{n}$ is a supersolution of $F_n = 0$, and Proposition 5.1 in Azcue and Muler [2] implies the bound of V_n in (3.18). ■

In the following theorem, we show that V_n converges to V_D at a rate of order $\mathcal{O}(n^{-1/2})$.

Theorem 3.4. *If* (3.11) *and* (3.17) *hold, then there exists* $C' > 0$ *such that, for all* $n > \max(N, N')$ *and* $x \geq 0$,

$$(3.22) \qquad\qquad \big|V_n(x) - V_D(x)\big| \leq \frac{C'}{\sqrt{n}}.$$

Proof. From Propositions 3.2 and 3.3, it follows that

$$V_D(x) - \frac{q}{\sqrt{n}} \leq V_n(x) \leq V_D(x) + \frac{p}{\sqrt{n}}$$

with $q > 0$ and $p > 0$ given in the proofs of those propositions. Subtracting $V_D(x)$ from each side yields

$$-\frac{q}{\sqrt{n}} \leq V_n(x) - V_D(x) \leq \frac{p}{\sqrt{n}} \, .$$

Thus, if we set $C' = \max(q, p)$, inequality (3.22) follows. ∎

Remark 3.5. Our assumption that the moment generating function of Y is finite in a neighborhood of 0 stems from the fact that we estimate G_n for the function V_D, which includes an exponential term. Although our techniques cannot handle random variables with infinite moment generating functions (for example, Pareto and lognormal), observe that not only do we prove convergence, we also provide the rate of convergence. Specifically, Theorem 3.4 asserts that the rate of convergence of V_n to V_D is of order $\mathcal{O}(n^{-1/2})$ and, moreover, that the convergences is uniform over $x \in [0, \infty)$. By using probabilistic techniques and relying on convergence in distribution of the underlying processes, Theorem 3.6 in Bäuerle [3] proves the pointwise convergence $\lim_{n\to\infty} V_n(x) = V_D(x)$ without estimating the rate of convergence, but, as mentioned in the introduction, these techniques are only valid for the case of bounded rates of dividend payments. We leave determining the rate of convergence in more general setups for future research.

We end this section with an example in which we calculate C' in Theorem 3.4.

Example 3.6. Let $Y \sim Gamma(2, 1)$ with probability density function $f_Y(y) = ye^{-y}$ for $y \geq 0$, $\lambda = 10$, $\theta = 0.07$, and $\delta = 0.10$, which is the example Azcue and Muler [2] consider in section 10.1 of their paper. Azcue and Muler [2] give the following value function for $n = 1$; note that V_1 embodies a nonbarrier band strategy:

$$(3.23) \quad V_1(x) = \begin{cases} x + 2.119, & 0 \leq x < 1.803, \\ 0.0944e^{-1.48825x} - 9.431e^{-0.079355x} + 11.257e^{0.039567x}, & 1.803, \leq x < 10.22, \\ x + 2.456, & x \geq 10.22. \end{cases}$$

For this example, $\gamma_1 = 0.03894$, $\gamma_2 = 0.08561$, and $b_D = 12.650$. In (3.14), we may set $N = 1$, from which it follows that

$$A = \frac{1}{6C} \left\{ \gamma_1^3 e^{\gamma_1 b_D} \mathbb{E}(Y^3) + \gamma_2^3 \mathbb{E}\left(Y^3 e^{\gamma_2 Y}\right) \right\} = 0.04651$$

and gives $q = \lambda A / \delta = 100A = 4.651$. Also,

$$\sup_{d \geq 0} \mathbb{E}(Y - d | Y > d) = \sup_{d \geq 0} \frac{2 + d}{1 + d} = 2,$$

which implies that we can set

$$p > 2 \left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}} = 2.687.$$

It follows that $C' = \max(q, p) = 4.651$.

The numerical scheme for computing V_n for $n \in \mathbb{N}$ is prescribed on pages 95–96 in Schmidli [24]. For completeness, we describe it here. From Gerber [14], we know that the optimal policy is a band policy. Now, for the initial capital $x = 0$, either (a) dividends are paid, in which case $V_n(0) = \lambda(\sqrt{n} + \theta)/(n\lambda + \delta)$; or, (b) no dividends are paid immediately and there is a value $b_0 = \inf\{x > 0 : V_n'(x) = 1\} > 0$, and when surplus lies in the band $[0, b_0]$, no dividends are paid. The value function in this case would be derived by taking a solution u of $G_n\big(x, u(x), u_x(x), u(\cdot)\big) = 0$ on $[0, b_0]$ with the initial condition $u(0) = 1$ and by setting $V_n(x) = u(x)/u_x(b_0)$ for $x \in [0, b_0]$. Then, one repeats this process on $[b_0, \infty)$.

In our case, assume that dividends are paid in the band that includes $x = 0$; then, by differentiating the integro-differential equation $G_n = 0$ twice, one obtains the differential equation

$$0 = n\delta u(x) + 2\big(\sqrt{n}\delta - \lambda\theta n\big)u_x(x) + \big(-3\lambda n - 4\lambda\sqrt{n} + \delta\big)u_{xx}(x)$$
$$- 2\lambda(\sqrt{n} + \theta)u_{xxx}(x),$$

which implies the general form of the solution $u(x) = \sum_{i=1}^{3} a_i e^{\alpha_i x}$. By substituting this ansatz into $G_n = 0$ and by using the initial condition $u(0) = 1$, one obtains, for $n = 1$,

$$u(x) = 5.947983e^{0.039567x} - 5.058731e^{-0.079355x} + 0.110748e^{-1.48825x}.$$

However, the minimum of u_x is attained at $x = 0$. Hence, $u(x)/u(b_0) < 1$ for any $b_0 > 0$. Therefore, we deduce that dividends are paid on the band that includes $x = 0$. That is, $V_1(x) = x + \lambda(\sqrt{n} + \theta)/(n\lambda + \delta)$ in a neighborhood of $x = 0$. To find the first band's upper threshold b_0, we define for any $b > 0$ the function $V_1^b(x) : [b, \infty) \to \mathbb{R}$ such that V_1^b solves the integro-differential equation with the initial condition $V_1^b(b) = V_1(b) = b + \lambda(\sqrt{n} + \theta)/(n\lambda + \delta)$. Then, we set

$$b_0 = \underset{x \geq b}{\operatorname{argmin}} \, (V_1^b)_x(x) = 1,$$

and b_1 equals the value of x for which $(V_1^{b_0})_x(x) = 1$. The latter is the upper bound of the second band from the bottom. In our case, $b_0 = 1.80303$ and $b_1 = 10.2162$. Above $x = b_1$, it is always optimal to pay dividends. In conclusion, V_1 is given by (3.23).

We repeated this procedure for $n = 4, 9,$ and 25 and obtained the following value functions:

$$V_4(x) = \begin{cases} x + 1.66, & 0 \leq x < 0.63, \\ 0.04413e^{-2.9883x} - 10.07e^{-0.0082368x} + 10.8631e^{0.039254x}, & 0.63 \leq x < 10.8, \\ x + 1.799, & x \geq 10.8. \end{cases}$$

$$V_9(x) = \begin{cases} x + 0.94746, & 0 \leq x < 0.266, \\ 0.073e^{-4.4883x} - 10.257e^{-0.08342x} + 10.841e^{0.039415x}, & 0.266 \leq x < 13.243, \\ x + 1.66343, & x \geq 13.24343. \end{cases}$$

$$V_{25}(x) = \begin{cases} x + 0.51043, & 0 \leq x < 0.105, \\ 0.04413e^{-7.4883x} - 10.51e^{-0.00842816x} + 10.86e^{0.039x}, & 0.105 \leq x < 12.11, \\ x + 1.537, & x \geq 12.11. \end{cases}$$

The comparison between these three V_n's and the respective bounds $V_D + p/\sqrt{n}$ and $V_D - q/\sqrt{n}$ is illustrated in Figure 1.

$n = 4$

$n = 9$

$n = 25$
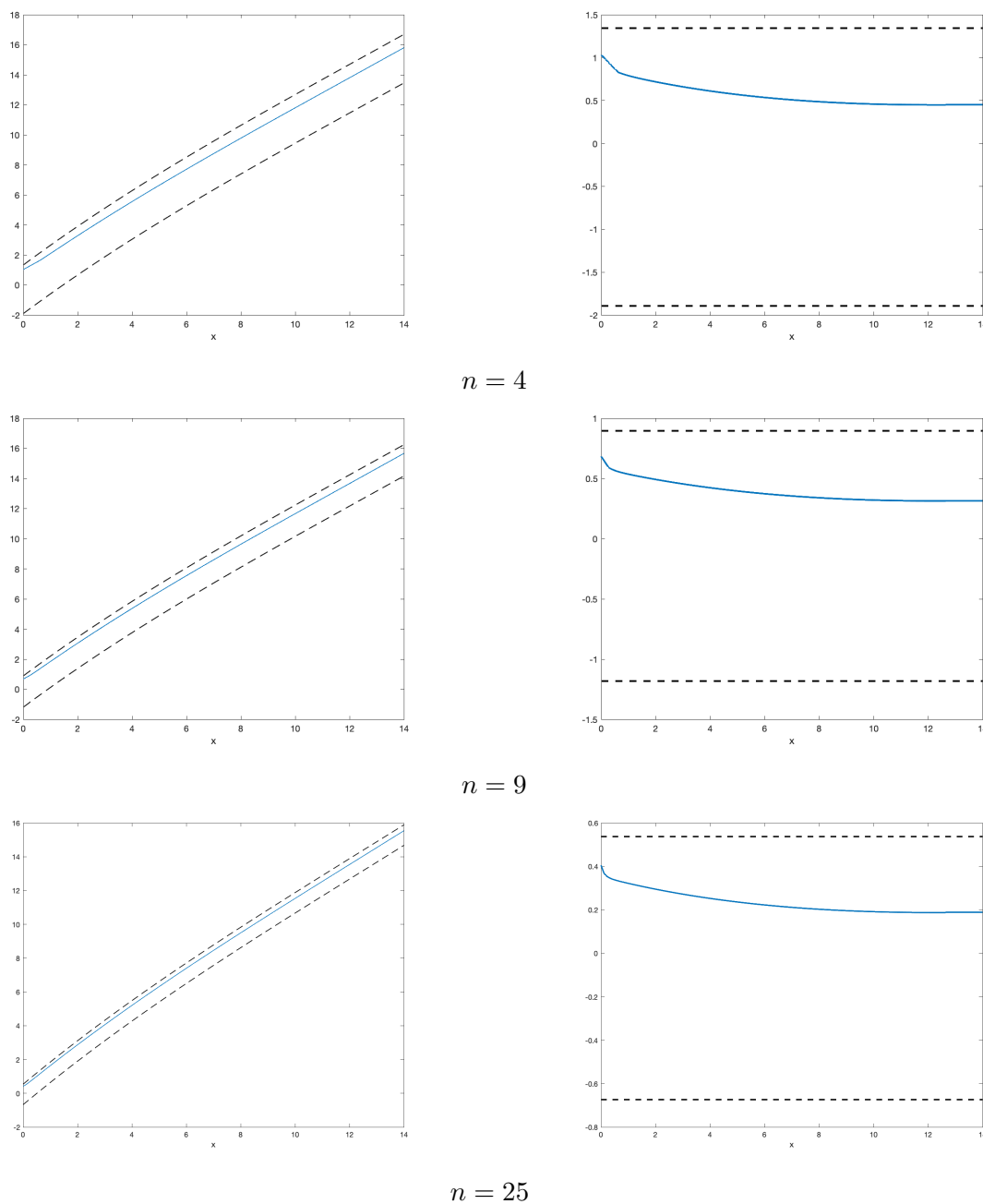
Figure 1. *For Example* 3.6 *the three graphs on the left-hand side are of V_n (solid line) and $V_D + \frac{p}{\sqrt{n}}$ and $V_D - \frac{q}{\sqrt{n}}$ (dashed). On the right-hand side the function $V_n - V_D$ is in solid and the dashed boundaries are $\frac{p}{\sqrt{n}}$ and $-\frac{q}{\sqrt{n}}$.*

3.3. $\mathcal{O}(n^{-1/2})$-optimality for using the barrier b_D for the scaled CL model. In this section, we show that using b_D as a barrier strategy for the scaled CL model, in place of the optimal band strategy, is $\mathcal{O}(n^{-1/2})$-optimal. Specifically, we show that there exists $C'' > 0$

and $N > 0$ such that $n > N$ implies

$$\left| V_n(x) - V_{D,n}(x) \right| < \frac{C''}{\sqrt{n}}$$

for all $x > 0$, in which $V_{D,n}$ denotes the (expected) payoff function for the scaled problem when we use the barrier b_D. From Lemma 2.48 of Schmidli [24], we know that there is a unique solution $g_n \in \mathcal{C}^1(\mathbb{R}+)$ of the integro-differential equation $G_n = 0$ with $g_n(0) = 1$. Moreover, the proof of this lemma shows that g_n is strictly increasing. We use g_n to construct an expression for $V_{D,n}$ as follows:

$$(3.24) \qquad V_{D,n}(x) = \begin{cases} \dfrac{g_n(x)}{g_n'(b_D)}, & 0 \le x \le b_D, \\ V_{D,n}(b_D) + (x - b_D), & x > b_D. \end{cases}$$

Note that $V_{D,n} \in \mathcal{C}^1(\mathbb{R}^+)$ with $V_{D,n}'(b_D) = 1$.

We connect $V_{D,n}$ and V_n via V_D because (1) $V_{D,n}$ and V_D are (expected) payoff functions for two different problems (scaled CL model versus its diffusion approximation) but with the *same* barrier b_D, and (2) we have a demonstrated relationship between V_D and V_n in (3.22) in Theorem 3.4. We begin by proving a theorem that is parallel to Theorem 3.4, but, first, we prove a comparison lemma for G_n on $[0, b_D]$, which we use to prove the parallel theorem.

Lemma 3.7. *Suppose $u, v \in \mathcal{C}^1([0, b_D])$ satisfying the following conditions:*
(i) *$u(0) \le v(0)$.*
(ii) *$G_n\big(x, u(x), u_x(x), u(\cdot)\big) \le G_n\big(x, v(x), v_x(x), v(\cdot)\big)$ for all $x \in (0, b_D]$.*
(iii) *$u_x(b_D) = v_x(b_D)$.*
Then, $u \le v$ on $[0, b_D]$.

Proof. Suppose, on the contrary, that $u(x) > v(x)$ for some value of $x \in (0, b_D]$. Then, there exists $x_0 \in (0, b_D]$ at which $u - v$ achieves a positive maximum, with $u_x(x_0) = v_x(x_0)$. Note that condition (iii) of the lemma ensures $u_x(x_0) = v_x(x_0)$ if x_0 equals the endpoint b_D. Then,

$$0 \le G_n\big(x_0, v(x_0), v_x(x_0), v(\cdot)\big) - G_n\big(x_0, u(x_0), u_x(x_0), u(\cdot)\big)$$

$$= (n\lambda + \delta)v(x_0) - \lambda\big(\sqrt{n} + \theta\big)\mathbb{E}Y\, v_x(x_0) - n\lambda \int_0^{\sqrt{n}x_0} v\Big(x_0 - \frac{t}{\sqrt{n}}\Big) dF_Y(t)$$

$$\quad - (n\lambda + \delta)u(x_0) + \lambda\big(\sqrt{n} + \theta\big)\mathbb{E}Y\, u_x(x_0) + n\lambda \int_0^{\sqrt{n}x_0} u\Big(x_0 - \frac{t}{\sqrt{n}}\Big) dF_Y(t)$$

$$= (n\lambda + \delta)\big(v(x_0) - u(x_0)\big) - n\lambda \int_0^{\sqrt{n}x_0} \left(v\Big(x_0 - \frac{t}{\sqrt{n}}\Big) - u\Big(x_0 - \frac{t}{\sqrt{n}}\Big) \right) dF_Y(t)$$

$$= -\big(n\lambda S_Y(\sqrt{n}x_0) + \delta\big)\big(u(x_0) - v(x_0)\big)$$

$$\quad - n\lambda \int_0^{\sqrt{n}x_0} \left(\big(u(x_0) - v(x_0)\big) - \Big(u\Big(x_0 - \frac{t}{\sqrt{n}}\Big) - v\Big(x_0 - \frac{t}{\sqrt{n}}\Big)\Big) \right) dF_Y(t)$$

$$< 0,$$

in which the inequality follows because $u - v$ achieves a positive maximum at x_0. Thus, we have a contradiction, so $u \leq v$ on $[0, b_D]$. ∎

In the following theorem, we use Lemma 3.7 to show that $V_{D,n}$ converges to V_D at a rate of order $\mathcal{O}(n^{-1/2})$. Theorem 3.7 in Bäuerle [3] proves the pointwise convergence $\lim_{n \to \infty} V_{D,n}(x) = V_D(x)$ without estimating the rate of convergence.

Theorem 3.8. *If* (3.11) *and* (3.17) *hold, then there exists* $C' > 0$ *such that, for all* $n > \max(N, N')$ *and* $x \geq 0$,

$$(3.25) \qquad \left| V_{D,n}(x) - V_D(x) \right| \leq \frac{C'}{\sqrt{n}}.$$

Proof. Note that $V_{D,n} \leq V_n$ on $\mathbb{R}^+$ because they are payoff functions for the same problem, and V_n is the maximum as the value function. Also, from Proposition 3.3, there exists $p > 0$ such that $V_n \leq V_D + p/\sqrt{n}$ on $\mathbb{R}^+$; thus, $V_{D,n} \leq V_D + p/\sqrt{n}$ on $\mathbb{R}^+$.

Next, compare $V_D - q/\sqrt{n}$ and $V_{D,n}$ on $[0, b_D]$ for $q > 0$ given in Proposition 3.2. In the proof of that proposition, we show that G_n evaluated at $V_D - q/\sqrt{n}$ is negative on $[0, b_D]$, and the construction in (3.24) shows that G_n evaluated at $V_{D,n}$ is zero on $[0, b_D]$. Because $u = V_D - q/\sqrt{n}$ and $v = V_{D,n}$ satisfy the conditions of Lemma 3.7, we deduce that $V_D - q/\sqrt{n} \leq V_{D,n}$ on $[0, b_D]$. Furthermore, because $V_{D,n}$ and $V_D - q/\sqrt{n}$ have slope identically equal to 1 for $x \geq b_D$, we deduce $V_D - q/\sqrt{n} \leq V_{D,n}$ on $\mathbb{R}^+$.

Thus, if we set $C' = \max(q, p)$, as in the proof of Theorem 3.4, inequality (3.25) follows. ∎

The following theorem is the main result of this section.

Theorem 3.9. *If* (3.11) *and* (3.17) *hold, then the barrier strategy with barrier* b_D *is* $\mathcal{O}(n^{-1/2})$-*optimal for the scaled problem. Specifically, then there exists* $C'' > 0$ *such that, for all* $n > \max(N, N')$ *and* $x \geq 0$,

$$(3.26) \qquad \left| V_n(x) - V_{D,n}(x) \right| \leq \frac{C''}{\sqrt{n}}.$$

Proof. If we set $C'' = 2C'$, then (3.26) follows from (3.22), (3.25), and the triangle inequality. ∎

Remark 3.10. Theorem 3.9 proves the $\mathcal{O}(n^{-1/2})$-optimality of using the optimal barrier for the diffusion approximation (namely, b_D) in place of the optimal band strategy for the n-scaled problem. This result supports the common practice in the mathematical finance and insurance literature of using the diffusion approximation in place of the classical risk model.

Although Theorem 3.9 proves that using the barrier strategy with barrier b_D is $\mathcal{O}(n^{-1/2})$-optimal for V_n's problem, we do not know that the barriers of V_n's band strategy converge to b_D; in fact, it does not appear to be true generally. Indeed, consider the gamma example on pages 95 and 96 of Schmidli [24]: $V'(x) = 1$ in a neighborhood of $x = 0$, so we hypothesize that, as $n \to \infty$, the smallest barrier goes to 0.

We end this paper with an example: let $Y \sim Exp(1)$, which implies Y_n is exponentially distributed with mean $1/\sqrt{n}$. As is well known, the optimal dividend strategy for V_n's problem is a barrier strategy; see, for example, Chapter 10 of Gerber [13] or Avanzi [1]. Moreover, we

have explicit expressions for the value function V_n and for its corresponding barrier b_n, and it is the latter in which we are interested. From equation (2.28) in Avanzi [1], the optimal barrier equals

$$(3.27) \qquad b_n = \frac{1}{r_1 + r_2} \ln\left(\frac{r_2^2(\sqrt{n} - r_2)}{r_1^2(\sqrt{n} + r_1)}\right)$$

in which

$$(3.28) \qquad r_1 = \frac{1}{2\lambda(\sqrt{n} + \theta)}\left[\sqrt{(\sqrt{n}\lambda\theta + \delta)^2 + 4n\delta\lambda} - (\sqrt{n}\lambda\theta - \delta)\right]$$

and

$$(3.29) \qquad r_2 = \frac{1}{2\lambda(\sqrt{n} + \theta)}\left[\sqrt{(\sqrt{n}\lambda\theta + \delta)^2 + 4n\delta\lambda} + (\sqrt{n}\lambda\theta - \delta)\right].$$

It is easy to see that

$$\lim_{n\to\infty} r_1 = \frac{1}{2}\left[\sqrt{\theta^2 + 4\delta/\lambda} - \theta\right] = \gamma_1$$

and

$$\lim_{n\to\infty} r_2 = \frac{1}{2}\left[\sqrt{\theta^2 + 4\delta/\lambda} + \theta\right] = \gamma_2$$

in which γ_1 and γ_2 are given in (3.4) and (3.5), respectively. Thus,

$$\lim_{n\to\infty} b_n = \frac{1}{\gamma_1 + \gamma_2} \ln\left(\frac{\gamma_2^2}{\gamma_1^2}\right) = \frac{2}{\gamma_1 + \gamma_2} \ln\left(\frac{\gamma_2}{\gamma_1}\right) = b_D,$$

as expected, and the rate of convergence is of order $\mathcal{O}(n^{-1/2})$. Indeed,

$$|b_n - b_D| \le \left|\frac{1}{r_1 + r_2} - \frac{1}{\gamma_1 + \gamma_2}\right| \times \left|\ln\left(\frac{r_2^2}{r_1^2}\right) + \ln\left(\frac{\sqrt{n} - r_2}{\sqrt{n} + r_1}\right)\right|$$
$$+ \frac{1}{\gamma_1 + \gamma_2} \times \left|\ln\left(\frac{r_2^2(\sqrt{n} - r_2)}{r_1^2(\sqrt{n} + r_1)}\right) - \ln\left(\frac{\gamma_2^2}{\gamma_1^2}\right)\right|.$$

Because $\ln(r_2^2/r_1^2)$ converges to $\ln(\gamma_2^2/\gamma_1^2)$ as $n \to \infty$, it follows that $|\ln(r_2^2/r_1^2)|$ is uniformly bounded over n. Also, $\ln(\frac{\sqrt{n}-r_2}{\sqrt{n}+r_1})$ is of order $\mathcal{O}(n^{-1/2})$. Hence, it suffices to estimate the following terms:

$$\left|\frac{1}{r_1 + r_2} - \frac{1}{\gamma_1 + \gamma_2}\right| \qquad \text{and} \qquad \left|\ln\left(\frac{r_2^2(\sqrt{n} - r_2)}{r_1^2(\sqrt{n} + r_1)}\right) - \ln\left(\frac{\gamma_2^2}{\gamma_1^2}\right)\right|.$$

Starting with the first term, there exists a constant $C > 0$ such that for any $n > 0$,

$$\left| \frac{1}{r_1 + r_2} - \frac{1}{\gamma_1 + \gamma_2} \right| = \lambda \left| \frac{\sqrt{n} + \theta}{\sqrt{n}\sqrt{(\lambda\theta + \delta/\sqrt{n})^2 + 4\lambda\delta}} - \frac{1}{\sqrt{(\lambda\theta)^2 + 4\lambda\delta}} \right|$$

$$\leq \lambda \left| \frac{1}{\sqrt{(\lambda\theta + \delta/\sqrt{n})^2 + 4\lambda\delta}} - \frac{1}{\sqrt{(\lambda\theta)^2 + 4\lambda\delta}} \right|$$

$$+ \frac{1}{\sqrt{n}} \cdot \frac{\lambda\theta}{\sqrt{(\lambda\theta + \delta/\sqrt{n})^2 + 4\lambda\delta}}$$

$$\leq \frac{C}{\sqrt{n}}.$$

Next,

$$\left| \ln\left(\frac{r_2^2(\sqrt{n} - r_2)}{r_1^2(\sqrt{n} + r_1)} \right) - \ln\left(\frac{\gamma_2^2}{\gamma_1^2} \right) \right| \leq \left| \ln\left(\frac{\sqrt{n} - r_2}{\sqrt{n} + r_1} \right) \right| + 2\left| \ln\left(\frac{r_1}{\gamma_1} \right) \right| + 2\left| \ln\left(\frac{r_2}{\gamma_2} \right) \right|.$$

The first term on the right side is of order $\mathcal{O}(n^{-1/2})$. The estimations for the last two terms are similar; hence, we provide details only for the last one:

$$\left| \ln\left(\frac{r_2}{\gamma_2} \right) \right| = \left| \ln\left(\frac{\sqrt{n}}{\sqrt{n} + \theta} \cdot \frac{\sqrt{(\lambda\theta + \delta/\sqrt{n})^2 + 4\delta\lambda} + \lambda\theta - \delta/\sqrt{n}}{\sqrt{(\lambda\theta)^2 + 4\delta\lambda} + \lambda\theta} \right) \right|$$

$$\leq \ln\left(1 + \frac{\theta}{\sqrt{n}} \right) + \left| \ln\left(\frac{\sqrt{(\lambda\theta + \delta/\sqrt{n})^2 + 4\delta\lambda} + \lambda\theta - \delta/\sqrt{n}}{\sqrt{(\lambda\theta)^2 + 4\delta\lambda} + \lambda\theta} \right) \right|.$$

The right side is bounded from above by $C\ln(1 + C/\sqrt{n})$ for some positive constant C, independent of n, and this bound is of order $\mathcal{O}(n^{-1/2})$.

Appendix A. G_n evaluated at V_D on $[0, b_D]$. In this appendix, we present the calculations that inspired Proposition 3.2. Recall that

$$F_n\big(x, u(x), u_x(x), u(\cdot)\big) = \min\Big\{ G_n\big(x, u(x), u_x(x), u(\cdot)\big), \, u_x(x) - 1 \Big\}$$

in which the operator G_n is defined in (3.10). We now evaluate G_n at V_D for $0 \leq x \leq b_D$:

$$
\begin{aligned}
(A.1) \quad & G_n\big(x, V_D(x), V_D'(x), V_D(\cdot)\big) \\
&= \lambda\bigg\{ \int_0^\infty \bigg[n\left(V_D(x) - V_D\left(x - \frac{t}{\sqrt{n}}\right) \right) \\
&\qquad\qquad - \sqrt{n}\,\mathbb{E}Y\, V_D'(x) + \left(\frac{\delta}{\lambda} V_D(x) - \theta\mathbb{E}Y V_D'(x) \right) \bigg] dF_Y(t) \\
&\qquad + \int_{\sqrt{n}x}^\infty n V_D\left(x - \frac{t}{\sqrt{n}}\right) dF_Y(t) \bigg\}
\end{aligned}
$$

in which we extend $V_D(x)$ to $x < 0$ via the first expression in (3.3). Note that $V_D(x) < 0$ for $x < 0$, which implies that the second integral above is nonpositive. In the first integral, we write $V_D(x) = \left(e^{\gamma_1 x} - e^{-\gamma_2 x}\right)/C$ in which C equals

$$C = (\gamma_1 + \gamma_2)\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}}.$$

The first integral is linear in V_D and, hence, equals the difference of two integrals: one with $e^{\gamma_1 x}/C$ replacing $V_D(x)$, and the second with $e^{-\gamma_2 x}/C$ replacing $V_D(x)$. We obtain the second of these integrals from the first one by substituting $-\gamma_2$ for γ_1; thus, we show the details only for γ_1:

$$\frac{n}{C}\int_0^\infty \left[e^{\gamma_1 x} - e^{\gamma_1\left(x - \frac{t}{\sqrt{n}}\right)} - \frac{1}{\sqrt{n}}\,\mathbb{E}Y\gamma_1 e^{\gamma_1 x} + \frac{1}{n}\left(\frac{\delta}{\lambda}\,e^{\gamma_1 x} - \theta\mathbb{E}Y\gamma_1 e^{\gamma_1 x}\right)\right]dF_Y(t)$$

$$= -\frac{n}{C}\,e^{\gamma_1 x}\int_0^\infty\left[e^{-\gamma_1 t/\sqrt{n}} - \left(1 - \frac{\gamma_1 t}{\sqrt{n}} + \frac{\gamma_1^2 t^2}{2n}\right)\right]dF_Y(t)$$

$$= \frac{\gamma_1^3}{2C\sqrt{n}}\,e^{\gamma_1 x}\int_0^1 (1-\omega)^2\,\mathbb{E}\left[Y^3 e^{\frac{-\gamma_1\omega}{\sqrt{n}}Y}\right]d\omega$$

in which the first and second equalities, respectively, follow from the identities

$$\frac{\delta}{\lambda} - \theta\mathbb{E}Y\gamma_1 = \frac{\gamma_1^2\,\mathbb{E}(Y^2)}{2}$$

and

(A.2) $$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{n!}\int_0^1 (1-\omega)^n e^{\omega x}d\omega.$$

A similar analysis for $e^{-\gamma_2 x}/C$ yields that the first integral on the right side of (A.1) equals

(A.3) $$\frac{\gamma_1^3}{2C\sqrt{n}}\,e^{\gamma_1 x}\int_0^1 (1-\omega)^2\,\mathbb{E}\left[Y^3 e^{\frac{-\gamma_1\omega}{\sqrt{n}}Y}\right]d\omega + \frac{\gamma_2^3}{2C\sqrt{n}}\,e^{-\gamma_2 x}\int_0^1 (1-\omega)^2\,\mathbb{E}\left[Y^3 e^{\frac{\gamma_2\omega}{\sqrt{n}}Y}\right]d\omega.$$

Because of the bound in (3.11), we can rewrite and bound the first integral in (A.1) on $[0, b_D]$ as follows: for $n \geq N$, in which N is such that inequality (3.14) holds,

$$\frac{1}{2C\sqrt{n}}\int_0^1 (1-\omega)^2\left\{\gamma_1^3 e^{\gamma_1 x}\mathbb{E}\left(Y^3 e^{\frac{-\gamma_1\omega}{\sqrt{n}}Y}\right) + \gamma_2^3 e^{-\gamma_2 x}\mathbb{E}\left(Y^3 e^{\frac{\gamma_2\omega}{\sqrt{n}}Y}\right)\right\}d\omega$$

$$\leq \frac{1}{2C\sqrt{n}}\int_0^1 (1-\omega)^2\left\{\gamma_1^3 e^{\gamma_1 b_D}\mathbb{E}(Y^3) + \gamma_2^3\,\mathbb{E}\left(Y^3 e^{\frac{\gamma_2}{\sqrt{n}}Y}\right)\right\}d\omega$$

(A.4) $$= \frac{1}{6C\sqrt{n}}\left\{\gamma_1^3 e^{\gamma_1 b_D}\mathbb{E}(Y^3) + \gamma_2^3\,\mathbb{E}\left(Y^3 e^{\frac{\gamma_2}{\sqrt{N}}Y}\right)\right\} =: \frac{A}{\sqrt{n}}.$$

Thus, for $n \geq N$ and $x \in [0, b_D]$, we have

(A.5) $$G_n\left(x, V_D(x), V_D'(x), V_D(\cdot)\right) \leq \frac{\lambda A}{\sqrt{n}}.$$

REFERENCES

[1] B. AVANZI, *Strategies for dividend distribution: a review*, N. Am. Actuar. J., 13 (2009), pp. 217–251, https://doi.org/10.1080/10920277.2009.10597549.

[2] P. AZCUE AND N. MULER, *Optimal reinsurance and dividend distribution policies in the Cramér-Lundberg model*, Math. Finance, 15 (2005), pp. 261–308, https://doi.org/10.1111/j.0960-1627.2005.00220.x.

[3] N. BÄUERLE, *Approximation of optimal reinsurance and dividend payout policies*, Math. Finance, 14 (2004), pp. 99–113, https://doi.org/10.1111/j.0960-1627.2004.00183.x.

[4] A. BUDHIRAJA AND A. P. GHOSH, *Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function*, Ann. Appl. Probab., 16 (2006), pp. 1962–2006, https://doi.org/10.1214/105051606000000457.

[5] A. BUDHIRAJA AND K. ROSS, *Existence of optimal controls for singular control problems with state constraints*, Ann. Appl. Probab., 16 (2006), pp. 2235–2255, https://doi.org/10.1214/105051606000000556.

[6] A. BUDHIRAJA AND K. ROSS, *Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem*, SIAM J. Control Optim., 45 (2007), pp. 2169–2206, https://doi.org/10.1137/050640515.

[7] A. COHEN, *Asymptotic analysis of a multiclass queueing control problem under heavy traffic with model uncertainty*, Stoch. Syst., 9 (2019), pp. 359–391, https://doi.org/10.1287/stsy.2019.0034.

[8] A. COHEN, *Brownian control problems for a multiclass M/M/1 queueing problem with model uncertainty*, Math. Oper. Res., 44 (2019), pp. 739–766.

[9] A. COHEN, *On Singular control problems, the time-stretching method, and the weak-M1 topology*, SIAM J. Control Optim., to appear.

[10] A. COHEN AND V. R. YOUNG, *Rate of convergence of the probability of ruin in the Cramér-Lundberg model to its diffusion approximation*, Insurance Math. Econom., 93 (2020), pp. 333–340, https://doi.org/10.1016/j.insmatheco.2020.06.003.

[11] C. COSTANTINI AND T. G. KURTZ, *Existence and uniqueness of reflecting diffusions in cusps*, Electron. J. Probab., 23 (2018), 84, https://doi.org/10.1214/18-EJP204.

[12] C. COSTANTINI AND T. G. KURTZ, *Markov selection for constrained martingale problems*, Electron. J. Probab., 24 (2019), 135, https://doi.org/10.1214/19-ejp393.

[13] H. U. GERBER, *An Introduction to Mathematical Risk Theory*, S.S. Heubner Foundation Monograph Series 8, S.S. Huebner Foundation for Insurance Education, Homewood, IL, 1979.

[14] H. U. GERBER, *Entscheidungskriterien für den zusammengesetzten Poisson-Prozess*, PhD thesis, ETH Zurich, Zurich, 1969.

[15] H. U. GERBER AND E. S. W. SHIU, *Optimal dividends: Analysis with Brownian motion*, N. Am. Actuar. J., 8 (2004), pp. 1–20, https://doi.org/10.1080/10920277.2004.10596125.

[16] H. U. GERBER, E. S. W. SHIU, AND N. SMITH, *Methods for estimating the optimal dividend barrier and the probability of ruin*, Insurance Math. Econom., 42 (2008), pp. 243–254, https://doi.org/10.1016/j.insmatheco.2007.02.002.

[17] D. L. IGLEHART, *Diffusion approximations in collective risk theory*, J. Appl. Probab., 6 (1969), pp. 285–292, https://doi.org/10.2307/3211999.

[18] T. G. KURTZ, *Martingale problems for constrained Markov problems*, in Recent Advances in Stochastic Calculus (College Park, MD, 1987), Progr. Automat. Info. Systems, Springer, New York, 1990, pp. 151–168.

[19] T. G. KURTZ, *A control formulation for constrained Markov processes*, in Mathematics of Random Media (Blacksburg, VA, 1989), Lectures in Appl. Math. 27, American Mathematical Society, Providence, RI, 1991, pp. 139–150.

[20] T. G. KURTZ, *Random time changes and convergence in distribution under the Meyer-Zheng conditions*, Ann. Probab., 19 (1991), pp. 1010–1034.

[21] H. J. KUSHNER AND L. F. MARTINS, *Numerical methods for stochastic singular control problems*, SIAM J. Control Optim., 29 (1991), pp. 1443–1475, https://doi.org/10.1137/0329073.

[22] L. F. MARTINS AND H. J. KUSHNER, *Routing and singular control for queueing networks in heavy traffic*, SIAM J. Control Optim., 28 (1990), pp. 1209–1233, https://doi.org/10.1137/0328065.

[23] P.-A. MEYER AND W. A. ZHENG, *Tightness criteria for laws of semimartingales*, Ann. Inst. Poincaré Probab. Stat., 20 (1984), pp. 353–372, http://www.numdam.org/item?id=AIHPB_1984__20_4_353_0.

[24] H. SCHMIDLI, *Stochastic control in insurance*, Probability and its Applications, Springer, London, 2008.

[25] H. SCHMIDLI, *Risk theory*, Springer Actuar., Springer, Cham, Switzerland, 2017, https://doi.org/10.1007/978-3-319-72005-0.