



Optimal reduction of public debt under partial observation of the economic growth

Giorgia Callegaro¹ · Claudia Ceci² ·
Giorgio Ferrari³

Received: 30 January 2019 / Accepted: 10 June 2020 / Published online: 15 September 2020
© The Author(s) 2020

Abstract We consider a government that aims at reducing the debt-to-(gross domestic product) (GDP) ratio of a country. The government observes the level of the debt-to-GDP ratio and an indicator of the state of the economy, but does not directly observe the development of the underlying macroeconomic conditions. The government's criterion is to minimise the sum of the total expected costs of holding debt and of debt reduction policies. We model this problem as a singular stochastic control problem under partial observation. The contribution of the paper is twofold. Firstly, we provide a general formulation of the model in which the level of the debt-to-GDP ratio and the value of the macroeconomic indicator evolve as a diffusion and a jump-diffusion, respectively, with coefficients depending on the regimes of the economy. The latter are described through a finite-state continuous-time Markov chain. We reduce the original problem via filtering techniques to an equivalent one with full information (the so-called separated problem), and we provide a general verification result in terms of a related optimal stopping problem under full information. Secondly, we specialise to a case study in which the economy faces only two regimes and the macroeconomic indicator has a suitable diffusive dynamics. In this setting, we provide the optimal debt reduction policy. This is given in terms of the contin-

✉ G. Ferrari
giorgio.ferrari@uni-bielefeld.de

G. Callegaro
gcallega@math.unipd.it

C. Ceci
claudia.ceci@unich.it

- ¹ Department of Mathematics, University of Padova, via Trieste 63, 35121 Padova, Italy
- ² Department of Economics, University "G. D'Annunzio" of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy
- ³ Center for Mathematical Economics (IMW), Bielefeld University, Universitätsstrasse 25, 33615 Bielefeld, Germany

uous free boundary arising in an auxiliary fully two-dimensional optimal stopping problem.

Keywords Singular stochastic control · Partial observation · Optimal stopping · Free boundary · Debt-to-GDP ratio

Mathematics Subject Classification (2010) 93E20 · 60G35 · 93E11 · 60G40 · 60J60 · 91B64

JEL Classification C61 · H63 · E32

1 Introduction

The question of optimally managing the debt-to-GDP ratio (also called “debt ratio”) of a country has become particularly important in the last years. Indeed, concurrently with the financial crisis started in 2007, debt-to-GDP ratios exploded from an average of 53% to circa 80% in developed countries. Clearly, the debt management policy of a government highly depends on the underlying macroeconomic conditions; indeed, these affect for example the growth rate of the GDP which in turn determines the growth rate of the debt-to-GDP ratio of a country. However, in practice, it is typically neither possible to measure in real time the growth rate of the GDP, nor can one directly observe the underlying business cycles. On August 24, 2018, during a speech at “Changing Market Structure and Implications for Monetary Policy” – a symposium sponsored by the Federal Reserve Bank of Kansas City in Jackson Hole, Wyoming –, the chairman of the Federal Reserve Jerome H. Powell said:

*... In conventional models of the economy, major economic quantities such as inflation, unemployment and the growth rate of the gross domestic product fluctuate around values that are considered “normal” or “natural” or “desired”. The FOMC (Federal Open Market Committee) has chosen a 2 percent inflation objective as one of these desired values. **The other values are not directly observed, nor can they be chosen by anyone...***

Following an idea that dates back to Hamilton [38], we assume in this paper that the GDP growth rate of a country is modulated by a continuous-time Markov chain that is not directly observable. The Markov chain has $Q \geq 2$ states modelling the different business cycles of the economy, so that a shift in the macroeconomic conditions induces a change in the value of the growth rate of the GDP. The government can observe only the current levels of the debt-to-GDP ratio and of a macroeconomic indicator. The latter might be e.g. one of the so-called “Big Four” which are usually considered proxies of the industrial production index, hence of the business conditions. These indicators constitute the Conference Board’s Index of Coincident Indicators; they are employment in non-agricultural businesses, industrial production, real personal income less transfers, and real manufacturing and trade sales. We refer to e.g. Stock and Watson [60], where the authors present a wide range of economic indicators and examine the forecasting performance of various of them in the recession of 2001.

Motivated by the recent aforementioned debt crisis, we consider a government that has the priority to *return debt to less dangerous levels, to move away from the dark corners* (O. Blanchard, former chief economist of the International Monetary Fund (2014)) e.g. through fiscal policies or imposing austerity policies in the form of spending cuts. In our model, we thus preclude the possibility for the government to increase the level of the debt ratio and we neglect any possible potential benefit resulting from holding debt, even if we acknowledge that a policy of debt reduction might not be always the most sensible approach, as also observed by Ostry et al. [55] (see also the discussion in Remark 2.6 below). We further assume that the loss resulting from holding debt can be measured through a convex and nondecreasing cost function, and that the debt ratio is instantaneously affected by any reduction. The latter need not necessarily be performed at rates, but also lump sum actions are allowed, and the cumulative amount of the debt ratio's decrease is the government's control variable. Any decrease of the debt ratio results in proportional costs, and the government aims at choosing a debt-reduction policy that minimises the total expected loss of holding debt, plus the total expected costs of interventions on the debt ratio. In line with recent papers on stochastic control methods for optimal debt management (see Cadenillas and Huamán-Aguilar [8, 9], Ferrari [31] and Ferrari and Rodosthenous [32]), we model the previous problem as a singular stochastic control problem. However, differently to all previous works, our problem is formulated in a partial observation setting, thus leading to a completely different mathematical analysis. In our model, the observations consist of the debt ratio and the macroeconomic indicator. The debt ratio is a linearly controlled geometric Brownian motion, and its drift is given in terms of the GDP growth rate, which is modulated by the unobservable continuous-time Markov chain Z . The macroeconomic indicator is a real-valued jump-diffusion which is correlated to the debt ratio process, and which has drift and both intensity and jump sizes depending on Z .

Our contributions. Our study of the optimal debt reduction problem is performed through three main steps.

First of all, via advanced filtering techniques with mixed-type observations, we reduce the original problem to an equivalent problem under full information, the so-called separated problem. This is a classical procedure used to handle optimal stochastic control problems under partial information (see e.g. Fleming and Pardoux [33], Bensoussan [4, Chap. 7.1] and Ceci and Gerardi [12]) The filtering problem consists in characterising the conditional distribution of the unobservable Markov chain Z at any time t , given observations up to time t . The case of diffusion observations has been widely studied in the literature, and textbook treatments can be found in Elliott et al. [29, Chap. 8], Kallianpur [44, Chap. 8] and Liptser and Shiryaev [49, Chap. 8]. There are also known results for pure-jump observations (see e.g. Brémaud [7, Chap. IV], Ceci and Gerardi [13, 14], Kliemann et al. [47] and references therein). More recently, filtering problems with mixed-type information which involve pure-jump processes and diffusions have been studied by Ceci and Colaneri [15, 16], among others.

Notice that also the economic and financial literature has experienced papers on models under partial observation where a reduction to a complete information setting is performed via filtering techniques and the problem is split into the so-called “two-

step procedure". We refer e.g. to the literature on portfolio selection in the seminal papers by Detemple [25] and Gennotte [36] (in a continuous-time setting, with diffusive observations leading to a Gaussian filter process); to Veronesi [62] for an equilibrium model with uncertain dividend drift in the field of market over- and under-reaction to information; to the more recent work by Luo [51], where different uncertainty models are analysed in a Gaussian setting with the aim of studying strategic consumption–portfolio rules when dealing with precautionary savings. Generally, the economic and financial literature refers to well-known results in filtering theory, in the case of diffusive observation processes such as additive Gaussian white noise (e.g. Detemple [25], Gennotte [36] and Luo [51]) or in the case of pure-jump observations (see Bäuerle and Rieder [3] and Ceci [11], among others). A few papers consider mixed-type information (see e.g. Callegaro et al. [10] and Frey and Schmidt [35], among others). In our paper, we deal with a more general setting with a two-dimensional observation process allowing jumps, for which known results cannot be invoked.

Usually, two filtering approaches are followed: the so-called reference probability approach (see the seminal paper by Zakai [65] and the more recent papers Frey and Runggaldier [34], Ceci and Colaneri [16] and Colaneri et al. [18], among others) and the innovation approach (see e.g. Brémaud [7, Chap. IV.1], Ceci and Colaneri [15], Eksi and Ku [27] and Frey and Schmidt [35]). Due to the general structure of our observations' dynamics, the innovation approach is more suitable to handle our filtering problem, and this leads to the so-called Kushner–Stratonovich equation. In particular, it turns out that the dynamics of our filter and of the observation process are coupled, thus making the proof of uniqueness of the solution to the Kushner–Stratonovich system more delicate. After providing such a result, we are then able to show that the original problem under partial observation and the separated problem are equivalent, that is, they share the same value and the same optimal control.

Secondly, we exploit the convex structure of the separated problem and provide a general probabilistic verification theorem. This result – which is in line with findings in Baldursson and Karatzas [2], De Angelis et al. [20] and Ferrari [31], among others – relates the optimal control process to the solution to an auxiliary optimal stopping problem. Moreover, it proves that the value function of the separated problem is the integral – with respect to the controlled state variable – of the value function of the optimal stopping problem. The stopping problem thus gives the optimal timing at which debt should be marginally reduced.

Finally, by specifying a setting in which the continuous-time Markov chain faces only two regimes (a fast growth or slow growth phase) and the macroeconomic indicator is a suitable diffusion process, we are able to characterise the optimal debt reduction policy. In this framework, the filter process is a two-dimensional process $(\pi_t, 1 - \pi_t)_{t \geq 0}$, where π_t is the conditional probability at time t that the economy enjoys the fast growth phase. We prove that the optimal control prescribes to keep at any time the debt ratio below an endogenously determined curve that is a function of the government's belief about the current state of the economy. Such a debt ceiling is the free boundary of the fully two-dimensional optimal stopping problem that is related to the separated problem (in the sense of the previously discussed verification theorem). By using almost exclusively probabilistic arguments, we are able to show that the value function of the auxiliary optimal stopping problem is a C^1 -function of

its arguments, and thus enjoys the so-called smooth-fit property. Moreover, the free boundary is a continuous, bounded and increasing function of the filter process. This last monotonicity property has also a clear economic interpretation: the more the government believes that the economy enjoys a regime of fast growth, the less strict the optimal debt reduction policy should be.

As a remarkable byproduct of the regularity of the value function of the optimal stopping problem, we also obtain that the value function of the singular stochastic control problem is a classical solution to its associated Hamilton–Jacobi–Bellman (HJB) equation. The latter takes the form of a variational inequality involving an elliptic second-order partial differential equation (PDE). It is worth noticing that the C^2 -regularity of the value function implies the validity of a second-order principle of smooth fit, usually observed in one-dimensional problems.

We believe that the study of the auxiliary fully two-dimensional optimal stopping problem is a valuable contribution to the literature on its own. Indeed, while the literature on one-dimensional optimal stopping problems is very rich, the problem of characterising the optimal stopping rule in multi-dimensional settings has been so far rarely explored in the literature (see the recent work by Christensen et al. [17], De Angelis et al. [20] as well as Johnson and Peskir [43] among the very few papers dealing with multi-dimensional stopping problems). This discrepancy is due to the fact that a standard guess-and-verify approach, based on the construction of an explicit solution to the variational inequality arising in the considered optimal stopping problem, is no longer applicable in multi-dimensional settings where the variational inequality involves a PDE rather than an ordinary differential equation.

Related literature. As already noticed above, our paper is placed among those recent works addressing the problem of optimal debt management via continuous-time stochastic control techniques. In particular, Cadenillas and Huamán-Aguilar [8, 9] model an optimal debt reduction problem as a one-dimensional control problem with singular and bounded-velocity controls, respectively. In the work by Ferrari and Rosthenous [32], the government is allowed to increase and decrease the current level of the debt ratio, and the interest rate on debt is modulated by a continuous-time observable Markov chain. The mathematical formulation leads to a one-dimensional bounded-variation stochastic control problem with regime switching. In the model by Ferrari [31], when optimally reducing the debt ratio, the government takes into consideration the evolution of the inflation rate of the country. The latter evolves as an uncontrolled diffusion process and affects the growth rate of the debt ratio, which is a process of bounded variation. In this setting, the debt reduction problem is formulated as a two-dimensional singular stochastic control problem whose HJB equation involves a second-order linear parabolic partial differential equation. All the previous papers are formulated in a full information setting, while ours is under partial observation.

The literature on singular stochastic control problems under partial observation is also still quite limited. Theoretical results on the PDE characterisation of the value function of a two-dimensional optimal correction problem under partial observation are obtained by Menaldi and Robin [53], whereas a general maximum principle for a not necessarily Markovian singular stochastic control problem under partial information has more recently been derived by Øksendal and Sulem [54]. We also refer

to De Angelis [19] and Decámps and Villeneuve [23] who provide a thorough study of the optimal dividend strategy in models in which the surplus process evolves as a drifted Brownian motion with unknown drift that can take only two constant values, with given probabilities.

Outline of the paper. The rest of the paper is organised as follows. In Sect. 2, we introduce the setting and formulate the problem. The reduction of the problem under partial observation to the separated problem is performed in Sect. 3; in particular, the filtering results are presented in Sect. 3.1. The probabilistic verification theorem connecting the separated problem to one of optimal stopping is then proved in Sect. 3.3. In Sect. 4, we consider a case study in which the economy faces only two regimes. Its solution, presented in Sects. 4.2 and 4.3, hinges on the study of a two-dimensional optimal stopping problem that is performed in Sect. 4.1. Finally, Appendix A collects the proofs of some technical filtering results.

2 Setting and problem formulation

2.1 The setting

Consider the complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ capturing all the uncertainty of our setting. Here, $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ denotes the *full information filtration*. We suppose that it satisfies the usual hypotheses of completeness and right-continuity.

We denote by Z a continuous-time finite-state Markov chain describing the different states of the economy. For $Q \geq 2$, let $S := \{1, 2, \dots, Q\}$ be the state space of Z and $(\lambda_{ij})_{1 \leq i, j \leq Q}$ its generator matrix. Here λ_{ij} , $i \neq j$, gives the intensity of a transition from state i to state j and is such that $\lambda_{ij} \geq 0$ for $i \neq j$ and $\sum_{j=1, j \neq i}^Q \lambda_{ij} = -\lambda_{ii}$. For any time $t \geq 0$, Z_t is \mathcal{F}_t -measurable.

In the absence of any intervention by the government, we assume that the (uncontrolled) debt-to-GDP ratio evolves as

$$dX_t^0 = (r - g(Z_t))X_t^0 dt + \sigma X_t^0 dW_t, \quad X_0^0 = x \in (0, \infty), \quad (2.1)$$

where W is a standard \mathbb{F} -Brownian motion on (Ω, \mathcal{F}) independent of Z , $r \geq 0$ and $\sigma > 0$ are constants and $g : S \rightarrow \mathbb{R}$. The constant r is the real interest rate on debt, σ is the debt's volatility and $g(i) \in \mathbb{R}$ is the rate of the GDP growth when the economy is in state $i \in S$.

It is clear that (2.1) admits a unique strong solution, and when needed, we denote it by $X^{x,0}$ for any $x > 0$. The current level of the debt-to-GDP ratio is known to the government at any time t , and $X^{x,0}$ is therefore the first component of the so-called *observation process*.

The government also observes a macroeconomic stochastic indicator η , e.g. one of the so-called “Big Four”, which we interpret as a proxy of the business conditions. We assume that η is a jump-diffusion process solving the stochastic differential equation

$$\begin{aligned} d\eta_t &= b_1(\eta_t, Z_t)dt + \sigma_1(\eta_t)dW_t + \sigma_2(\eta_t)dB_t + c(\eta_{t-}, Z_{t-})dN_t, \\ \eta_0 &= q \in \mathcal{I}, \end{aligned} \quad (2.2)$$

where $\sigma_1 > 0$, $\sigma_2 > 0$ and b_1, c are measurable functions of their arguments and $\mathcal{I} \subseteq \mathbb{R}$ is the state space of η . Here, B is an \mathbb{F} -standard Brownian motion independent of W and Z . Moreover, N is an \mathbb{F} -adapted point process, without common jump times with Z , independent of W and B . The predictable intensity of N is denoted by $(\lambda^N(Z_{t-}))_{t \geq 0}$ and depends on the current state of the economy, with $\lambda^N(\cdot) > 0$ being a measurable function. **From now on, we make the following assumptions** that ensure strong existence and uniqueness of the solution to (2.2) (by a standard localising argument, one can indeed argue as e.g. in the proof of Xi and Zhu [64, Theorem 2.1], performed in a setting more general than ours by employing Ikeda and Watanabe [40, Theorem IV.9.1]).

Assumption 2.1 The functions $b_1 : \mathcal{I} \times S \rightarrow \mathbb{R}$, $\sigma_1 : \mathcal{I} \rightarrow (0, \infty)$, $\sigma_2 : \mathcal{I} \rightarrow (0, \infty)$ and $c : \mathcal{I} \times S \rightarrow \mathbb{R}$ are such that for any $i \in S$,

- (i) (continuity) $b_1(\cdot, i)$, $\sigma_1(\cdot)$, $\sigma_2(\cdot)$ and $c(\cdot, i)$ are continuous;
- (ii) (local Lipschitz conditions) for any $R > 0$, there exists a constant $L_R > 0$ such that if $|q| < R$, $|q'| < R$, $q, q' \in \mathcal{I}$, then

$$|b_1(q, i) - b_1(q', i)| + |\sigma_1(q) - \sigma_1(q')| + |\sigma_2(q) - \sigma_2(q')| + |c(q, i) - c(q', i)| \leq L_R |q - q'|;$$

- (iii) (growth conditions) there exists a constant $C > 0$ such that

$$|b_1(q, i)|^2 + |\sigma_1(q)|^2 + |\sigma_2(q)|^2 + |c(q, i)|^2 \leq C(1 + |q|^2).$$

The dynamics proposed in (2.2) is of jump-diffusive type and allows size and intensity of the jumps to be affected by the state of the economy. It is therefore flexible enough to describe a large class of stochastic factors which may exhibit jumps.

The *observation filtration* $\mathbb{H} = (\mathcal{H}_t)_{t \geq 0}$ is defined as

$$\mathbb{H} := \mathbb{F}^{X^0} \vee \mathbb{F}^\eta,$$

where \mathbb{F}^{X^0} and \mathbb{F}^η denote the natural filtrations generated by X^0 and η , respectively, as usual augmented by \mathbb{P} -null sets. Clearly, (X^0, η) is adapted to both \mathbb{H} and \mathbb{F} , and

$$\mathbb{H} \subseteq \mathbb{F}.$$

The above inclusion means that the government cannot directly observe the state Z of the economy, but that this has to be inferred through the observation of (X^0, η) . We are therefore working in a *partial information setting*.

2.2 The optimal debt reduction problem

The government can reduce the level of the debt-to-GDP ratio by intervening on the primary budget balance (i.e., the overall difference between government revenues and spending), for example through austerity policies in the form of spending cuts. When doing so, the debt ratio dynamics becomes

$$dX_t^v = (r - g(Z_t))X_t^v dt + \sigma X_t^v dW_t - dv_t, \quad X_{0-}^v = x > 0. \tag{2.3}$$

The process ν is the control that the government chooses based on the information at its disposal. More precisely, ν_t defines the cumulative reduction of the debt-to-GDP ratio made by the government up to time t , and ν is therefore a nondecreasing process belonging to the set

$$\begin{aligned} \mathcal{M}(x, \underline{y}, q) := \{ \nu : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ : (\nu_t(\omega) := \nu(\omega, t))_{t \geq 0} \text{ is nondecreasing,} \\ \text{right-continuous, } \mathbb{H}\text{-adapted and such that} \\ X_t^\nu \geq 0 \text{ for every } t \geq 0, X_{0-}^\nu = x, \\ \mathbb{P}[Z_0 = i] = y_i, i \in S, \eta_0 = q \text{ a.s.} \}, \end{aligned}$$

for any given and fixed initial value $x \in (0, \infty)$ of X^ν , initial value $q \in \mathcal{I}$ of η , and $\underline{y} \in \mathcal{Y}$. Here

$$\mathcal{Y} := \left\{ \underline{y} = (y_1, \dots, y_Q) : y_i \in [0, 1], i = 1, \dots, Q, \sum_{i=1}^Q y_i = 1 \right\}$$

is the probability simplex on \mathbb{R}^Q , representing the space of initial distributions of the process Z . **From now on, we set $\nu_{0-} = 0$ a.s.** for any $\nu \in \mathcal{M}(x, \underline{y}, q)$.

Remark 2.2 Notice that in the definition of the set \mathcal{M} above, as well as in (2.4) and in (2.5) below, we have stressed the dependency on the initial data (x, \underline{y}, q) just for notational convenience, not to indicate any Markovian nature of the considered problem, which is in fact not given.

For any $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$ and $\nu \in \mathcal{M}(x, \underline{y}, q)$, there exists a unique solution to (2.3), denoted by $X_t^{x, \nu}$, that is given by

$$X_t^{x, \nu} = X_t^{1,0} \left(x - \int_0^t \frac{d\nu_s}{X_s^{1,0}} \right), \quad t \geq 0, \quad X_{0-}^{x, \nu} = x,$$

where

$$X_t^{1,0} = e^{\int_0^t (r-g(Z_s))ds - \frac{1}{2}\sigma^2 t + \sigma W_t}, \quad t \geq 0.$$

Here and in the rest of this paper, we use the notation $\int_0^t (\cdot) d\nu_s = \int_{[0,t]} (\cdot) d\nu_s$ for the Lebesgue–Stieltjes integral with respect to the random measure $d\nu$ induced by the nondecreasing process ν on $[0, \infty)$.

Remark 2.3 The dynamics (2.3) might be justified in the following way. Suppose that the public debt (in real terms) D and the GDP Y follow the classical dynamics

$$\begin{cases} dD_t = rD_t dt - d\xi_t, & D_{0-} = d > 0, \\ dY_t = g(Z_t)Y_t dt + \sigma Y_t d\tilde{W}_t, & Y_0 = y > 0, \end{cases}$$

where ξ_t is the cumulative real budget balance up to time t and \tilde{W} is a Brownian motion. An easy application of Itô’s formula and a change of measure then gives that the ratio $X := D/Y$ evolves as in (2.3), upon setting $\nu := \int_0^\cdot d\xi_s/Y_s$ and $x := d/y$.

The government aims at reducing the level of the debt ratio. Having a level of $X_t = x$ at time $t \geq 0$ when the state of the economy is $Z_t = i$, the government incurs an instantaneous cost (loss) $h(x, i)$. This may be interpreted as an opportunity cost resulting from private investments' crowding out, less room for financing public investments, and from a tendency to suffer low subsequent growth (see the technical report [30] and the work by Woo and Kumar [63], among others). The cost function $h : \mathbb{R} \times S \mapsto \mathbb{R}_+$ fulfils the following requirements (see also Cadenillas and Huamán-Aguilar [8] and Ferrari [31]).

Assumption 2.4 (i) For any $i \in S$, the mapping $x \mapsto h(x, i)$ is strictly convex, continuously differentiable and nondecreasing on \mathbb{R}_+ . Moreover, $h(0, i) = 0$.

(ii) For any given $x \in (0, \infty)$ and $i \in S$, one has

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} h(X_t^{x,0}, i) dt \right] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} X_t^{1,0} h_x(X_t^{x,0}, i) dt \right] < \infty.$$

Remark 2.5 1) As an example, the power function given by $h(x, i) = \vartheta_i x^{n_i+1}$ for $(x, i) \in [0, \infty) \times S$, $\vartheta_i > 0$, $n_i \geq 1$ satisfies Assumption 2.4 (for a suitable $\rho > 0$ taking care of requirement (ii) above). Inspired by the careful discussion of Cadenillas and Huamán-Aguilar in [8, Sect. 2], n_i is a subjective regime-dependent parameter capturing the government's aversion/intolerance towards the debt ratio. On the other hand, the parameter ϑ_i can be thought of as a measure (in monetary terms) of the importance of debt: the better the debt's characteristics (for example, a larger portion of debt is domestic rather than external, cf. Japan), the lower the parameter ϑ_i (relative to marginal cost of intervention, see below). A power cost function as the one above is in line with the usual quadratic loss function adopted in the economic literature (see the influential paper by Tabellini [61], among many others).

2) Notice that the integrability conditions in Assumption 2.4 (ii) ensure that the expected cost and marginal cost of having debt and not intervening on it are finite for any possible regime of the economy. In particular, the finiteness of the second expectation in Assumption 2.4 (ii) guarantees that the stopping functional considered in Sect. 3.3 below is finite.

Whenever the government intervenes in order to reduce the debt-to-GDP ratio, it incurs a proportional cost. This might be seen as a measure of the social and financial consequences deriving from a debt-reduction policy, and the associated regime-dependent marginal cost $\kappa(Z_t)$ allows to express it in monetary terms (we refer e.g. to Lukkezen and Suyker [50] for an empirical evaluation of those costs). We assume that $\kappa(\cdot) > 0$ is a measurable finite function.

Given an intertemporal discount rate $\rho > 0$, for any given and fixed triple $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, the government thus aims to minimise the expected total cost functional

$$\mathcal{J}_{x, \underline{y}, q}(v) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} h(X_t^{x,v}, Z_t) dt + \int_0^\infty e^{-\rho t} \kappa(Z_t) dv_t \right] \tag{2.4}$$

for $v \in \mathcal{M}(x, \underline{y}, q)$. The government's problem under partial observation can be therefore defined as

$$V_{\text{po}}(x, \underline{y}, q) := \inf_{v \in \mathcal{M}(x, \underline{y}, q)} \mathcal{J}_{x, \underline{y}, q}(v), \quad (x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}. \quad (2.5)$$

Remark 2.6 1) We provide here some comments on our formulation of the optimal debt reduction problem. In line with the recent literature [8, 9, 31, 32] on stochastic control models for debt management, the cost/loss function h appearing in the government's objective functional is nondecreasing and null when the debt level is zero. While the latter requirement can be made without loss of generality, the former implicitly means that the government believes that disadvantages arising from debt far outweigh the advantages, and therefore neglects any potential social and financial benefit arising from having debt (cf. Holmström and Tirole [39]). One could think that this assumption is more appropriate for those countries that have faced severe debt crises during the last financial crisis and whose governments trust that high government debt has a negative effect on the long-term economic growth, makes the economy less resilient to macroeconomic shocks (e.g. sovereign default risks and liquidity shocks) and poses limits to the adoption of counter-cyclical fiscal policies (see e.g. the book by Blanchard [5, Chap. 22], the technical report [30] and Won and Kumar [63] for empirical studies).

However, it is also worth noticing that the general results of Sect. 3 of this paper still hold if we take $x \mapsto h(x, i)$ convex and bounded from below and remove the condition of being nondecreasing on \mathbb{R}_+ (thus allowing potential benefits arising from debt). On the other hand, the monotonicity of $h(\cdot, i)$ has an important role in our analysis of Sects. 3.3 and 4 (see Propositions 4.4 and 4.6).

2) In our model, we do not allow policies that might lead to an increase of the debt like e.g. investments in infrastructure, healthcare, education and research, and we neglect any possible social and financial benefit that those economic measures might induce (see Ostry et al. [55]). From a mathematical point of view, allowing policies of debt increase would lead to a singular stochastic control problem with controls of bounded variation, where the two nondecreasing processes giving the minimal decomposition of any admissible control represent the cumulative amount of the debt's increase and decrease. In this case, one might also allow in the government's objective functional the total expected social and financial benefits arising from a policy of debt expansion. We refer to Ferrari and Rodosthenous [32] where a similar setting has been considered in a problem of debt management under complete observation.

The function V_{po} is well defined and finite. Indeed, it is nonnegative due to the nonnegativity of h ; moreover, since the admissible policy "instantaneously reduce at initial time the debt ratio to 0" is a priori suboptimal and has cost x , we have $V_{\text{po}} \leq x$.

We should like to stress once more that any $v \in \mathcal{M}(x, \underline{y}, q)$ is \mathbb{H} -adapted, and therefore (2.5) is a stochastic control problem under partial observation. In particular, it is a *singular stochastic control problem under partial observation*, that is, an optimal control problem in which the random measures induced by the nondecreasing control processes on $[0, \infty)$ might be singular with respect to Lebesgue measure,

and in which one component Z of the state variable is not directly observable by the controller.

In its current formulation, the optimal debt reduction problem is not Markovian and the dynamic programming approach via an HJB equation is not applicable. In the next section, by using techniques from filtering theory, we introduce an equivalent problem under complete information, the so-called *separated problem*. This enjoys a Markovian structure, and its solution is characterised in Sect. 3.3 through a Markovian optimal stopping problem.

3 Reduction to an equivalent problem under complete information

In this section, we derive the *separated problem*. To this end, we first study the filtering problem arising in our model. As already discussed in the introduction, results on such a filtering problem cannot be directly obtained from existing literature due to the structure of our dynamics.

3.1 The filtering problem

The filtering problem consists in finding the best mean-squared estimate of $f(Z_t)$, for any t and any measurable function f , on the basis of the information available up to time t . In our setting, that information flow is given by the filtration \mathbb{H} . The estimate can be described through the filter process $(\pi_t)_{t \geq 0}$ which provides the conditional distribution of Z_t given \mathcal{H}_t for any time t (see for instance Liptser and Shiryaev [49, Chap. 8]). For any probability measure μ on $S = \{1, \dots, Q\}$ and any function f on S , we write $\mu(f) := \int_S f d\mu = \sum_{i=1}^Q f(i)\mu(\{i\})$. It is known that there exists a càdlàg (right-continuous with left limits) process taking values in the space of probability measures on $S = \{1, \dots, Q\}$ such that for any measurable function f on S ,

$$\pi_t(f) = \mathbb{E}[f(Z_t)|\mathcal{H}_t]; \quad (3.1)$$

see for further details Kurtz and Ocone [48, Lemma 1.1]. Moreover, since Z takes only a finite number of values, the filter is completely described by the vector

$$\pi_t(f_i) = \mathbb{P}[Z_t = i|\mathcal{H}_t], \quad i \in S,$$

where $f_i(z) := \mathbb{1}_{\{z=i\}}$, $i \in S$. With a slight abuse of notation, we denote in the following by $\pi(i)$ the process $\pi(f_i)$, so that for all measurable functions f , (3.1) gives

$$\pi_t(f) = \sum_{i=1}^Q f(i)\pi_t(i).$$

Setting $\beta(Z_t) := r - g(Z_t)$ and $\beta(i) := r - g(i)$, $i \in S$, notice that β is clearly a bounded function. Then we define two processes I and I^1 such that for any $t \geq 0$,

$$\begin{aligned}
 I_t &:= W_t - \int_0^t \sigma^{-1}(\pi_s(\beta) - \beta(Z_s))ds, \\
 I_t^1 &:= B_t - \int_0^t \left(\pi_s(\alpha(\eta_s, \cdot)) - \alpha(\eta_s, Z_s) \right) ds,
 \end{aligned}
 \tag{3.2}$$

where

$$\alpha(q, i) := \sigma_2(q)^{-1}(b_1(q, i) - \sigma^{-1}\beta(i)\sigma_1(q)), \quad (q, i) \in \mathcal{I} \times S. \tag{3.3}$$

Henceforth, we work under the following Novikov condition.

Assumption 3.1

$$\mathbb{E}\left[e^{\frac{1}{2} \int_0^t \alpha^2(\eta_s, Z_s) ds}\right] < \infty \quad \text{for any } t \geq 0.$$

Under Assumption 3.1, by classical results from filtering theory (see e.g. [49, Chap. 7]), the *innovation processes* I and I^1 are Brownian motions with respect to the filtration \mathbb{H} . Moreover, given the assumed independence of B and W , they turn out to be independent.

The integer-valued random measure associated to the jumps of η is defined as

$$m(dt, dq) := \sum_{s: \Delta\eta_s \neq 0} \delta_{(s, \Delta\eta_s)}(ds, dq), \tag{3.4}$$

where $\delta_{(a_1, a_2)}$ denotes the Dirac measure at the point $(a_1, a_2) \in \mathbb{R}_+ \times \mathbb{R}$. Notice that the \mathbb{H} -adapted random measure m is such that

$$\int_0^t c(\eta_{s-}, Z_{s-}) \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \neq 0\}} dN_s = \int_0^t \int_{\mathbb{R}} q m(ds, dq), \quad t > 0.$$

To proceed further we need the following useful definitions.

Definition 3.2 For any filtration \mathbb{G} , we denote by $\mathcal{P}(\mathbb{G})$ the predictable σ -field on the product space $(0, \infty) \times \Omega$. Moreover, let $\mathcal{B}(\mathbb{R})$ be the Borel σ -algebra on \mathbb{R} . Any mapping $H : (0, \infty) \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ which is $\mathcal{P}(\mathbb{G}) \times \mathcal{B}(\mathbb{R})$ -measurable is called a \mathbb{G} -predictable process indexed by \mathbb{R} .

Letting

$$\mathcal{F}_t^m := \sigma\{m((0, s] \times A) : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{R})\}, \tag{3.5}$$

we denote by $\mathbb{F}^m := (\mathcal{F}_t^m)_{t \geq 0}$ the filtration which is generated by the random measure $m(dt, dq)$. It is right-continuous by [7, Theorem T25 in Appendix A2].

Definition 3.3 Given any filtration \mathbb{G} with $\mathbb{F}^m \subseteq \mathbb{G}$, the \mathbb{G} -dual predictable projection of m , denoted by $m^{p, \mathbb{G}}(dt, dq)$, is the unique positive \mathbb{G} -predictable random measure such that for any nonnegative \mathbb{G} -predictable process Φ indexed by \mathbb{R} ,

$$\mathbb{E}\left[\int_0^\infty \int_{\mathbb{R}} \Phi(s, q) m(ds, dq)\right] = \mathbb{E}\left[\int_0^\infty \int_{\mathbb{R}} \Phi(s, q) m^{p, \mathbb{G}}(ds, dq)\right]. \tag{3.6}$$

To prove that a given positive \mathbb{G} -predictable random measure is the \mathbb{G} -dual predictable projection of m , it suffices to verify (3.6) for any process which has the form $\Phi(t, q) = C_t \mathbb{1}_A(q)$ with C a nonnegative \mathbb{G} -predictable process and $A \in \mathcal{B}(\mathbb{R})$. For further details, we refer to the books by Brémaud [7, Sect. VIII.4] and Jacod [41, Sect. III.1].

We now aim at deriving an equation for the evolution of the filter (the *filtering equation*). To this end, we use the so-called *innovation approach* (see Brémaud [7, Chap. IV.1], Liptser and Shiryaev [49, Chaps. 7.4 and 10.1.5] and Ceci and Colaneri [15], among others), which in our setting requires the introduction of the \mathbb{H} -compensated jump measure of η ,

$$m^\pi(dt, dq) := m(dt, dq) - m^{p, \mathbb{H}}(dt, dq). \tag{3.7}$$

The triplet (I, I^1, m^π) also represents a building block for the construction of \mathbb{H} -martingales as shown in Proposition 3.5 below. We start by determining the form of $m^{p, \mathbb{H}}$.

Proposition 3.4 *The \mathbb{H} -dual predictable projection of m is given by*

$$m^{p, \mathbb{H}}(dt, dq) = \sum_{i=1}^Q \pi_{t-}(i) \lambda^N(i) \mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} \delta_{c(\eta_{t-}, i)}(dq) dt, \tag{3.8}$$

where δ_a denotes the Dirac measure at the point $a \in \mathbb{R}$.

Proof 1) We first prove that the \mathbb{F} -dual predictable projection of m is given by

$$m^{p, \mathbb{F}}(dt, dq) := \lambda^N(Z_{t-}) \mathbb{1}_{\{c(\eta_{t-}, Z_{t-}) \neq 0\}} \delta_{c(\eta_{t-}, Z_{t-})}(dq) dt. \tag{3.9}$$

Let $A \in \mathcal{B}(\mathbb{R})$ and introduce

$$\mathcal{N}_t(A) := m((0, t] \times A) = \sum_{s \leq t} \mathbb{1}_{\{\Delta \eta_s \in A \setminus \{0\}\}}, \quad t \geq 0. \tag{3.10}$$

Then $\mathcal{N}(A)$ is the point process counting the number of jumps of η up to time t with jump size in the set A . Since (2.2) implies that $\Delta \eta_s = c(\eta_{s-}, Z_{s-}) \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \neq 0\}} \Delta N_s$ for $s \geq 0$ and N is a point process with \mathbb{F} -predictable intensity given by $(\lambda^N(Z_{t-}))_{t \geq 0}$, we obtain for each nonnegative \mathbb{F} -predictable process C that

$$\begin{aligned} \mathbb{E} \left[\int_0^t C_s d\mathcal{N}_s(A) \right] &= \mathbb{E} \left[\int_0^t C_s \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \in A \setminus \{0\}\}} dN_s \right] \\ &= \mathbb{E} \left[\int_0^t C_s \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \in A \setminus \{0\}\}} \lambda^N(Z_{s-}) ds \right]. \end{aligned}$$

So for any $A \in \mathcal{B}(\mathbb{R})$, $(\lambda^N(Z_{t-}) \mathbb{1}_{\{c(\eta_{t-}, Z_{t-}) \in A \setminus \{0\}\}})_{t \geq 0}$ provides the \mathbb{F} -predictable intensity of the counting process $\mathcal{N}(A)$. Recalling (3.10) and Definition 3.3, this implies that $m^{p, \mathbb{F}}(dt, dq)$ in (3.9) coincides with the \mathbb{F} -dual predictable projection of m , since (3.6) holds with the choice $\mathbb{G} = \mathbb{F}$ and $\Phi(t, q) = C_t \mathbb{1}_A(q)$.

2) As in Ceci [11, Proposition 2.3], we can now derive the \mathbb{H} -dual predictable projection of $m^{p,\mathbb{F}}$ by projecting $m^{p,\mathbb{F}}$ onto the observation flow \mathbb{H} . More precisely, the \mathbb{H} -predictable intensity of the point process $\mathcal{N}(A)$, $A \in \mathcal{B}(\mathbb{R})$, is given by

$$\pi_{t-}(\lambda^N(\cdot)\mathbb{1}_{\{c(\eta_{t-},\cdot) \in A \setminus \{0\}\}}) = \sum_{i=1}^Q \pi_{t-}(i)\lambda^N(i)\mathbb{1}_{\{c(\eta_{t-},i) \in A \setminus \{0\}\}}, \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

This implies that $m^{p,\mathbb{H}}(dt, dq)$ is given by (3.8), since (3.6) is satisfied with the choice $\mathbb{G} = \mathbb{H}$, $\Phi(t, q) = C_t \mathbb{1}_A(q)$. □

An essential tool to prove that the original problem under partial information is equivalent to the separated one is the characterisation of the filter as the unique solution to the filtering equation (see El Karoui et al. [28], Mazliak [52] and Ceci and Gerardi [12]). In order to derive the filtering equation solved by π , we first give a representation theorem for \mathbb{H} -martingales. The proof of the following technical result is given in Appendix A.

Proposition 3.5 *Under Assumptions 2.1 and 3.1, every \mathbb{H} -local martingale M admits the decomposition*

$$M_t = M_0 + \int_0^t \varphi_s dI_s + \int_0^t \psi_s dI_s^1 + \int_0^t \int_{\mathbb{R}} w(s, q) m^\pi(ds, dq),$$

where φ and ψ are \mathbb{H} -predictable processes and w is an \mathbb{H} -predictable process indexed by \mathbb{R} such that a.s.

$$\int_0^t \varphi_s^2 ds < \infty, \quad \int_0^t \psi_s^2 ds < \infty, \quad \int_0^t \int_{\mathbb{R}} |w(s, q)| m^{p,\mathbb{H}}(ds, dq) < \infty, \quad t \geq 0.$$

We are now in the position to prove the following fundamental result, whose proof is postponed to Appendix A.

Theorem 3.6 *Recall (3.7), let $\underline{y} \in \mathcal{Y}$ be the initial distribution of Z and let Assumptions 2.1 and 3.1 hold. Then the filter $(\underline{\pi}_t)_{t \geq 0} := (\pi_t(i); i \in S)_{t \geq 0}$ solves the Kushner–Stratonovich system*

$$\begin{aligned} \pi_t(i) = & y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left(\beta(i) - \sum_{j=1}^Q \beta(j) \pi_s(j) \right) dI_s \\ & + \int_0^t \pi_s(i) \left(\alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right) dI_s^1 \\ & + \int_0^t \int_{\mathbb{R}} (w_i^\pi(s, q) - \pi_{s-}(i)) m^\pi(ds, dq) \end{aligned} \tag{3.11}$$

for any $i \in S$. Here, $\beta(i) = r - g(i)$ and

$$w_i^\pi(s, q) := \frac{d\lambda^N(i)\pi_{s-}(i)\mathbb{1}_{\{c(\eta_{s-},i)\neq 0\}}\delta_{c(\eta_{s-},i)}(dq)}{d(\sum_{j=1}^Q \pi_{s-}(j)\lambda^N(j)\mathbb{1}_{\{c(\eta_{s-},j)\neq 0\}}\delta_{c(\eta_{s-},j)}(dq))} \tag{3.12}$$

denotes the Radon–Nikodým derivative of $\lambda^N(i)\pi_{s-}(i)\mathbb{1}_{\{c(\eta_{s-},i)\neq 0\}}\delta_{c(\eta_{s-},i)}(dq)$ with respect to $\sum_{j=1}^Q \pi_{s-}(j)\lambda^N(j)\mathbb{1}_{\{c(\eta_{s-},j)\neq 0\}}\delta_{c(\eta_{s-},j)}(dq)$.

Let us introduce the sequence of jump times and jump sizes of the process η , denoted by $(T_n, \zeta_n)_{n \geq 1}$ and recursively defined, with $T_0 := 0$, as

$$T_{n+1} := \inf \left\{ t > T_n : \int_{T_n}^t c(\eta_{s-}, Z_{s-}) dN_s \neq 0 \right\},$$

$$\zeta_n := \eta_{T_n} - \eta_{T_n-} = c(\eta_{T_n-}, Z_{T_n-}), \quad n \geq 1.$$

We use the standard convention that $\inf \emptyset = +\infty$. Then the integer-valued measure associated to the jumps of η (cf. (3.4)) can also be written as

$$m(dt, dq) = \sum_{n \geq 1} \delta_{(T_n, \zeta_n)}(dt, dq) \mathbb{1}_{\{T_n < +\infty\}}. \tag{3.13}$$

The filtering system (3.11) has a natural recursive structure in terms of the sequence $(T_n)_{n \geq 1}$, as shown in the next proposition.

Proposition 3.7 *Between two consecutive jump times, i.e., for $t \in [T_n, T_{n+1})$, the filtering system (3.11) reads as*

$$\begin{aligned} \pi_t(i) = & \pi_{T_n}(i) + \int_{T_n}^t \left(\sum_{j=1}^Q \lambda_{ji} \pi_s(j) - \pi_s(i) \left(\lambda^N(i) \mathbb{1}_{\{c(\eta_{s-},i)\neq 0\}} \right. \right. \\ & \left. \left. - \sum_{j=1}^Q \lambda^N(j) \pi_s(j) \mathbb{1}_{\{c(\eta_{s-},j)\neq 0\}} \right) \right) ds \\ & + \int_{T_n}^t \sigma^{-1} \pi_s(i) \left(\beta(i) - \sum_{j=1}^Q \beta(j) \pi_s(j) \right) dI_s \\ & + \int_{T_n}^t \pi_s(i) \left(\alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right) dI_s^1 \end{aligned} \tag{3.14}$$

for any $i \in S$. At a jump time T_n of η , $(\underline{\pi}_t)_{t \geq 0} = (\pi_t(i); i \in S)_{t \geq 0}$ jumps as well, and its value is given by

$$\pi_{T_n}(i) = \frac{\lambda^N(i)\pi_{T_n-}(i)\mathbb{1}_{\{\zeta_n=c(\eta_{T_n-},i)\}}}{\sum_{j=1}^Q \lambda^N(j)\pi_{T_n-}(j)\mathbb{1}_{\{\zeta_n=c(\eta_{T_n-},j)\}}}, \quad i \in S. \tag{3.15}$$

Proof First, recalling that $m^\pi(dt, dq) = m(dt, dq) - m^{p, \mathbb{H}}(dt, dq)$ and

$$m^{p, \mathbb{H}}(dt, dq) = \sum_{j=1}^Q \pi_{t-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{t-}, j) \neq 0\}} \delta_{c(\eta_{t-}, j)}(dq) dt,$$

we obtain that

$$\begin{aligned} & \int_0^t \int_{\mathbb{R}} (w_i^\pi(s, q) - \pi_{s-}(i)) m^{p, \mathbb{H}}(ds, dq) \\ &= \int_0^t \pi_s(i) \left(\lambda^N(i) \mathbb{1}_{\{c(\eta_{s-}, i) \neq 0\}} - \sum_{j=1}^Q \lambda^N(j) \pi_s(j) \mathbb{1}_{\{c(\eta_{s-}, j) \neq 0\}} \right) ds, \end{aligned}$$

which from (3.11) implies that $\pi_t(i)$ solves (3.14) for any $t \in [T_n, T_{n+1})$. Finally, (3.15) follows by (3.12) and

$$\pi_{T_n}(i) = w_i^\pi(T_n, \zeta_n) = \frac{\lambda^N(i) \pi_{T_n-}(i) \mathbb{1}_{\{c(\eta_{T_n-}, i) \neq 0\}} \delta_{c(\eta_{T_n-}, i)}(\zeta_n)}{\sum_{j=1}^Q \pi_{T_n-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{T_n-}, j) \neq 0\}} \delta_{c(\eta_{T_n-}, j)}(\zeta_n)}. \quad \square$$

We want to stress that (3.15) shows that the vector $\underline{\pi}_{T_n}$ is completely determined by the observed data η and the knowledge of $\underline{\pi}_t$ for $t \in [T_{n-1}, T_n)$, since $\pi_{T_n-}(i) := \lim_{t \uparrow T_n} \pi_t(i)$, $i \in S$.

Example 3.8 1) In the case $c(q, i) \equiv c \neq 0$ for any $i \in S$ and $q \in \mathcal{I}$, the sequences of jump times of η and N coincide and the filtering system (3.11) reduces to (for $i \in S$)

$$\begin{aligned} \pi_t(i) &= y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left(\beta(i) - \sum_{j=1}^Q \beta(j) \pi_s(j) \right) dI_s \\ &+ \int_0^t \pi_s(i) \left(\alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right) dI_s^1 \\ &+ \int_0^t \left(\frac{\lambda^N(i) \pi_{s-}(i)}{\sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j)} - \pi_{s-}(i) \right) \left(dN_s - \sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j) ds \right). \end{aligned}$$

2) In the case $\alpha(q, i) = \alpha(i)$ and $c(q, i) \equiv 0$ for any $i \in S$ and $q \in \mathcal{I}$, the filtering system (3.11) does not depend explicitly on the process η . In particular, one has

$$\begin{aligned} \pi_t(i) &= y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left(\beta(i) - \sum_{j=1}^Q \beta(j) \pi_s(j) \right) dI_s \\ &+ \int_0^t \pi_s(i) \left(\alpha(i) - \sum_{j=1}^Q \alpha(j) \pi_s(j) \right) dI_s^1, \quad i \in S, \end{aligned}$$

where we have set $\alpha(i) := \sigma_2^{-1}(b_1(i) - \sigma^{-1}\beta(i)\sigma_1)$. In Sect. 4, we provide the explicit solution to the optimal debt reduction problem within this setting. With reference to (2.2) and (3.3), this setting corresponds e.g. to the purely diffusive arithmetic case $c(q, i) = 0, b_1(q, i) = b_1(i)$ and $\sigma_1(q) = \sigma_1 > 0, \sigma_2(q) = \sigma_2 > 0$ for any $i \in S$ and $q \in \mathcal{I}$, or to the purely diffusive geometric case $c(q, i) = 0, b_1(q, i) = b_1(i)q$ and $\sigma_1(q) = \sigma_1q, \sigma_2(q) = \sigma_2q$ for any $i \in S$ and $q \in \mathcal{I}$.

3.2 The separated problem

Thanks to the introduction of the filter, (2.1)–(2.3) can now be rewritten in terms of observable processes. In particular, we have that

$$\begin{aligned} dX_t^0 &= \pi_t(\beta)X_t^0 dt + \sigma X_t^0 dI_t, \\ X_0^0 &= x > 0, \end{aligned} \tag{3.16}$$

$$\begin{aligned} d\eta_t &= \pi_t(b_1(\eta_t, \cdot))dt + \sigma_1(\eta_t)dI_t + \sigma_2(\eta_t)dI_t^1 + \int_{\mathbb{R}} \zeta m(dt, d\zeta), \\ \eta_0 &= q \in \mathcal{I}, \end{aligned} \tag{3.17}$$

$$\begin{aligned} dX_t^v &= \pi_t(\beta)X_t^v dt + \sigma X_t^v dI_t - dv_t, \\ X_{0-}^v &= x > 0. \end{aligned} \tag{3.18}$$

Notice that for any $v \in \mathcal{M}(x, \underline{y}, q)$, the process X^v turns out to be \mathbb{H} -adapted and depends on the vector $(\underline{\pi}_t)_{t \geq 0} = (\pi_t(i); i \in S)_{t \geq 0}$ with $\underline{\pi}_0 = \underline{y} \in \mathcal{Y}$.

Definition 3.9 We say that a process $(\tilde{\pi}_t, \tilde{\eta}_t)_{t \geq 0}$ with values in $\mathcal{Y} \times \mathcal{I}$ is a *strong solution* to (3.11) and (3.17) if it satisfies those equations pathwise. We say that *strong uniqueness* for the system (3.11) and (3.17) holds if for any strong solution $(\tilde{\pi}_t, \tilde{\eta}_t)_{t \geq 0}$ to (3.11) and (3.17), one has $\tilde{\pi}_t = \underline{\pi}_t$ and $\tilde{\eta}_t = \eta_t$ a.s. for all $t \geq 0$.

Proposition 3.10 *Let Assumptions 2.1 and 3.1 hold and suppose that $\alpha(\cdot, i)$ is locally Lipschitz for any $i \in S$ and there exists $M > 0$ such that $|\alpha(q, i)| \leq M(1 + |q|)$ for any $q \in \mathcal{I}$ and any $i \in S$. Then the system (3.11) and (3.17) admits a unique strong solution.*

The proof of Proposition 3.10 is postponed to Appendix A. Notice that under Assumption 2.1, the requirement on α of Proposition 3.10 is verified e.g. whenever $\sigma_2(q) \geq \underline{\sigma}$ for some $\underline{\sigma} > 0$ and for any $q \in \mathcal{I}$, or if b_1/σ_2 and σ_1/σ_2 are locally Lipschitz in $q \in \mathcal{I}$ and have sublinear growth. As a byproduct of Proposition 3.10, we also have strong uniqueness of the solution to (3.18). In the following, when there is a need to stress the dependence with respect to the initial value $x > 0$, we denote the solution to (3.16) and (3.18) by $X^{x,0}$ and $X^{x,v}$, respectively. Since

$$\mathbb{E}[\pi_t(h(X_t^{x,v}, \cdot))] = \mathbb{E}[\mathbb{E}[h(X_t^{x,v}, Z_t)|\mathcal{H}_t]],$$

an application of the Fubini–Tonelli theorem allows writing

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} h(X_t^{x,v}, Z_t) dt \right] = \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} \pi_t(h(X_t^v, \cdot)) dt \right],$$

where $\mathbb{E}_{(x, \underline{y}, q)}$ denotes the expectation conditioned on $X_{0-}^v = x > 0$, $\underline{x}_0 = \underline{y} \in \mathcal{Y}$, and $\eta_0 = q \in \mathcal{I}$. Also, because $\pi(\kappa)$ is the \mathbb{H} -optional projection of the process $\kappa(Z)$ (cf. [48, Lemma 1.1]) and any admissible control v is increasing and \mathbb{H} -adapted, an application of Dellacherie and Meyer [24, Theorem VI.57, in particular (VI.57.1)] yields

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} \kappa(Z_t) dv_t \right] = \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} \pi_t(\kappa(\cdot)) dv_t \right].$$

Hence, the cost functional of (2.4) can be rewritten in terms of observable quantities as

$$\mathcal{J}_{x, \underline{y}, q}(v) = \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} \pi_t(h(X_t^v, \cdot)) dt + \int_0^\infty e^{-\rho t} \pi_t(\kappa(\cdot)) dv_t \right].$$

Notice that the latter expression does not depend on the unobservable process Z any more, and this allows us to introduce a control problem with complete information, *the separated problem*, in which the new state variable is given by the triplet $(X^v, \underline{x}, \eta)$. For this problem, we introduce the set $\mathcal{A}(x, \underline{y}, q)$ of admissible controls, given in terms of the observable processes in (3.11), (3.17) and (3.18) as

$$\begin{aligned} \mathcal{A}(x, \underline{y}, q) := \{ & v : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ : (v_t(\omega) := v(\omega, t))_{t \geq 0}, \text{ is nondecreasing,} \\ & \text{right-continuous, } \mathbb{H}\text{-adapted and such that} \\ & X_t^{x,v} \geq 0 \text{ for every } t \geq 0, X_{0-}^{x,v} = x, \\ & \underline{x}_0 = \underline{y}, \eta_0 = q \text{ a.s.} \} \end{aligned}$$

for every initial value $x \in (0, \infty)$ of $X^{x,v}$ defined in (3.18), any initial value $\underline{y} \in \mathcal{Y}$ of the process $(\underline{x}_t)_{t \geq 0} = (\pi_t(i); i \in S)_{t \geq 0}$ solving (3.11) and any initial value $q \in \mathcal{I}$ of η . In the following, we set $v_{0-} = 0$ a.s. for any $v \in \mathcal{A}(x, \underline{y}, q)$.

Given $v \in \mathcal{A}(x, \underline{y}, q)$, the triplet $(X_t^{x,v}, \underline{x}_t, \eta_t)_{t \geq 0}$ solves (3.18), (3.11) and (3.17) and the jump measure associated to η has \mathbb{H} -predictable dual projection given by (3.8). Hence, the process $(X_t^{x,v}, \underline{x}_t, \eta_t)_{t \geq 0}$ is an \mathbb{H} -Markov process and we therefore define the Markovian separated problem as

$$\left\{ \begin{aligned} V(x, \underline{y}, q) &:= \inf_{v \in \mathcal{A}(x, \underline{y}, q)} \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} \pi_t(h(X_t^v, \cdot)) dt \right. \\ &\quad \left. + \int_0^\infty e^{-\rho t} \pi_t(\kappa(\cdot)) dv_t \right] \\ dX_t^{x,v} &= \pi_t(\beta) X_t^{x,v} dt + \sigma X_t^{x,v} dI_t - dv_t, \quad X_{0-}^{x,v} = x > 0, \\ &(\underline{x}, \eta) \text{ solution to (3.11) and (3.17).} \end{aligned} \right. \tag{3.19}$$

This is now a *singular stochastic problem under complete information*, since all the processes involved are \mathbb{H} -adapted.

The next proposition immediately follows from the construction of the separated problem and the strong uniqueness of the solutions to (3.11), (3.17) and (3.18).

Proposition 3.11 *Assume strong uniqueness for the system of (3.11) and (3.17), and let $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$ be the initial values of the process (X, Z, η) in the problem (2.5) under partial observation. Then*

$$V_{po}(x, \underline{y}, q) = V(x, \underline{y}, q).$$

Moreover, $\mathcal{A}(x, \underline{y}, q) = \mathcal{M}(x, \underline{y}, q)$ and v^* is an optimal control for the separated problem (3.19) if and only if it is optimal for the original problem (2.5) under partial observation.

Remark 3.12 Notice that in the setting of Example 3.8, 2), the pair $(X^{x,v}, \underline{\pi})$ solving (3.18) and (3.11) is an \mathbb{H} -Markov process for any given control $v \in \mathcal{A}(x, \underline{y}, q)$, $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$. As a consequence, since the cost functional and the set of admissible controls do not depend explicitly on the process η , the value function of the separated problem (3.19) does not depend on the variable q . We consider this setting in Sect. 4.

3.3 A probabilistic verification theorem via reduction to optimal stopping

In this section, we relate the separated problem to a Markovian optimal stopping problem and show that the solution to the latter is directly related to the optimal control of the former. The following analysis is fully probabilistic and based on a change-of-variable formula for Lebesgue–Stieltjes integrals that has already been employed in singular control problems (see e.g. Baldursson and Karatzas [2] and Ferrari [31]). The result of this section is then employed in Sect. 4 where in a case study, we determine the optimal debt reduction policy by solving an auxiliary optimal stopping problem.

With regard to (3.19), notice that we can write $\pi_t(\kappa(\cdot)) = \sum_{i=1}^Q \pi_t(i)\kappa(i)$ as well as $\pi_t(h(X_t^{x,v}, \cdot)) = \sum_{i=1}^Q \pi_t(i)h(X_t^{x,v}, i)$ a.s. for any $t \geq 0$. For any $(x, \underline{\pi})$ in $(0, \infty) \times \mathcal{Y}$, set

$$\widehat{h}(x, \underline{\pi}) := \sum_{i=1}^Q \pi(i)h(x, i), \quad \widehat{\kappa}(\underline{\pi}) := \sum_{i=1}^Q \pi(i)\kappa(i),$$

and given $z \in (0, \infty)$, we introduce the optimal stopping problem

$$\begin{aligned} \widetilde{U}_t(z) := \operatorname{ess\,inf}_{\tau \geq t} \mathbb{E} & \left[\int_t^\tau e^{-\rho(s-t)} X_s^{1,0} \widehat{h}_x(X_s^{z,0}, \underline{\pi}_s) ds \right. \\ & \left. + e^{-\rho(\tau-t)} \widehat{\kappa}(\underline{\pi}_\tau) X_\tau^{1,0} \middle| \mathcal{H}_t \right], \quad t \geq 0, \end{aligned} \tag{3.20}$$

where the optimisation is taken over all \mathbb{H} -stopping times $\tau \geq t$.

Under Assumption 2.4, the expectation in (3.20) is finite for any \mathbb{H} -stopping time $\tau \geq t$, for any $t \geq 0$. Observing that $\kappa(i) < \infty$ for any $i \in S$, in order to take care of the event $\{\tau = \infty\}$, we use in (3.20) the convention

$$e^{-\rho\tau} X_\tau^{1,0} := \liminf_{t \uparrow \infty} e^{-\rho t} X_t^{1,0} \quad \text{on } \{\tau = \infty\}. \tag{3.21}$$

Denote by $U(z)$ a càdlàg modification of $\tilde{U}(z)$ (which under our assumptions exists due to the results in Karatzas and Shreve [46, Appendix D]), and observe that $0 \leq U_t(z) \leq \widehat{\kappa}(\underline{\pi}_t) X_t^{1,0}$ for any $t \geq 0$, a.s. Also, define the stopping time

$$\tau_t^*(z) := \inf\{s \geq t : U_s(z) \geq \widehat{\kappa}(\underline{\pi}_s) X_s^{1,0}\}, \quad z \in (0, \infty), \tag{3.22}$$

with the usual convention $\inf \emptyset = \infty$. Then by [46, Theorem D.12], $\tau_t^*(z)$ is an optimal stopping time for (3.20). In particular, $\tau^*(z) := \tau_0^*(z)$ is optimal for the problem

$$U_0(z) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} X_t^{1,0} \widehat{h}_x(X_t^{z,0}, \underline{\pi}_t) dt + e^{-\rho\tau} \widehat{\kappa}(\underline{\pi}_\tau) X_\tau^{1,0} \right].$$

Notice that since $h_x(\cdot, \underline{\pi})$ is a.s. increasing, $z \mapsto \tau^*(z)$ is a.s. decreasing. This monotonicity of $\tau^*(\cdot)$ will be important in the sequel as we need to consider its generalised inverse. Moreover, since the triplet $(X^{z,0}, \underline{\pi}, \eta)$ is a homogeneous \mathbb{H} -Markov process, there exists a measurable function $U : (0, \infty) \times \mathcal{Y} \times \mathcal{I} \rightarrow \mathbb{R}$ such that $U_t(z) = U(X_t^{z,0}, \underline{\pi}_t, \eta_t)$ for any $t \geq 0$, a.s. Hence $U_0(z) = U(z, \underline{y}, q)$, and for any $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, we define

$$\tilde{V}(x, \underline{y}, q) := \int_0^x U(z, \underline{y}, q) dz. \tag{3.23}$$

Moreover, we introduce the nondecreasing right-continuous process

$$\bar{v}_t^* := \sup\{\alpha \in [0, x] : \tau^*(x - \alpha) \leq t\}, \quad t \geq 0, \bar{v}_{0-}^* = 0, \tag{3.24}$$

and then also the process

$$v_t^* := \int_0^t X_s^{1,0} d\bar{v}_s^*, \quad t > 0, v_{0-}^* = 0.$$

Notice that \bar{v}^* is the right-continuous inverse of $\tau^*(\cdot)$.

Theorem 3.13 *Let \tilde{V} be as in (3.23) and V as in the definition (3.19). Then $\tilde{V} = V$, and v^* is the (unique) optimal control for (3.19).*

Proof 1) Let $x > 0$, $\underline{y} \in \mathcal{Y}$ and $q \in \mathcal{I}$ be given and fixed. For $v \in \mathcal{A}(x, \underline{y}, q)$, we introduce the process \bar{v} such that $\bar{v}_t := \int_0^t \frac{dv_x}{X_s^{1,0}}$, $t \geq 0$, and define its inverse (see e.g. Revuz and Yor [59, Sect. 0.4]) by

$$\tau^{\bar{v}}(z) := \inf\{t \geq 0 : x - \bar{v}_t < z\}, \quad 0 < z \leq x.$$

Notice that the process $(\tau^{\bar{v}}(z))_{z \leq x}$ has decreasing left-continuous sample paths, and hence it admits right limits

$$\tau_+^{\bar{v}}(z) := \inf\{t \geq 0 : x - \bar{v}_t \leq z\}, \quad z \leq x. \tag{3.25}$$

Moreover, the set of points $z \in \mathbb{R}$ at which $\tau^{\bar{v}}(z)(\omega) \neq \tau_+^{\bar{v}}(z)(\omega)$ is a.s. countable for a.e. $\omega \in \Omega$. The random time $\tau^{\bar{v}}(z)$ is actually an \mathbb{H} -stopping time because it is the entry time into an open set of the right-continuous process \bar{v} , and \mathbb{H} is right-continuous. Moreover, since $\tau_+^{\bar{v}}(z)$ is the first entry time of the right-continuous process \bar{v} into a closed set, it is an \mathbb{H} -stopping time as well for any $z \leq x$.

Proceeding then as in Ferrari [31, Step 1 of the proof of Theorem 3.1], by employing the change-of-variable formula in [59, Proposition 0.4.9], one finds that

$$\tilde{V}(x, \underline{y}, q) = \int_0^x U(z, \underline{y}, q) dz \leq \mathcal{J}_{x, \underline{y}, q}(v).$$

Hence, since v was arbitrary, we find that

$$\tilde{V}(x, \underline{y}, q) \leq V(x, \underline{y}, q), \quad (x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}. \tag{3.26}$$

2) To complete the proof, we have to show the reverse inequality. Let $x \in (0, \infty)$, $\underline{y} \in \mathcal{Y}$ and $q \in \mathcal{I}$ be initial values of $X^{x, v}$, $\underline{\pi}$ and η . We first notice that $v^* \in \mathcal{A}(x, \underline{y}, q)$. Indeed, v^* is nondecreasing, right-continuous and such that $X_t^{x, v^*} = X_t^{1, 0}(x - \bar{v}_t^*) \geq 0$ a.s. for all $t \geq 0$, since $\bar{v}_t^* \leq x$ a.s. by definition. Moreover, for any $0 < z \leq x$, we can write by (3.24) and (3.25) that

$$\tau_+^{\bar{v}^*}(z) \leq t \iff \bar{v}_t^* \geq x - z \iff \tau^*(z) \leq t.$$

Then recalling that $\tau_+^{\bar{v}^*}(z) = \tau^{\bar{v}^*}(z)$ \mathbb{P} -a.s. for almost every $z \leq x$, we pick $v = v^*$ (equivalently, $\bar{v} = \bar{v}^*$) and following [31, Step 2 in the proof of Theorem 3.1], we obtain $\tilde{V}(x, \underline{y}, q) = \mathcal{J}_{x, \underline{y}, q}(v^*) \geq V(x, \underline{y}, q)$, where the last inequality is due to the admissibility of v^* . Hence, by (3.26), we have $\tilde{V} = V$ and v^* is optimal. In fact, by strict convexity of $\mathcal{J}_{x, \underline{y}, q}(\cdot)$, v^* is the unique optimal control in the class of controls belonging to $\mathcal{A}(x, \underline{y}, q)$ and such that $\mathcal{J}_{x, \underline{y}, q}(v) < \infty$. \square

Remark 3.14 For any given $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, define the Markovian optimal stopping problem

$$v(x, \underline{y}, q) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} X_t^{x, 0} \widehat{h}_x(X_t^{x, 0}, \underline{\pi}_t^{\underline{y}}) dt + e^{-\rho \tau} \widehat{\kappa}(\underline{\pi}_\tau^{\underline{y}}) X_\tau^{x, 0} \right],$$

where $\underline{\pi}^{\underline{y}}$ denotes the filter process starting at time zero from $\underline{y} \in \mathcal{Y}$. Then, since $X^{x, 0} = x X^{1, 0}$ by (3.16) and $U_0(z) = U(z, \underline{y}, q)$ for some measurable function U , one can easily see that $v(x, \underline{y}, q) = x U(x, \underline{y}, q)$. Moreover, the previous considerations together with (3.22) (evaluated at $t = 0$) ensure that the stopping time

$$\tau^*(x, \underline{y}, q) := \inf\{t \geq 0 : v(X_t^{x, 0}, \underline{\pi}_t^{\underline{y}}, \eta_t^q) \geq \widehat{\kappa}(\underline{\pi}_t^{\underline{y}}) X_t^{x, 0}\}$$

is optimal for $v(x, \underline{y}, q)$, where $q = \eta_0$.

4 The solution in a case study with $Q = 2$ economic regimes

In this section, we build on the general filtering analysis developed in the previous sections and on the result of Theorem 3.13, and provide the form of the optimal debt reduction policy in a case study defined through the following **standing assumption**.

Assumption 4.1 1) Z takes values in $S = \{1, 2\}$, and with reference to (2.3), we assume $g_2 := g(2) < g(1) =: g_1$.

2) For any $q \in \mathcal{I}$ and $i \in \{1, 2\}$, one has $c(q, i) = 0$, and for α as in (3.3), we assume $\alpha(q, i) = \alpha(i)$.

3) $h(x, i) = h(x)$ for all $(x, i) \in (0, \infty) \times \{1, 2\}$, with $h : \mathbb{R} \rightarrow \mathbb{R}$ such that

(i) $x \mapsto h(x)$ is strictly convex, twice continuously differentiable and nondecreasing on \mathbb{R}_+ with $h(0) = 0$ and $\lim_{x \uparrow \infty} h(x) = \infty$;

(ii) there exist $\gamma > 1, 0 < K_o < K$ and $K_1, K_2 > 0$ such that

$$K_o|x^+|^\gamma - K \leq h(x) \leq K(1 + |x|^\gamma),$$

$$|h'(x)| \leq K_1(1 + |x|^{\gamma-1}),$$

$$|h''(x)| \leq K_2(1 + |x|^{(\gamma-2)^+}).$$

4) $\kappa(i) = 1$ for $i \in \{1, 2\}$.

Notice that under Assumption 4.1, 2), the macroeconomic indicator η has a suitable diffusive dynamics whose coefficients b_1, σ_1, σ_2 are such that the function α is independent of q . As discussed in Example 3.8, 2), this is the case of a geometric or arithmetic diffusive dynamics for η . In this setting, the Kushner–Stratonovich system (3.11) reduces to

$$\begin{aligned} d\pi_t(1) &= (\lambda_2 - (\lambda_1 + \lambda_2)\pi_t(1))dt \\ &+ \pi_t(1)(1 - \pi_t(1))\left(\frac{\beta_1 - \beta_2}{\sigma}dI_t + (\alpha_1 - \alpha_2)dI_t^1\right) \end{aligned} \tag{4.1}$$

and $\pi_t(2) = 1 - \pi_t(1)$. Here, $\lambda_1 := \lambda_{12} > 0$ and $\lambda_2 := \lambda_{21} > 0$.

Setting $\pi_t := \pi_t(1), t \geq 0$, (3.19) then reads as

$$\left\{ \begin{aligned} V(x, y) &= \inf_{v \in \mathcal{A}(x,y)} \mathbb{E}_{(x,y)} \left[\int_0^\infty e^{-\rho t} h(X_t^v) dt + \int_0^\infty e^{-\rho t} dv_t \right], \\ dX_t^{x,y,v} &= (\beta_2 + \pi_t^y(g_2 - g_1))X_t^{x,y,v} dt + \sigma X_t^{x,y,v} dI_t - dv_t, \\ d\pi_t^y &= (\lambda_2 - (\lambda_1 + \lambda_2)\pi_t^y)dt \\ &+ \pi_t^y(1 - \pi_t^y)\left(\frac{g_2 - g_1}{\sigma}dI_t + (\alpha_1 - \alpha_2)dI_t^1\right), \end{aligned} \right. \tag{4.2}$$

with initial conditions $X_{0-}^{x,y,v} = x > 0, \pi_0 = y \in (0, 1)$, and where $g_i = r - \beta_i$ denotes the rate of economic growth in the state $i, i = 1, 2$. Note that we switch here from arguments (i) to subscripts i .

It is worth noticing that there is no need to involve the process η in the Markovian formulation (4.2). This is due to the fact that the couple (X^ν, π) solving the two stochastic differential equations above is a strong Markov process and the cost functional and the set of admissible controls $\mathcal{A}(x, y, q)$ do not depend explicitly on η ; hence we simply write $\mathcal{A}(x, y)$ instead of $\mathcal{A}(x, y, q)$ (cf. (4.2) above). For this reason, the value function of (4.2) does not depend on the initial value q of the process η . However, the memory of the macroeconomic indicator process η appears in the filter π through the constant term $\alpha_1 - \alpha_2$ in its dynamics.

Since (4.1) admits a unique strong solution, Proposition 3.11 implies the following result.

Proposition 4.2 *Under Assumption 4.1, solving (4.2) is equivalent to solving the original problem (2.5). That is,*

$$V_{\text{po}}(x, y) = V(x, y) \quad \text{for any given and fixed } (x, y) \in (0, \infty) \times (0, 1),$$

and a control is optimal for the separated problem (4.2) if and only if it is optimal for the original problem (2.5) under partial observation.

In the following analysis, we need (for technical reasons due to the infinite horizon of our problem) to take a sufficiently large discount factor. Namely, defining

$$\begin{aligned} \rho_o := & \left(\beta_2 + \frac{1}{2}\sigma^2 \right) \vee \left(\gamma\beta_2 + \frac{1}{2}\sigma^2\gamma(\gamma - 1) \right) \vee (2\beta_2 + \sigma^2) \vee (24\theta^2 - (\lambda_1 + \lambda_2)) \\ & \vee (4\beta_2 + 6\sigma^2) \vee \left(4\beta_2(2 \vee \gamma) + 2\sigma^2(2 \vee \gamma)(4(2 \vee \gamma) - 1) \right), \end{aligned}$$

with $\theta^2 := \frac{1}{2} \left(\frac{(g_1 - g_2)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2 \right)$, we assume the following.

Assumption 4.3 One has $\rho > \rho_o^+$.

Due to the growth condition on h , let us notice that Assumption 4.3 in particular ensures that $\rho > \gamma\beta_2 + \frac{1}{2}\sigma^2\gamma(\gamma - 1)$ so that the (trivial) admissible control $\nu \equiv 0$ has a finite total expected cost.

4.1 The related optimal stopping problem

Motivated by the results of the previous sections (in particular Theorem 3.13), we now aim at solving (4.2) through the study of an auxiliary optimal stopping problem whose value function can be interpreted as the marginal value of the optimal debt reduction problem (cf. (3.23) and Theorem 3.13). Therefore, we can think informally of the solution to that optimal stopping problem as the optimal time at which the government should marginally reduce the debt ratio. The optimal stopping problem involves a two-dimensional diffusive process, and in the sequel, we provide an almost exclusively probabilistic analysis.

4.1.1 Formulation and preliminary results

Recall that $(I_t, I_t^1)_{t \geq 0}$ is a two-dimensional standard \mathbb{H} -Brownian motion, and introduce the two-dimensional diffusion process $(\widehat{X}, \pi) := (\widehat{X}_t, \pi_t)_{t \geq 0}$ solving the stochastic differential equations (SDEs)

$$\begin{cases} d\widehat{X}_t = \widehat{X}_t(\beta_2 + (g_2 - g_1)\pi_t)dt + \sigma \widehat{X}_t dI_t, \\ d\pi_t = (\lambda_2 - (\lambda_1 + \lambda_2)\pi_t)dt + \pi_t(1 - \pi_t) \left(\frac{g_2 - g_1}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1 \right) \end{cases} \quad (4.3)$$

with initial conditions $\widehat{X}_0 = x, \pi_0 = y$ for any $(x, y) \in \mathcal{O} := (0, \infty) \times (0, 1)$. Recall that $\beta_2 = r - g_2$.

Since the process π is bounded, classical results on SDEs ensure that (4.3) admits a unique strong solution that, when needed, we denote by $(\widehat{X}^{x,y}, \pi^y)$ to stress its dependence on the initial datum $(x, y) \in \mathcal{O}$. In particular, one easily obtains

$$\widehat{X}_t^{x,y} = x e^{(\beta_2 - \frac{1}{2}\sigma^2)t + \sigma I_t + (g_2 - g_1) \int_0^t \pi_s^y ds}, \quad t \geq 0. \quad (4.4)$$

Moreover, it can be shown that Feller’s test of explosion (see e.g. Karatzas and Shreve [45, Chap. 5.5]) gives $1 = \mathbb{P}[\pi_t^y \in (0, 1), \forall t \geq 0]$ for all $y \in (0, 1)$. In fact, the boundary points 0 and 1 are classified as “entrance-not-exit”, hence are unattainable for the process π . In other words, the diffusion π can start from 0 and 1, but it cannot reach any of those two points when starting from $y \in (0, 1)$ (we refer to Borodin and Salminen [6, Sect. II.6] for further details on boundary classification).

With regard to Remark 3.14, we study here the fully two-dimensional Markovian optimal stopping problem with value function

$$\begin{aligned} v(x, y) &:= \inf_{\tau \geq 0} \mathbb{E}_{(x,y)} \left[\int_0^\tau e^{-\rho t} \widehat{X}_t h'(\widehat{X}_t) dt + e^{-\rho \tau} \widehat{X}_\tau \right] \\ &=: \inf_{\tau \geq 0} \widehat{\mathcal{J}}_{(x,y)}(\tau), \quad (x, y) \in \mathcal{O}. \end{aligned} \quad (4.5)$$

In (4.5), the optimisation is taken over all \mathbb{H} -stopping times, and $\mathbb{E}_{(x,y)}$ denotes the expectation under the probability measure $\mathbb{P}_{(x,y)}[\cdot] := \mathbb{P}[\cdot | \widehat{X}_0 = x, \pi_0 = y]$.

Because π is positive, $g_2 - g_1 < 0$ and $\rho > \beta_2$ by Assumption 4.3, (4.4) gives

$$\liminf_{t \uparrow \infty} e^{-\rho t} \widehat{X}_t = 0 \quad \mathbb{P}_{(x,y)\text{-a.s.}},$$

which implies via the convention (3.21) that $e^{-\rho \tau} \widehat{X}_\tau = 0$ on $\{\tau = \infty\}$ for any \mathbb{H} -stopping time τ .

Clearly, $v \geq 0$ since \widehat{X} is positive and h is increasing on \mathbb{R}_+ . Also, $v \leq x$ on \mathcal{O} , and we can therefore define the continuation region and the stopping region as

$$\mathcal{C} := \{(x, y) \in \mathcal{O} : v(x, y) < x\}, \quad \mathcal{S} := \{(x, y) \in \mathcal{O} : v(x, y) = x\}. \quad (4.6)$$

Notice that integrating by parts the term $e^{-\rho \tau} \widehat{X}_\tau$, taking expectations and exploiting that $\mathbb{E}[\int_0^\tau e^{-\rho s} \widehat{X}_s dI_s] = 0$ for any \mathbb{H} -stopping time τ (because $\rho > \beta_2 + \frac{1}{2}\sigma^2$ by

Assumption 4.3), we can equivalently rewrite (4.5) as

$$v(x, y) := x + \inf_{\tau \geq 0} \mathbb{E}_{(x,y)} \left[\int_0^\tau e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt \right] \tag{4.7}$$

for any $(x, y) \in \mathcal{O}$. From (4.7), it is readily seen that

$$\{(x, y) \in \mathcal{O} : h'(x) - (\rho - \beta_2 - (g_2 - g_1)y) < 0\} \subseteq \mathcal{C},$$

which implies

$$\mathcal{S} \subseteq \{(x, y) \in \mathcal{O} : h'(x) - (\rho - \beta_2 - (g_2 - g_1)y) \geq 0\}. \tag{4.8}$$

Moreover, since ρ satisfies Assumption 4.3 and $0 \leq \pi_t \leq 1$ for any $(x, y) \in \mathcal{O}$, one has that

$$\mathbb{E}_{(x,y)} \left[\int_0^\infty e^{-\rho t} \widehat{X}_t (h'(\widehat{X}_t) + \rho + |\beta_2| + |g_2 - g_1|) dt \right] < \infty, \tag{4.9}$$

and the family of random variables

$$\left\{ \int_0^\tau e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt : \tau \text{ is an } \mathbb{H}\text{-stopping time} \right\}$$

is therefore \mathbb{H} -uniformly integrable under $\mathbb{P}_{(x,y)}$.

Preliminary properties of v are given in the next proposition.

Proposition 4.4 *The following hold:*

- (i) $x \mapsto v(x, y)$ is increasing for any $y \in (0, 1)$.
- (ii) $y \mapsto v(x, y)$ is decreasing for any $x \in (0, \infty)$.
- (iii) $(x, y) \mapsto v(x, y)$ is continuous in \mathcal{O} .

Proof (i) Recall (4.5). By the strict convexity and monotonicity of h and (4.4), it follows that $x \mapsto \widehat{\mathcal{J}}_{(x,y)}(\tau)$ is increasing for any \mathbb{H} -stopping time τ and any $y \in (0, 1)$. Hence the claim is proved.

(ii) This is due to the fact that $y \mapsto \widehat{\mathcal{J}}_{(x,y)}(\tau)$ is decreasing for any stopping time τ and $x \in (0, \infty)$. Indeed, the mapping $y \mapsto \widehat{X}_t^{x,y}$ is a.s. decreasing for any $t \geq 0$ because $y \mapsto \pi_t^y$ is a.s. increasing by the comparison theorem of Yamada and Watanabe (see e.g. Karatzas and Shreve [45, Proposition 5.2.18]) and $g_2 - g_1 < 0$, and $x \mapsto xh'(x)$ is increasing.

(iii) Since $(x, y) \mapsto (\widehat{X}_t^{x,y}, \pi_t^y)$ is a.s. continuous for any $t \geq 0$, it is not hard to verify that $(x, y) \mapsto \widehat{\mathcal{J}}_{(x,y)}(\tau)$ is continuous for any given $\tau \geq 0$. Hence v is upper semicontinuous. We now show that it is also lower semicontinuous.

Let $(x, y) \in \mathcal{O}$ and let $(x_n, y_n)_{n \in \mathbb{N}} \subseteq \mathcal{O}$ be any sequence converging to (x, y) . Without loss of generality, we may take $(x_n, y_n) \in (x - \delta, x + \delta) \times (y - \delta, y + \delta)$ for a suitable $\delta > 0$. Letting $\tau_\varepsilon^n := \tau_\varepsilon^n(x_n, y_n)$ be ε -optimal for $v(x_n, y_n)$, but suboptimal for $v(x, y)$, we can then write

$$v(x, y) - v(x_n, y_n) \leq \mathbb{E} \left[\int_0^{\tau_\varepsilon^n} e^{-\rho t} (\widehat{X}_t^{x,y} h'(\widehat{X}_t^{x,y}) - \widehat{X}_t^{x_n, y_n} h'(\widehat{X}_t^{x_n, y_n})) dt \right] + \mathbb{E} [e^{-\rho \tau_\varepsilon^n} (\widehat{X}_{\tau_\varepsilon^n}^{x,y} - \widehat{X}_{\tau_\varepsilon^n}^{x_n, y_n})] + \varepsilon.$$

Notice now that a.s.

$$\int_0^{\tau_\varepsilon^n} e^{-\rho t} |\widehat{X}_t^{x,y} h'(\widehat{X}_t^{x,y}) - \widehat{X}_t^{x_n, y_n} h'(\widehat{X}_t^{x_n, y_n})| dt \leq \int_0^\infty e^{-\rho t} (\widehat{X}_t^{x,y} h'(\widehat{X}_t^{x,y}) + \widehat{X}_t^{x+\delta, y-\delta} h'(\widehat{X}_t^{x+\delta, y-\delta})) dt,$$

where we have used that $x \mapsto \widehat{X}^{x,y}$ is increasing, $y \mapsto \widehat{X}^{x,y}$ is decreasing and $x \mapsto xh'(x)$ is positive and increasing. The random variable on the right-hand side above is independent of n and integrable due to (4.9). Also, using integration by parts and performing standard estimates, we can write that a.s.

$$e^{-\rho \tau_\varepsilon^n} (\widehat{X}_{\tau_\varepsilon^n}^{x,y} - \widehat{X}_{\tau_\varepsilon^n}^{x_n, y_n}) \leq |x - x_n| + \int_0^\infty e^{-\rho s} (\rho + |\beta_2| + |g_2 - g_1|)(\widehat{X}_s^{x,y} + \widehat{X}_s^{x+\delta, y-\delta}) ds,$$

and the last integral above is independent of n and has finite expectation due to (4.9). Then taking limits as $n \uparrow \infty$, invoking the dominated convergence theorem thanks to the previous estimates and using that $(x, y) \mapsto (\widehat{X}_t^{x,y}, \pi_t^y)$ is a.s. continuous for any $t \geq 0$, we find (after rearranging terms) that

$$\liminf_{n \uparrow \infty} v(x_n, y_n) \geq v(x, y) - \varepsilon.$$

We thus conclude that v is lower semicontinuous at (x, y) by arbitrariness of ε . Since $(x, y) \in \mathcal{O}$ was arbitrary as well, v is lower semicontinuous on \mathcal{O} . \square

Due to Proposition 4.4 (iii), the stopping region is closed whereas the continuation region is open. Moreover, thanks to (4.9) and the $\widehat{\mathbb{P}}_{(x,y)}$ -a.s. continuity of the paths of the process $(\int_0^t e^{-\rho s} \widehat{X}_s (h'(\widehat{X}_s) - (\rho - \beta_2 - (g_2 - g_1))\pi_s) ds)_{t \geq 0}$, we can apply Karatzas and Shreve [46, Theorem D.12] to obtain that the first entry time of (\widehat{X}, π) into \mathcal{S} is optimal for (4.5), that is,

$$\tau^*(x, y) := \inf\{t \geq 0 : (\widehat{X}_t, \pi_t) \in \mathcal{S}\} \quad \mathbb{P}_{(x,y)\text{-a.s.}}, (x, y) \in \mathcal{O}, \tag{4.10}$$

attains the infimum in (4.5) (with the usual convention $\inf \emptyset = \infty$). Also, standard arguments based on the strong Markov property of (\widehat{X}, π) (see e.g. Peskir and Shiryaev [57, Theorem I.2.4]) allow one to show that $\mathbb{P}_{(x,y)$ -a.s., the process $S := (S_t)_{t \geq 0}$ with

$$S_t := e^{-\rho t} v(\widehat{X}_t, \pi_t) + \int_0^t e^{-\rho s} \widehat{X}_s h'(\widehat{X}_s) dt$$

is an \mathbb{H} -submartingale, and the stopped process $(S_{t \wedge \tau^*})_{t \geq 0}$ is an \mathbb{H} -martingale. The latter two conditions are usually referred to as the *subharmonic characterisation* of the value function v .

We now rule out the possibility of an empty stopping region.

Lemma 4.5 *The stopping region of (4.6) is not empty.*

Proof We argue by contradiction and suppose that $\mathcal{S} = \emptyset$. Hence for any $(x, y) \in \mathcal{O}$, we can write

$$\begin{aligned} x > v(x, y) &= \mathbb{E}_{(x,y)} \left[\int_0^\infty e^{-\rho t} \widehat{X}_t h'(\widehat{X}_t) dt \right] \\ &\geq K_o x^\gamma \mathbb{E}_{(1,y)} \left[\int_0^\infty e^{-\rho t} \widehat{X}_t^\gamma dt \right] - \frac{K}{\rho}, \end{aligned}$$

where the inequality $xh'(x) \geq h(x)$, due to convexity of h , and the growth condition assumed on h (cf. Assumption 4.1) have been used. Now by taking x sufficiently large, we reach a contradiction since $\gamma > 1$ by assumption. Hence $\mathcal{S} \neq \emptyset$. \square

Proposition 4.6 *For any $y \in (0, 1)$, let*

$$\bar{x}(y) := \inf\{x > 0 : v(x, y) \geq x\} \tag{4.11}$$

with the convention $\inf \emptyset = \infty$. Then:

(i) *We have*

$$\mathcal{C} = \{(x, y) \in \mathcal{O} : x < \bar{x}(y)\}, \quad \mathcal{S} = \{(x, y) \in \mathcal{O} : x \geq \bar{x}(y)\}. \tag{4.12}$$

(ii) $y \mapsto \bar{x}(y)$ *is increasing and left-continuous.*

(iii) *There exist $0 < x_\star < x^\star < \infty$ such that for any $y \in [0, 1]$,*

$$(h')^{-1}(\rho - \beta_2) \vee x_\star \leq \bar{x}(y) \leq x^\star.$$

Proof (i) To show (4.12), it suffices to show that if $(x_1, y) \in \mathcal{S}$, then $(x_2, y) \in \mathcal{S}$ for any $x_2 \geq x_1$. Let $\tau^\varepsilon := \tau^\varepsilon(x_2, y)$ be an ε -optimal stopping time for $v(x_2, y)$. Then exploiting $\widehat{X}_t^{x_2,y} = \frac{x_2}{x_1} \widehat{X}_t^{x_1,y} \geq \widehat{X}_t^{x_1,y}$ a.s. and monotonicity of h' , (4.7) yields

$$\begin{aligned} 0 &\geq v(x_2, y) - x_2 \\ &\geq \mathbb{E} \left[\int_0^{\tau^\varepsilon} e^{-\rho t} \widehat{X}_t^{x_2,y} \left(h'(\widehat{X}_t^{x_2,y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_t^y) \right) dt \right] - \varepsilon \\ &\geq \frac{x_2}{x_1} \mathbb{E} \left[\int_0^{\tau^\varepsilon} e^{-\rho t} \widehat{X}_t^{x_1,y} \left(h'(\widehat{X}_t^{x_1,y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_t^y) \right) dt \right] - \varepsilon \\ &\geq \frac{x_2}{x_1} (v(x_1, y) - x_1) - \varepsilon = -\varepsilon. \end{aligned}$$

Therefore, by arbitrariness of ε , we conclude that $(x_2, y) \in \mathcal{S}$ as well, and therefore that \bar{x} in (4.11) splits \mathcal{C} and \mathcal{S} as in (4.12).

(ii) Let $(x, y_1) \in \mathcal{C}$. Since $y \mapsto v(x, y)$ is decreasing by Proposition 4.4 (ii), it follows that $(x, y_2) \in \mathcal{C}$ for any $y_2 \geq y_1$. This in turn implies that $y \mapsto \bar{x}(y)$ is increasing.

The monotonicity of $y \mapsto \bar{x}(y)$ together with the fact that \mathcal{S} is closed then gives the claimed left-continuity by standard arguments.

(iii) Let $\Theta_t^x := x \exp((\beta_2 - \frac{1}{2}\sigma^2 + (g_2 - g_1))t + \sigma I_t)$ and introduce the one-dimensional optimal stopping problem

$$v^*(x) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \Theta_t^x h'(\Theta_t^x) dt + e^{-\rho \tau} \Theta_\tau^x \right], \quad x > 0.$$

Because $g_2 - g_1 < 0$, h' is increasing and $\pi_t^y \leq 1$ a.s. for all $t \geq 0$ and $y \in (0, 1)$, it is not hard to see that $v(x, y) \geq v^*(x)$ for any $(x, y) \in \mathcal{O}$.

By arguments similar to those employed to prove (i), one can show that there exists x^* such that $\{x \in (0, \infty) : v^*(x) \geq x\} = \{x \in (0, \infty) : x \geq x^*\}$. In fact, by arguing as in the proof of Lemma 4.5, the latter set is not empty. Then we have the inclusions

$$\{x \in (0, \infty) : x \geq x^*\} \subseteq \{(x, y) \in \mathcal{O} : v(x, y) \geq x\} = \{(x, y) \in \mathcal{O} : x \geq \bar{x}(y)\},$$

which in turn show that $\bar{x}(y) \leq x^*$ for all $y \in (0, 1)$. Hence also $\bar{x}(y) \leq x^*$ for all $y \in [0, 1]$, setting $\bar{x}(0+) := \lim_{y \downarrow 0} \bar{x}(y)$ by monotonicity and $\bar{x}(1) := \lim_{y \uparrow 1} \bar{x}(y)$ by left-continuity. As for the lower bound of \bar{x} , notice that (4.8) implies

$$\bar{x}(y) \geq (h')^{-1}(\rho - \beta_2 - (g_2 - g_1)y) =: \zeta(y), \quad y \in (0, 1), \tag{4.13}$$

where $(h')^{-1}(\cdot)$ is the inverse of the strictly increasing function $h' : [0, \infty) \rightarrow (0, \infty)$ (notice that $\rho - \beta_2 - (g_2 - g_1)y \geq 0$ since $\rho > \beta_2$, $g_2 - g_1 < 0$ and $y > 0$). Since $(h')^{-1}$ is strictly increasing and $-(g_2 - g_1)y \geq 0$, we can conclude from (4.13) that $\bar{x}(y) \geq (h')^{-1}(\rho - \beta_2)$ for every $y \in [0, 1]$. Moreover, setting

$$\Psi_t^x := x \left(\left(\beta_2 - \frac{1}{2}\sigma^2 \right) t + \sigma I_t \right)$$

and introducing the one-dimensional optimal stopping problem

$$v_*(x) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \Psi_t^x h'(\Psi_t^x) dt + e^{-\rho \tau} \Psi_\tau^x \right], \quad x > 0,$$

one has $v(x, y) \leq v_*(x)$ for any $(x, y) \in \mathcal{O}$. Following arguments as those employed above and defining $x_* := \inf\{x > 0 : v_*(x) \geq x\} \in (0, \infty)$, the last inequality implies that $\bar{x}(y) \geq x_*$ for all $y \in [0, 1]$. □

4.1.2 Smooth-fit property and continuity of the free boundary

We now aim at proving further regularity of v and the free boundary \bar{x} .

The second-order linear elliptic differential operator

$$\begin{aligned} \mathbb{L} := & (\beta_2 + (g_2 - g_1)y)x \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + (\lambda_2 - (\lambda_1 + \lambda_2)y) \frac{\partial}{\partial y} \\ & + \frac{1}{2} \left((\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2} \right) y^2 (1 - y)^2 \frac{\partial^2}{\partial y^2}, \end{aligned} \tag{4.14}$$

acting on any function $f \in C^2(\mathcal{O})$, is the infinitesimal generator of the process (\widehat{X}, π) . The nondegeneracy of the process (\widehat{X}, π) and the smoothness of the coefficients in (4.14) together with the subharmonic characterisation of v allow proving by standard arguments (see e.g. [57, Sect. 3.7.1]) and classical regularity results for elliptic partial differential equations (see e.g. Gilbarg and Trudinger [37, Sect. 6.6.3]) the following result.

Lemma 4.7 *The value function v of (4.5) belongs to C^2 separately in the interior of \mathcal{C} and in the interior of \mathcal{S} (i.e., away from the boundary $\partial\mathcal{C}$ of \mathcal{C}). Moreover, in the interior of \mathcal{C} , it satisfies*

$$(\mathbb{L} - \rho)v(x, y) = -xh'(x),$$

with \mathbb{L} as in (4.14).

We continue our analysis by proving that the value function of (4.5) belongs to the class $C^1((0, \infty) \times (0, 1))$. This will be obtained through probabilistic methods that rely on the regularity (in the sense of diffusions) of the stopping set \mathcal{S} for the process (\widehat{X}, π) (see De Angelis and Peskir [22] where this methodology has recently been developed in a general context; for other examples, refer to De Angelis et al. [21] as well as to Johnson and Peskir [43]). Recall that the boundary points are regular for \mathcal{S} with respect to (\widehat{X}, π) if (cf. Karatzas and Shreve [45, Definition 4.2.9])

$$\widehat{\tau}(x_o, y_o) := \inf\{t > 0 : (\widehat{X}_t^{x_o, y_o}, \pi_t^{y_o}) \in \mathcal{S}\} = 0 \quad \text{a.s.}, \quad \forall (x_o, y_o) \in \partial\mathcal{C}. \quad (4.15)$$

The time $\widehat{\tau}(x_o, y_o)$ is the first hitting time of $(\widehat{X}^{x_o, y_o}, \pi^{y_o})$ to \mathcal{S} .

Notice that defining $U_t := \ln \widehat{X}_t$, one has

$$dU_t = \left(\beta_2 + (g_2 - g_1)\pi_t - \frac{1}{2}\sigma^2 \right) dt + \sigma dI_t$$

as well as $\mathbb{E}_{(x,y)}[f(\widehat{X}_t, \pi_t)] = \mathbb{E}_{(u,y)}[f(e^{U_t}, \pi_t)]$ for every bounded Borel function $f : \mathbb{R}^2 \mapsto \mathbb{R}$, where $u := \ln x$. Due the nondegeneracy of the process (U, π) and the smoothness and boundedness of its coefficients, the pair (U, π) has a continuous transition density $\widehat{p}(\cdot, \cdot, \cdot; u, y)$, $(u, y) \in \mathbb{R} \times (0, 1)$, such that for any $(u', y') \in \mathbb{R} \times (0, 1)$ and $t \geq 0$ (see e.g. Aronson [1]),

$$\begin{aligned} \frac{M}{t} \exp\left(-\lambda \frac{(u - u')^2 + (y - y')^2}{t}\right) &\geq \widehat{p}(t, u', y'; u, y) \\ &\geq \frac{m}{t} \exp\left(-\Lambda \frac{(u - u')^2 + (y - y')^2}{t}\right), \end{aligned} \quad (4.16)$$

for some constants $M > m > 0$ and $\Lambda > \lambda > 0$. It thus follows that the mapping $(u, y) \mapsto \mathbb{E}_{(u,y)}[f(e^{U_t}, \pi_t)]$ is continuous, so that (U, π) is a strong Feller process. Hence (\widehat{X}, π) is strong Feller as well, and we can therefore conclude that (4.15) holds if and only if (see Dynkin [26, Chap. 13.1-2])

$$\tau^*(x_n, y_n) \rightarrow 0 \quad \text{a.s.} \quad \text{whenever } \mathcal{C} \ni (x_n, y_n)_{n \in \mathbb{N}} \rightarrow (x_o, y_o) \in \partial\mathcal{C},$$

where τ^* is as in (4.10).

The next proposition shows the validity of (4.15).

Proposition 4.8 *The boundary points in $\partial\mathcal{C}$ are regular for \mathcal{S} with respect to (\widehat{X}, π) , that is, (4.15) holds.*

Proof Let $(x_o, y_o) \in \partial\mathcal{C}$ and set $u_o := \ln x_o$. We set $\widehat{\sigma}(u_o, y_o) := \widehat{\tau}(e^{u_o}, y_o)$ for any given $(u_o, y_o) \in \mathbb{R} \times (0, 1)$ and equivalently rewrite (4.15) in terms of the process (U, π) , with U as defined above, as

$$\begin{aligned} \widehat{\sigma}(u_o, y_o) &= \inf\{t > 0 : U_t^{u_o, y_o} \geq \ln \bar{x}(\pi_t^{y_o})\} \\ &= 0 \quad \text{a.s., for all } (u_o, y_o) \text{ such that } u_o = \ln \bar{x}(y_o). \end{aligned}$$

Given that $y \mapsto \ln \bar{x}(y)$ is increasing like $y \mapsto \bar{x}(y)$, the region

$$\widehat{\mathcal{S}} := \{(u, y) \in \mathbb{R} \times (0, 1) : u \geq \ln \bar{x}(y)\}$$

enjoys the so-called *cone property* (see Karatzas and Shreve [45, Definition 4.2.18]). In particular, we can always construct a cone C_o with vertex in (u_o, y_o) and aperture $0 \leq \phi \leq \pi/2$ such that $C_o \cap (\mathbb{R} \times (0, 1)) \subseteq \widehat{\mathcal{S}}$ and for any $t_o \geq 0$, we have

$$\mathbb{P}[\widehat{\sigma}(u_o, y_o) \leq t_o] \geq \mathbb{P}[(U_{t_o}^{u_o, y_o}, \pi_{t_o}^{y_o}) \in C_o]. \tag{4.17}$$

Then using (4.16), one has

$$\begin{aligned} \mathbb{P}[(U_{t_o}^{u_o, y_o}, \pi_{t_o}^{y_o}) \in C_o] &= \int_{C_o} \widehat{p}(t_o, u_o, y_o; u, y) du dy \\ &\geq \int_{C_o} \frac{m}{t_o} e^{-\Lambda \frac{(u-u_o)^2 + (y-y_o)^2}{t_o}} du dy \\ &= m \int_{C_o} e^{-\Lambda((u')^2 + (y')^2)} du' dy' =: \ell > 0, \end{aligned} \tag{4.18}$$

using that the change of variables $u' := (u - u_o)/\sqrt{t_o}$ and $y' := (y - y_o)/\sqrt{t_o}$ maps the cone C_o into itself. The number ℓ above depends on u_o, y_o , but is independent of t_o . From (4.17) and (4.18), we thus have $\mathbb{P}[\widehat{\sigma}(u_o, y_o) \leq t_o] \geq \ell$, and letting $t_o \downarrow 0$ yields $\mathbb{P}[\widehat{\sigma}(u_o, y_o) = 0] \geq \ell > 0$. However, $\{\widehat{\sigma}(u_o, y_o) = 0\} \in \mathcal{H}_o$, and by the Blumenthal 0–1 law, we obtain $\mathbb{P}[\widehat{\sigma}(u_o, y_o) = 0] = 1$, which completes the proof. \square

Theorem 4.9 *One has that $v \in C^1(\mathcal{O})$.*

Proof The value function belongs to C^2 in the interior of the continuation region due to Lemma 4.7, and it is C^∞ in the interior of the stopping region where $v(x, y) = x$. It thus only remains to prove that v is continuously differentiable across $\partial\mathcal{C}$. In the sequel, we prove that (i) the function $\bar{w}(x, y) := \frac{1}{x}(v(x, y) - x)$ has a continuous derivative with respect to x across $\partial\mathcal{C}$ (and this clearly implies the continuity of v_x across $\partial\mathcal{C}$); (ii) the function v_y is continuous across $\partial\mathcal{C}$.

(i) *Continuity of v_x across $\partial\mathcal{C}$* : For the subsequent arguments, it is useful to notice that the function \bar{w} admits the representation (recall (4.7))

$$\bar{w}(x, y) = \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \widehat{X}_s^{1,y} \left(h'(\widehat{X}_s^{x,y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_s^y) \right) ds \right] \tag{4.19}$$

and to bear in mind that the optimal stopping time τ^* for v in (4.10) is also optimal for \bar{w} since $v \geq x$ if and only if $\bar{w} \geq 0$. We now prove that \bar{w}_x is continuous across $\partial\mathcal{C}$, thus implying continuity of v_x across $\partial\mathcal{C}$.

Take $(x, y) \in \mathcal{C}$ and let $\varepsilon > 0$ be such that $x - \varepsilon > 0$. Since $x \mapsto \bar{w}(x, y)$ is increasing due to the monotonicity of h' , it is clear that $(x - \varepsilon, y) \in \mathcal{C}$ as well. Denote by $\tau_\varepsilon^*(x, y) := \tau^*(x - \varepsilon, y)$ the optimal stopping time for $\bar{w}(x - \varepsilon, y)$ and notice that $\tau_\varepsilon^*(x, y)$ is suboptimal for $\bar{w}(x, y)$ and $\tau_\varepsilon^*(x, y) \rightarrow \tau^*(x, y)$ a.s. as $\varepsilon \downarrow 0$. To simplify the exposition, we write $\tau_\varepsilon^* := \tau_\varepsilon^*(x, y)$ and $\tau^* := \tau^*(x, y)$ in the sequel. We then have from (4.19) that

$$\begin{aligned} 0 \leq \frac{\bar{w}(x, y) - \bar{w}(x - \varepsilon, y)}{\varepsilon} &\leq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{1,y} (h'(\widehat{X}_t^{x,y}) - h'(\widehat{X}_t^{x-\varepsilon,y})) dt \right] \\ &= \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\widehat{X}_t^{1,y})^2 h''(\widehat{X}_t^{\xi_\varepsilon,y}) dt \right] \end{aligned} \tag{4.20}$$

for some $\xi_\varepsilon \in (x - \varepsilon, x)$, where we have used in the last step the mean value theorem and the fact that $\widehat{X}_t^{x,y} - \widehat{X}_t^{x-\varepsilon,y} = \varepsilon \widehat{X}_t^{1,y}$. Letting $\varepsilon \downarrow 0$, invoking dominated convergence (thanks to the fact that $\rho > (\gamma\beta_2 + \frac{1}{2}\sigma^2\gamma - 1) \vee (2\beta_2 + \sigma^2)$ by Assumption 4.3) and using that $\bar{w} \in C^1(\mathcal{C})$ (since $v \in C^1(\mathcal{C})$), we then find from (4.20) that

$$0 \leq \bar{w}_x(x, y) \leq \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} (\widehat{X}_t^{1,y})^2 h''(\widehat{X}_t^{x,y}) dt \right]. \tag{4.21}$$

Now let (x_o, y_o) be an arbitrary point belonging to $\partial\mathcal{C}$. Taking limits $(x, y) \rightarrow (x_o, y_o)$ in (4.21), using dominated convergence and Proposition 4.8, we obtain

$$0 \leq \liminf_{(x,y) \rightarrow (x_o,y_o) \in \partial\mathcal{C}} \bar{w}_x(x, y) \leq \limsup_{(x,y) \rightarrow (x_o,y_o) \in \partial\mathcal{C}} \bar{w}_x(x, y) \leq 0,$$

thus proving that \bar{w}_x is continuous across $\partial\mathcal{C}$. This immediately implies the continuity of v_x across $\partial\mathcal{C}$, upon recalling that $v(x, y) = x(\bar{w}(x, y) + 1)$.

(ii) *Continuity of v_y across $\partial\mathcal{C}$* : Take again $(x, y) \in \mathcal{C}$ and $\varepsilon > 0$ such that $y + \varepsilon < 1$. Since $y \mapsto v(x, y)$ is decreasing by Proposition 4.4 (ii), it is clear that $(x, y + \varepsilon) \in \mathcal{C}$ as well. Denote by $\tau_\varepsilon^*(x, y) := \tau^*(x, y + \varepsilon)$ the optimal stopping time for $v(x, y + \varepsilon)$ and notice that $\tau_\varepsilon^*(x, y)$ is suboptimal for $v(x, y)$ and $\tau_\varepsilon^*(x, y) \rightarrow \tau^*(x, y)$ a.s. as $\varepsilon \downarrow 0$. To simplify the notation, we write τ_ε^* instead of $\tau_\varepsilon^*(x, y)$ in the sequel. From Proposition 4.4 (ii) and (4.7), we then have

$$\begin{aligned}
 0 &\geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \\
 &\geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} \left(h'(\widehat{X}_t^{x, y + \varepsilon}) - (\rho - \beta_2 - \pi_t^{y + \varepsilon} (g_2 - g_1)) \right) dt \right] \\
 &\quad - \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y} \left(h'(\widehat{X}_t^{x, y}) - (\rho - \beta_2 - \pi_t^y (g_2 - g_1)) \right) dt \right] \\
 &= \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \left((\widehat{X}_t^{x, y + \varepsilon} h'(\widehat{X}_t^{x, y + \varepsilon}) - \widehat{X}_t^{x, y} h'(\widehat{X}_t^{x, y})) \right. \right. \\
 &\quad \left. \left. - (\rho - \beta_2)(\widehat{X}_t^{x, y + \varepsilon} - \widehat{X}_t^{x, y}) \right) dt \right] \\
 &\quad + \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (g_2 - g_1) (\widehat{X}_t^{x, y + \varepsilon} \pi_t^{y + \varepsilon} - \widehat{X}_t^{x, y} \pi_t^y) dt \right].
 \end{aligned}$$

Now add and subtract on the right-hand side both $\mathbb{E}[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} h'(\widehat{X}_t^{x, y}) dt]$ and $(g_2 - g_1)\mathbb{E}[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} \pi_t^y dt]$ and recall that $g_2 - g_1 < 0$, $\widehat{X}_t^{x, y} \geq 0$ a.s. and $\pi_t^{y + \varepsilon} - \pi_t^y \geq 0$ a.s., for every $t \geq 0$. Then after rearranging terms and using the integral mean value theorem for some $L_t^\varepsilon \in (\widehat{X}_t^{x, y + \varepsilon}, \widehat{X}_t^{x, y})$ a.s., we obtain that

$$\begin{aligned}
 0 &\geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \\
 &\geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} (h'(\widehat{X}_t^{x, y + \varepsilon}) - h'(\widehat{X}_t^{x, y})) dt \right] \\
 &\quad + \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\widehat{X}_t^{x, y + \varepsilon} - \widehat{X}_t^{x, y}) (h'(\widehat{X}_t^{x, y}) - (\rho - \beta_2 - \pi_t^y (g_2 - g_1))) dt \right] \\
 &\quad - \frac{1}{\varepsilon} |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} (\pi_t^{y + \varepsilon} - \pi_t^y) dt \right] \tag{4.22} \\
 &\geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\widehat{X}_t^{x, y + \varepsilon} - \widehat{X}_t^{x, y}) (\widehat{X}_t^{x, y + \varepsilon} h''(L_t^\varepsilon) + h'(\widehat{X}_t^{x, y})) dt \right] \\
 &\quad - \frac{1}{\varepsilon} |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x, y + \varepsilon} (\pi_t^{y + \varepsilon} - \pi_t^y) dt \right].
 \end{aligned}$$

In the last inequality, we have used that $\rho - \beta_2 - \pi_t^y (g_2 - g_1) \geq 0$ since $\rho > \beta_2$ by Assumption 4.3, that $g_2 - g_1 < 0$ and that $\widehat{X}_t^{x, y + \varepsilon} \leq \widehat{X}_t^{x, y}$.

Define now $\Delta\pi_t^y := \frac{1}{\varepsilon}(\pi_t^{y + \varepsilon} - \pi_t^y)$, $t \geq 0$, and notice that by using the second equation in (4.3), we can write for any $t \geq 0$ that

$$\begin{aligned}
 \Delta\pi_t^y &= 1 - \int_0^t (\lambda_1 + \lambda_2) \Delta\pi_s^y ds \\
 &\quad + \int_0^t \Delta\pi_s^y (1 - \pi_s^{y + \varepsilon} - \pi_s^y) \left(\frac{g_2 - g_1}{\sigma} dI_s + (\alpha_1 - \alpha_2) dI_s^1 \right).
 \end{aligned}$$

With the help of Itô’s formula, it can easily be shown that

$$\begin{aligned} \Delta\pi_t^y &= \exp\left(-(\lambda_1 + \lambda_2)t - \theta^2 \int_0^t (1 - \pi_s^{y+\varepsilon} - \pi_s^y)^2 ds\right) \\ &\quad \times \exp\left(\int_0^t (1 - \pi_s^{y+\varepsilon} - \pi_s^y) \left(\frac{g_2 - g_1}{\sigma} dI_s + (\alpha_1 - \alpha_2) dI_s^1\right)\right) \end{aligned} \tag{4.23}$$

with $\theta^2 := \frac{1}{2} \left(\frac{(g_2 - g_1)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2\right)$. Also, by (4.4) and simple algebra,

$$\frac{1}{\varepsilon} (\widehat{X}_t^{x,y+\varepsilon} - \widehat{X}_t^{x,y}) = \widehat{X}_t^{x,y} \frac{e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon}. \tag{4.24}$$

Using the definition of $\Delta\pi_t^y$ and (4.24) in (4.22) and $\widehat{X}_t^{x,y+\varepsilon} \leq \widehat{X}_t^{x,y}$, one finds

$$\begin{aligned} 0 &\geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \\ &\geq \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x,y} \frac{e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon} (\widehat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\widehat{X}_t^{x,y})) dt \right] \\ &\quad - |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x,y} \Delta\pi_t^y dt \right]. \end{aligned} \tag{4.25}$$

We now aim at taking limits as $\varepsilon \downarrow 0$ in (4.25). To this end, notice that $\Delta\pi_t^y \rightarrow Z_t^y$ a.s. for all $t \geq 0$ as $\varepsilon \downarrow 0$, where by Protter [58, Theorem V.7.39], $(Z_t^y)_{t \geq 0}$ is the unique strong solution to

$$dZ_t^y = -(\lambda_1 + \lambda_2)Z_t^y dt + Z_t^y \left(1 - 2\pi_t^y\right) \left(\frac{g_2 - g_1}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1\right), \quad t > 0,$$

with $Z_0^y = 1$. Then, if we are allowed to invoke the dominated convergence theorem when taking limits as $\varepsilon \downarrow 0$ in (4.25), we obtain that

$$\begin{aligned} 0 &\geq v_y(x, y) \\ &\geq (g_2 - g_1) \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} \widehat{X}_t^{x,y} \left(\int_0^t Z_s^y ds\right) (\widehat{X}_t^{x,y} h''(\widehat{X}_t^{x,y}) + h'(\widehat{X}_t^{x,y})) dt \right] \\ &\quad - |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} \widehat{X}_t^{x,y} Z_t^y dt \right] \end{aligned} \tag{4.26}$$

upon recalling that $v \in C^2(\mathcal{C})$. Therefore, letting (x_o, y_o) be an arbitrary point belonging to $\partial\mathcal{C}$, by taking limits in (4.26) as $(x, y) \rightarrow (x_o, y_o)$ and using dominated convergence and Proposition 4.8, we obtain that

$$0 \geq \limsup_{(x,y) \rightarrow (x_o,y_o) \in \partial\mathcal{C}} v_y(x, y) \geq \liminf_{(x,y) \rightarrow (x_o,y_o) \in \partial\mathcal{C}} v_y(x, y) \geq 0,$$

thus proving that v_y is continuous across $\partial\mathcal{C}$.

To complete the proof, it only remains to show that the dominated convergence theorem can be applied when taking limits as $\varepsilon \downarrow 0$ in (4.25). We show this in the two following technical steps.

1) To prove that the dominated convergence theorem can be invoked when taking $\varepsilon \downarrow 0$ in the first expectation on the right-hand side of (4.25), we set

$$\Lambda_\varepsilon := \int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x,y} \frac{e^{\varepsilon(g_2-g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon} (\widehat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\widehat{X}_t^{x,y})) dt$$

and show that the family of random variables $\{\Lambda_\varepsilon : \varepsilon \in (0, 1 - y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

Notice that by Assumption 4.1 (ii) and the fact that $L_t^\varepsilon \leq \widehat{X}_t^{x,y}$ a.s., one has a.s. for any $t \geq 0$ that

$$\widehat{X}_t^{x,y} (\widehat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\widehat{X}_t^{x,y})) \leq \widehat{K} (1 + (\widehat{X}_t^{x,y})^{\gamma\vee 2})$$

for some constant $\widehat{K} > 0$ independent of ε so that by Jensen’s inequality,

$$|\Lambda_\varepsilon|^2 \leq \frac{2\widehat{K}^2}{\rho^2} \int_0^\infty \rho e^{-\rho t} \left(\frac{1 - e^{\varepsilon(g_2-g_1) \int_0^t \Delta\pi_s^y ds}}{\varepsilon} \right)^2 (1 + (\widehat{X}_t^{x,y})^{2\gamma\vee 4}) dt.$$

Taking expectations and applying Hölder’s inequality gives

$$\begin{aligned} \mathbb{E}[|\Lambda_\varepsilon|^2]^{\frac{1}{2}} &\leq K' \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\frac{1 - e^{\varepsilon(g_2-g_1) \int_0^t \Delta\pi_s^y ds}}{\varepsilon} \right)^4 dt \right]^{\frac{1}{4}} \\ &\quad \times \mathbb{E} \left[\int_0^\infty e^{-\rho t} (1 + (\widehat{X}_t^{x,y})^{4\gamma\vee 8}) dt \right]^{\frac{1}{4}} \end{aligned} \tag{4.27}$$

for some other constant $K' > 0$ independent of ε that in the sequel will vary from line to line. The standard inequality $1 - e^{-u} \leq u$ with $u = \varepsilon(g_1 - g_2) \int_0^t \Delta\pi_s^y ds \geq 0$ allows us to continue from (4.27) and write

$$\begin{aligned} \mathbb{E}[|\Lambda_\varepsilon|^2]^{\frac{1}{2}} &\leq K' \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\int_0^t \Delta\pi_s^y ds \right)^4 dt \right]^{\frac{1}{4}} \\ &\quad \times \mathbb{E} \left[\int_0^\infty e^{-\rho t} (1 + (\widehat{X}_t^{x,y})^{4(\gamma\vee 2)}) dt \right]^{\frac{1}{4}}. \end{aligned} \tag{4.28}$$

We now treat the two expectations in (4.28) separately. First, by Jensen’s inequality,

$$\left(\int_0^t \Delta\pi_s^y ds \right)^4 = \left(\frac{1}{t} \int_0^t t \Delta\pi_s^y ds \right)^4 \leq t^3 \int_0^t (\Delta\pi_s^y)^4 ds. \tag{4.29}$$

Second, thanks to the nonnegativity of $(\Delta\pi^y)^4$, we can invoke the Fubini–Tonelli theorem and using also (4.29), we obtain

$$\begin{aligned} & \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(\int_0^t \Delta\pi_s^y ds\right)^4 dt\right] \\ & \leq \mathbb{E}\left[\int_0^\infty e^{-\rho t} t^3 \int_0^t (\Delta\pi_s^y)^4 ds dt\right] \\ & = \frac{1}{\rho^4} \int_0^\infty e^{-\rho s} (\rho^3 s^3 + 3\rho^2 s^2 + 6\rho s + 6)\mathbb{E}[(\Delta\pi_s^y)^4] ds. \end{aligned} \tag{4.30}$$

To evaluate the expectation in the last integral above, notice that applying Itô’s formula to the process $\xi_t^y := (\Delta\pi_t^y)^4$ and using (4.23) gives for any $t > 0$ that

$$\begin{aligned} d\xi_t^y &= \xi_t^y(-(\lambda_1 + \lambda_2) + 12\theta^2(1 - \pi_t^{y+\varepsilon} - \pi_t^y)^2) dt \\ & \quad + 4\xi_t^y(1 - \pi_t^{y+\varepsilon} - \pi_t^y)\left(\frac{g_2 - g_1}{\sigma} dI_s + (\alpha_1 - \alpha_2)dI_s^1\right) \end{aligned}$$

with $\xi_0^y = 1$ and $\theta^2 = \frac{1}{2}\left(\frac{(g_2 - g_1)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2\right)$. Because $(1 - \pi_t^{y+\varepsilon} - \pi_t^y)^2 \leq 2$ a.s. for all $t \geq 0$ and

$$\xi_t^y = e^{-(\lambda_1 + \lambda_2)t + 12\theta^2 \int_0^t (1 - \pi_s^{y+\varepsilon} - \pi_s^y)^2 ds} M_t^y,$$

where $(M_t^y)_{t \geq 0}$ is an exponential martingale, it is easy to see that

$$\mathbb{E}[(\Delta\pi_t^y)^4] \leq e^{-(\lambda_1 + \lambda_2)t + 24\theta^2 t}, \quad t \geq 0. \tag{4.31}$$

Using the latter estimate in (4.30) together with Assumption 4.3, we deduce that

$$\sup_{\varepsilon \in (0, 1-y)} \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(\int_0^t \Delta\pi_s^y ds\right)^4 dt\right] < \infty. \tag{4.32}$$

For the second expectation in (4.28), Assumption 4.3 and standard estimates using (4.4) plus the fact that $(g_2 - g_1) \int_0^t \pi_s^y ds < 0$ guarantee that it is finite. Moreover, it is independent of ε . Combining this with (4.32), we thus find from (4.28) that

$$\sup_{\varepsilon \in (0, 1-y)} \mathbb{E}[|\Lambda_\varepsilon|^2]^{\frac{1}{2}} < \infty.$$

This implies that the family of random variables $\{\Lambda_\varepsilon : \varepsilon \in (0, 1 - y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

2) Consider the second expectation on the right-hand side of (4.25) and set

$$\Xi_\varepsilon := \int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x,y} \Delta\pi_t^y dt.$$

We aim at proving that the family of random variables $\{\Xi_\varepsilon : \varepsilon \in (0, 1 - y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

By Jensen’s inequality and then Hölder’s inequality, one finds that

$$\mathbb{E}[|\Xi_\varepsilon|^2]^{\frac{1}{2}} \leq \widehat{K} \mathbb{E}\left[\int_0^\infty e^{-\rho t} (\widehat{X}_t^{x,y})^4 dt\right]^{\frac{1}{4}} \mathbb{E}\left[\int_0^\infty e^{-\rho t} (\Delta\pi_t^y)^4 dt\right]^{\frac{1}{4}} \tag{4.33}$$

for some $\widehat{K} > 0$ independent of ε . The first expectation on the right-hand side of (4.33) is finite thanks to Assumption 4.3 and standard estimates using (4.4) plus the fact that $(g_2 - g_1) \int_0^t \pi_s^y ds < 0$. Moreover, it is independent of ε . For the second one, interchanging expectation and dt -integral by the Fubini–Tonelli theorem and using (4.31), we obtain

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} (\Delta \pi_t^y)^4 dt \right]^{\frac{1}{4}} \leq \frac{1}{(\rho + \lambda_1 + \lambda_2 - 24\theta^2)^{\frac{1}{4}}}$$

by Assumption 4.3. We thus conclude by (4.33) that $\sup_{\varepsilon \in (0, 1-y)} \mathbb{E}[|\Xi_\varepsilon|^2]^{\frac{1}{2}} < \infty$, which completes the proof. □

The previous theorem in particular implies the so-called *smooth-fit* property, a well-known optimality principle in optimal stopping theory. Moreover, by standard arguments based on the strong Markov property of (\widehat{X}, π) (see Peskir and Shiryaev [57, Chap. III]), it follows from the results collected so far that the couple (v, \bar{x}) solves the free-boundary problem

$$\begin{cases} (\mathbb{L} - \rho)v(x, y) = -xh'(x) & \text{on } \mathcal{C}, \\ v(x, y) = x & \text{on } \mathcal{S}, \\ v_x(x, y) = 1 & \text{at } x = \bar{x}(y), y \in (0, 1), \\ v_y(x, y) = 0 & \text{at } x = \bar{x}(y), y \in (0, 1), \end{cases}$$

with $v \in C^2(\mathcal{C})$.

An important consequence of Theorem 4.9 is the following result.

Proposition 4.10 *The mapping $y \mapsto \bar{x}(y)$ is continuous on $[0, 1]$.*

Proof Let $T > 0$ and define the probability measure \mathbb{Q} on (Ω, \mathcal{H}_T) by

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{H}_t} = e^{-\frac{1}{2}\sigma^2 t + \sigma I_t}, \quad t \in [0, T].$$

Under the new measure \mathbb{Q} , the process $\widehat{I}_t := I_t - \sigma t, t \in [0, T]$, is a standard Brownian motion, and the dynamics of (\widehat{X}, π) read

$$\begin{cases} d\widehat{X}_t = \widehat{X}_t(\beta_2 + \sigma^2 + (g_2 - g_1)\pi_t)dt + \sigma \widehat{X}_t d\widehat{I}_t, \\ d\pi_t = (\lambda_2 - (\lambda_1 + \lambda_2)\pi_t + (g_2 - g_1)\pi_t(1 - \pi_t))dt \\ \quad + \pi_t(1 - \pi_t) \left(\frac{g_2 - g_1}{\sigma} d\widehat{I}_t + (\alpha_1 - \alpha_2) dI_t^1 \right), \end{cases} \tag{4.34}$$

with initial conditions $\widehat{X}_0 = x, \pi_0 = y, (x, y) \in \mathcal{O}$. Now for any τ and (x, y) , we have

$$\begin{aligned} & \mathbb{E}_{(x,y)} \left[\int_0^{\tau \wedge T} e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt \right] \\ &= \mathbb{E}_{(x,y)}^{\mathbb{Q}} \left[\int_0^{\tau \wedge T} e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \pi_s ds} \widehat{H}(\widehat{X}_t, \pi_t) dt \right] \end{aligned} \tag{4.35}$$

with $\widehat{H}(x, y) := (\rho - \beta_2 - (g_2 - g_1)y - h'(x))$. We cannot directly take the limit $T \uparrow \infty$ in (4.35) since the measure change \mathbb{Q} depends on T . However, we notice that the right-hand side of (4.35) only depends on the law of $(\bar{X}, \bar{\pi})$ under \mathbb{Q} . Therefore, we can define a new probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ equipped with a two-dimensional Brownian motion (\bar{B}^1, \bar{B}^2) and a filtration $\bar{\mathbb{F}} := (\bar{\mathcal{F}}_t)_{t \geq 0}$ and let $(\bar{X}, \bar{\pi})$ be the unique strong solution to (4.34), driven by (\bar{B}^1, \bar{B}^2) . In this setting, we can then define the stopping problems

$$\begin{aligned} \bar{V}(x, y; T) &:= \sup_{\tau \geq 0} \bar{\mathbb{E}}_{(x,y)} \left[\int_0^{\tau \wedge T} e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \bar{\pi}_s ds} \widehat{H}(\bar{X}_t, \bar{\pi}_t) dt \right], \\ \bar{V}(x, y) &:= \sup_{\tau \geq 0} \bar{\mathbb{E}}_{(x,y)} \left[\int_0^{\tau} e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \bar{\pi}_s ds} \widehat{H}(\bar{X}_t, \bar{\pi}_t) dt \right], \end{aligned} \tag{4.36}$$

where $\bar{\mathbb{E}}_{(x,y)}$ is the expectation under $\bar{\mathbb{P}}$ conditionally on $(\bar{X}_0, \bar{\pi}_0) = (x, y)$. By arguing as in the proof of De Angelis [19, Proposition 4.2], one can show that

$$\lim_{T \uparrow \infty} \bar{V}(x, y; T) = \bar{V}(x, y) \quad \text{and} \quad \lim_{T \uparrow \infty} \widehat{V}(x, y; T) = \widehat{V}(x, y),$$

where we have set

$$\begin{aligned} \widehat{V}(x, y; T) &:= \sup_{\tau \geq 0} \mathbb{E}_{(x,y)} \left[\int_0^{\tau \wedge T} e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt \right], \\ \widehat{V}(x, y) &:= \sup_{\tau \geq 0} \mathbb{E}_{(x,y)} \left[\int_0^{\tau} e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt \right]. \end{aligned}$$

Since now

$$\begin{aligned} &\sup_{\tau \geq 0} \mathbb{E}_{(x,y)}^{\mathbb{Q}} \left[\int_0^{\tau \wedge T} e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \pi_s ds} \widehat{H}(\widehat{X}_t, \pi_t) dt \right] \\ &= \sup_{\tau \geq 0} \bar{\mathbb{E}}_{(x,y)} \left[\int_0^{\tau \wedge T} e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \bar{\pi}_s ds} \widehat{H}(\bar{X}_t, \bar{\pi}_t) dt \right], \end{aligned}$$

the equivalence in law of the processes $(\widehat{X}, \pi, \widehat{I}, I^1)$ under \mathbb{Q} and $(\bar{X}, \bar{\pi}, \bar{B}^1, \bar{B}^2)$ under $\bar{\mathbb{P}}$ on $[0, T]$ plus (4.35) allow us to write

$$\bar{V}(x, y) = \lim_{T \uparrow \infty} \bar{V}(x, y; T) = \lim_{T \uparrow \infty} \widehat{V}(x, y; T) = \widehat{V}(x, y).$$

In light of this last equality, it is then not difficult to see that v as in (4.7) is such that $v(x, y) := x - \bar{V}(x, y)$ for any $(x, y) \in \mathcal{O}$. Since

$$\{(x, y) \in \mathcal{O} : v(x, y) \geq x\} = \{(x, y) \in \mathcal{O} : \bar{V}(x, y) \leq 0\},$$

$\bar{x}(\cdot)$ is the optimal stopping boundary for the problem with value \bar{V} as well.

To prove the continuity of $\bar{x}(\cdot)$, we now aim at applying Peskir [56, Theorem 10] for (4.36). Notice that $\bar{V}_x \leq 0$ on \mathcal{O} since $x \mapsto h(x)$ is strictly convex. Moreover,

recalling $\theta^2 = \frac{1}{2}((\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2})$, we have $\partial_x \frac{\widehat{H}}{\theta^2 y^2 (1-y)^2} < 0$ on \mathcal{O} again thanks to the strict convexity of h . Also, \overline{V}_y is continuous across the boundary due to the C^1 -property shown in Theorem 4.9 for $v = x - \widehat{V}$; hence, the horizontal smooth-fit property holds. We can therefore apply [56, Theorem 10] (upon noticing that in [56], x is the horizontal axis and y the vertical one, while in our paper, x is the vertical axis and y the horizontal one) and conclude that \overline{x} cannot have discontinuities of the first kind at any point $y \in [0, 1)$. Finally, \overline{x} is also continuous at $y = 1$ since it is left-continuous by Proposition 4.6 (ii). \square

4.2 The optimal control for the problem (4.2)

In this section, we provide the form of the optimal debt reduction policy. It is given in terms of the free boundary studied in the previous section.

For \overline{x} as in (4.11), introduce under $\mathbb{P}_{(x,y)}$ the nondecreasing process

$$\overline{v}_t^* = \left(x - \inf_{0 \leq s \leq t} (\overline{x}(\pi_s) e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - \sigma I_s - (g_2 - g_1) \int_0^s \pi_u du}) \right) \vee 0, \quad t \geq 0, \quad (4.37)$$

with $\overline{v}_{0-}^* = 0$ and then the process

$$v_t^* := \int_0^t e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - \sigma I_s - (g_2 - g_1) \int_0^s \pi_u du} d\overline{v}_s^*, \quad t \geq 0, v_{0-}^* = 0. \quad (4.38)$$

Notice that since $\overline{v}_t^* \leq x$ a.s. for all $t \geq 0$ and $t \mapsto \overline{v}_t^*$ is nondecreasing, it follows from (4.38) that v^* is admissible. Moreover, $t \mapsto \overline{v}_t^*$ is continuous (with the exception of a possible initial jump at time 0), due to the continuity of $y \mapsto \overline{x}(y)$, $t \mapsto I_t$, $t \mapsto \pi_t$ and $t \mapsto \int_0^t \pi_s ds$.

Theorem 4.11 *Let $\widetilde{V}(x, y) := \int_0^x \frac{1}{z} v(z, y) dz$, $(x, y) \in [0, \infty) \times [0, 1]$. Then one has $\widetilde{V} = V$ on $[0, \infty) \times [0, 1]$, and v^* as in (4.38) is optimal for (4.2).*

Proof Recall $U = U_0$ as in (3.20) and notice that in our Markovian setting, one actually has $\frac{1}{z} v(z, y) = U(z)$. By the proof of Theorem 3.13, it suffices to show that the right-continuous inverse of the stopping time $\tau^*(z, y) = \inf\{t \geq 0 : \widehat{X}_t^{z,y} \geq \overline{x}(\pi_t^y)\}$ (which is optimal for $v(z, y)$, cf. (4.10)) coincides (up to a null set) with \overline{v}^* . For that, recall (3.25) from the proof of Theorem 3.13, fix $(x, y) \in (0, \infty) \times (0, 1)$, take $t \geq 0$ arbitrary and notice that by (4.10), we have $\mathbb{P}_{(z,y)}$ -a.s. the equivalences

$$\begin{aligned} \tau^*(z, y) \leq t &\iff \widehat{X}_\theta \geq \overline{x}(\pi_\theta) \text{ for some } \theta \in [0, t] \\ &\iff z \geq e^{-(\beta_2 - \frac{1}{2}\sigma^2)\theta - \sigma I_\theta - (g_2 - g_1) \int_0^\theta \pi_u du} \overline{x}(\pi_\theta) \text{ for some } \theta \in [0, t] \\ &\iff \left(x - \inf_{0 \leq s \leq t} (\overline{x}(\pi_s) e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - (g_2 - g_1) \int_0^s \pi_u du - \sigma I_s}) \right) \vee 0 \geq x - z \\ &\iff \overline{v}_t^* \geq x - z \\ &\iff \tau_+^{\overline{v}^*}(z) \leq t. \end{aligned}$$

Hence $\tau_+^{\bar{v}^*}(z) = \tau^*(z, y)$ a.s. and \bar{v}^* is the right-continuous inverse of $\tau^*(\cdot, y)$. As \bar{v}^* is admissible, the claim follows by arguing as in the proof of Theorem 3.13, part 2). \square

Notice that (4.38) and the equation for $X^{x,y,v}$ in the formulation of (4.2) yield

$$X_t^{x,y,v^*} = e^{(\beta_2 - \frac{1}{2}\sigma^2)t + (g_2 - g_1) \int_0^t \pi_s^y ds + \sigma I_t}(x - \bar{v}_t^*),$$

which with regard to (4.37) shows that

$$0 \leq X_t^{x,y,v^*} \leq \bar{x}(\pi_t^y), \quad t \geq 0, \mathbb{P}\text{-a.s.}$$

Moreover, it is easy to see that we can express \bar{v}^* of (4.37) as

$$\bar{v}_t^* = \sup_{0 \leq s \leq t} \left(\frac{X_s^{x,y,0} - \bar{x}(\pi_s^y)}{X_s^{1,y,0}} \right) \vee 0, \quad \bar{v}_{0-}^* = 0.$$

These equations allow us to make some remarks about the optimal debt management policy of our problem.

(i) If at the initial time 0, the level x of the debt ratio is above $\bar{x}(y)$, then an immediate lump sum reduction of $x - \bar{x}(y)$ is optimal.

(ii) At any time $t \geq 0$, it is optimal to keep the debt ratio level below the belief-dependent ceiling \bar{x} .

(iii) If the level of the debt ratio at time t is strictly below $\bar{x}(\pi_t)$, there is no need for interventions. The government should intervene to reduce its debt only at those (random) times t at which the debt ratio attempts to rise above $\bar{x}(\pi_t)$. These interventions are then minimal, in the sense that $(X^{x,y,v^*}, \pi^y, v^*)$ solves a Skorokhod reflection problem at the free boundary \bar{x} .

(iv) Recall that the debt ceiling \bar{x} is an increasing function of the government’s belief that the economy is enjoying a phase of fast growth. Then, with regard to the previous description of the optimal debt reduction rule, we have that the more the government believes that the economy is in good shape, the less strict the optimal debt reduction policy should be.

4.3 Regularity of the value function of (4.2) and related HJB equation

Combining the results collected so far, we are now able to prove that the value function V of the control problem (4.2) is a twice continuously differentiable function. As a byproduct, V is a classical solution to the corresponding Hamilton–Jacobi–Bellman (HJB) equation.

From Theorem 4.11, we know that we have $V(x, y) = \int_0^x \frac{1}{z} v(z, y) dz$ for all $(x, y) \in \bar{\mathcal{O}} := [0, \infty) \times [0, 1]$. Hence thanks to Theorem 4.9 and the dominated convergence theorem, we immediately obtain the following result.

Lemma 4.12 *One has that $V \in C^1(\mathcal{O}) \cap C(\bar{\mathcal{O}})$. Moreover, $V_{xx} \in C(\mathcal{O})$ as well as $V_{xy} \in C(\mathcal{O})$.*

To take care of the second derivative V_{yy} we follow ideas used in De Angelis [19]. In particular, we determine the second weak derivative of V (recall that V_y is continuous by Theorem 4.9) and then show that it is a continuous function. This is accomplished in the next proposition.

Proposition 4.13 *Let $\theta^2 := \frac{1}{2}((\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2})$. We have $V_{yy} \in C(\mathcal{O})$ with*

$$\begin{aligned}
 &V_{yy}(x, y) \\
 &= -\frac{1}{\theta^2 y^2 (1 - y)^2} \left(\left(\beta_2 + (g_2 - g_1)y - \frac{1}{2}\sigma^2 \right) \left(v(x \wedge \bar{x}(y), y) - v(0+, y) \right) \right. \\
 &\quad \left. + h(x \wedge \bar{x}(y)) + \frac{1}{2}\sigma^2 (x \wedge \bar{x}(y)) v_x(x \wedge \bar{x}(y), y) \right) \\
 &\quad + \frac{\lambda_2 - (\lambda_1 + \lambda_2)y}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v_y(z, y) dz \\
 &\quad - \frac{\rho}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v(z, y) dz. \tag{4.39}
 \end{aligned}$$

Proof Notice that $V_y(x, y) = \int_0^x \frac{1}{z} v_y(z, y) dz$ and therefore $V_y(x, \cdot)$ is a continuous function for all $x > 0$ by Theorem 4.9 (notice indeed that by the bounds in (4.26) and the multiplicative dependence of $\widehat{X}^{z,y}$ with respect to z , $\frac{1}{z} v_y(z, y)$ is integrable at zero). Hence its weak derivative with respect to y is a function $g \in L^1_{loc}(\mathcal{O})$ such that for any test function $\varphi \in C^\infty_c((0, 1))$, one has

$$\int_0^1 V_y(x, y) \varphi'(y) dy = - \int_0^1 g(x, y) \varphi(y) dy.$$

We now want to evaluate g and show that it coincides with the right-hand side of (4.39).

Denote by $m(x)$ for $x > 0$ the generalised right-continuous inverse of $\bar{x}(y)$ for $y \in [0, 1]$, that is, $m(x) := \inf\{y \in [0, 1] : \bar{x}(y) \geq x\}$. Then noticing that $v_y = 0$ on the set $\{(x, y) \in \mathcal{O} : x > \bar{x}(y)\}$ and using Fubini's theorem, we can write

$$\begin{aligned}
 \int_0^1 V_y(x, y) \varphi'(y) dy &= \int_0^1 \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v_y(z, y) dz \varphi'(y) dy \\
 &= \int_0^x \frac{1}{z} \int_{m(z)}^1 v_y(z, y) \varphi'(y) dy dz \\
 &= \int_0^x \frac{1}{z} \left(v_y(z, 1) \varphi(1) - v_y(z, m(z)) \varphi(m(z)) \right. \\
 &\quad \left. - \int_{m(z)}^1 v_{yy}(z, y) \varphi(y) dy \right) dz \\
 &= - \int_0^x \frac{1}{z} \int_{m(z)}^1 v_{yy}(z, y) \varphi(y) dy dz, \tag{4.40}
 \end{aligned}$$

where we have used $v_y(z, m(z)) = 0$ for all $z \in (0, x)$ and $x > 0$ as well as $\varphi(1) = 0$. By Lemma 4.7 (cf. also (4.14)), for any $y > m(z)$ with $z \in (0, x)$ and $x > 0$, we have

$$v_{yy}(z, y) = \frac{1}{\theta^2 y^2 (1 - y)^2} \left(\rho v(z, y) - (\lambda_2 - (\lambda_1 + \lambda_2)y)v_y(z, y) - zh'(z) - \frac{1}{2}\sigma^2 z^2 v_{xx}(z, y) - (\beta_2 + (g_2 - g_1)y)z v_x(z, y) \right).$$

Inserting this into the last integral term on the right-hand side of (4.40), using again Fubini’s theorem and then integrating the derivatives with respect to x , we find

$$\begin{aligned} & \int_0^1 V_y(x, y)\varphi'(y)dy \\ &= - \int_0^x \frac{1}{z} \int_{m(z)}^1 v_{yy}(z, y)\varphi(y)dydz \\ &= \int_0^1 \frac{\lambda_2 - (\lambda_1 + \lambda_2)y}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v_y(z, y) dz \varphi(y) dy \\ & \quad - \int_0^1 \frac{\rho}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v(z, y) dz \varphi(y) dy \\ & \quad + \int_0^1 \left(h(x \wedge \bar{x}(y)) + (\beta_2 + (g_2 - g_1)y) \left(v(x \wedge \bar{x}(y), y) - v(0+, y) \right) \right. \\ & \quad \quad \left. + \frac{1}{2}\sigma^2 (x \wedge \bar{x}(y)) v_x(x \wedge \bar{x}(y), y) - \frac{1}{2}\sigma^2 \left(v(x \wedge \bar{x}(y), y) - v(0+, y) \right) \right) \frac{\varphi(y)}{\theta^2 y^2 (1 - y)^2} dy, \end{aligned} \tag{4.41}$$

where we have also used that $h(0) = 0$. Finally, setting

$$\begin{aligned} g(x, y) &:= - \frac{1}{\theta^2 y^2 (1 - y)^2} \left(h(x \wedge \bar{x}(y)) \right. \\ & \quad \left. + \left(\beta_2 + (g_2 - g_1)y - \frac{1}{2}\sigma^2 \right) \left(v(x \wedge \bar{x}(y), y) - v(0+, y) \right) \right. \\ & \quad \left. + \frac{1}{2}\sigma^2 (x \wedge \bar{x}(y)) v_x(x \wedge \bar{x}(y), y) \right) \\ & \quad + \frac{\lambda_2 - (\lambda_1 + \lambda_2)y}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v_y(z, y) dz \\ & \quad - \frac{\rho}{\theta^2 y^2 (1 - y)^2} \int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v(z, y) dz, \end{aligned}$$

we see that (4.41) reads $\int_0^1 V_y(x, y)\varphi'(y)dy = -\int_0^1 g(x, y)\varphi(y)dy$ so that g can be identified with the second weak derivative of V with respect to y . Notice that g is continuous by the continuity of \bar{x} , v , v_x , h and the fact that $\int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v(z, y)dz$ and $\int_0^{x \wedge \bar{x}(y)} \frac{1}{z} v_y(z, y)dz$ are finite due to (4.4), (4.5) and (4.26). The proof is complete. \square

Thanks to Lemma 4.12 and Proposition 4.13, we have that $V \in C^2(\mathcal{O}) \cap C(\overline{\mathcal{O}})$. As a byproduct of this, by the dynamic programming principle and standard methods based on an application of Dynkin's formula, we obtain the next result.

Proposition 4.14 *Recall the second-order differential operator \mathbb{L} defined in (4.14). The value function V of (4.2) is a classical solution to the HJB equation*

$$\min\{(\mathbb{L} - \rho)V(x, y) + h(x), 1 - V_x(x, y)\} = 0, \quad (x, y) \in \mathcal{O},$$

with boundary condition $V(0, y) = 0$ for any $y \in [0, 1]$.

Acknowledgements The research of Claudia Ceci is partially supported by ‘‘Gruppo Nazionale per l’Analisi Matematica, la Probabilit  e le loro Applicazioni’’ (GNAMPA) of ‘‘Istituto Nazionale di Alta Matematica’’ (INdAM). Financial support by the German Research Foundation (DFG) through the Collaborative Research Centre 1283 is gratefully acknowledged by Giorgio Ferrari. This work was also supported by the University of Padova via the research grant BIRD172407/17 ‘‘New Perspectives in Stochastic Methods for Finance and Energy Markets’’.

We wish to thank Luciano Campi, Tiziano De Angelis, Paola Mannucci, Fabio Paronetto, Paavo Salminen and Wolfgang Runggaldier for useful discussions.

We are also grateful to anonymous referees, the Associate Editor, the Editor Prof. Pierre Collin-Dufresne and the Editor in Chief Prof. Martin Schweizer for comments and suggestions that helped improve an earlier version of this work.

Funding Note Open Access funding provided by Projekt DEAL.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Filtering results

Proof of Proposition 3.5 Since the innovation processes (I, I^1) from (3.2) and the random measure $m(dt, dq)$ from (3.4) are \mathbb{H} -adapted, we have $\mathbb{F}^I \vee \mathbb{F}^{I^1} \vee \mathbb{F}^m \subseteq \mathbb{H}$.

In general, the inclusion could be strict. Now consider the exponential \mathbb{F} -martingale with $L_0 = 1$ solving

$$dL_t = -L_t \left(\frac{\beta(Z_t)}{\sigma} dW_t + \alpha(\eta_t, Z_t) dB_t \right), \quad t \geq 0.$$

Let $T > 0$ and define the probability measure \mathbb{Q} on (Ω, \mathcal{F}_T) by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = L_t, \quad t \in [0, T].$$

Notice that Assumption 3.1 ensures that L is indeed an \mathbb{F} -martingale. By Girsanov’s theorem, the processes

$$\tilde{W}_t := W_t + \int_0^t \frac{\beta(Z_s)}{\sigma} ds, \quad \tilde{B}_t := B_t + \int_0^t \alpha(\eta_s, Z_s) ds, \quad t \geq 0, \quad (\text{A.1})$$

are independent (\mathbb{Q}, \mathbb{F}) -Brownian motions on $[0, T]$. We now prove that we have $\mathbb{F}^{\tilde{W}} \vee \mathbb{F}^{\tilde{B}} \vee \mathbb{F}^m = \mathbb{H}$. The inclusion $\mathbb{F}^{\tilde{W}} \vee \mathbb{F}^{\tilde{B}} \vee \mathbb{F}^m \subseteq \mathbb{H}$ follows from the fact that \tilde{W} and \tilde{B} turn out to be \mathbb{H} -adapted since they can be written as

$$\tilde{W}_t = I_t + \int_0^t \frac{\pi_s(\beta)}{\sigma} ds, \quad \tilde{B}_t = I_t^1 + \int_0^t \pi_s(\alpha(\eta_s, \cdot)) ds, \quad t \geq 0. \quad (\text{A.2})$$

To prove the converse, observe that under \mathbb{Q} , the processes X^0 and η solve on $[0, T]$ the stochastic differential equations

$$\begin{aligned} dX_t^0 &= X_t^0 \sigma d\tilde{W}_t, & X_0^0 &= x > 0, \\ d\eta_t &= \sigma_1(\eta_t) d\tilde{W}_t + \sigma_2(\eta_t) d\tilde{B}_t + \int_{\mathbb{R}} qm(dt, dq), & \eta_0 &= q \in \mathcal{I}. \end{aligned} \quad (\text{A.3})$$

Clearly, X^0 is $\mathbb{F}^{\tilde{W}}$ -adapted. Recalling (3.13), the solution to (A.3) can be constructed iteratively. More precisely, for $t \in [0, T_1)$, the process η solves

$$d\eta_t = \sigma_1(\eta_t) d\tilde{W}_t + \sigma_2(\eta_t) d\tilde{B}_t, \quad \eta_0 = q \in \mathcal{I},$$

and between two consecutive jump times, i.e., for $t \in [T_n, T_{n+1})$, $n \geq 1$, one has

$$d\eta_t = \sigma_1(\eta_t) d\tilde{W}_t + \sigma_2(\eta_t) d\tilde{B}_t, \quad \eta_{T_n} = \eta_{T_n-} + \zeta_n.$$

By Assumption 2.1, this sequence of stochastic differential equations has a unique strong solution on any interval $[T_n, T_{n+1})$, and this in turn gives the unique strong solution η to (A.3). Moreover, η turns out to be $F^{\tilde{W}} \vee \mathbb{F}^{\tilde{B}} \vee \mathbb{F}^m$ -adapted.

By Jacod and Shiryaev [42, Corollary III.4.3.1], every (\mathbb{Q}, \mathbb{H}) -local martingale \tilde{M} admits the decomposition

$$\tilde{M}_t = \tilde{M}_0 + \int_0^t \tilde{\varphi}_s d\tilde{W}_s + \int_0^t \tilde{\psi}_s d\tilde{B}_s + \int_0^t \int_{\mathbb{R}} \tilde{w}(s, q) m^\pi(ds, dq), \quad t \in [0, T],$$

where $\tilde{\varphi}$ and $\tilde{\psi}$ are \mathbb{H} -predictable processes and \tilde{w} is an \mathbb{H} -predictable process indexed by \mathbb{R} such that for all $t \geq 0$,

$$\int_0^t \tilde{\varphi}_s^2 ds < \infty, \quad \int_0^t \tilde{\psi}_s^2 ds < \infty, \quad \int_0^t \int_{\mathbb{R}} |\tilde{w}(s, q)| m^{p, \mathbb{H}}(ds, dq) < \infty, \quad \mathbb{Q}\text{-a.s.}$$

Now let M be a (\mathbb{P}, \mathbb{H}) -local martingale. Then $\tilde{M} := M\tilde{L}^{-1}$ is a (\mathbb{Q}, \mathbb{H}) -local martingale, where

$$\tilde{L}_t := \mathbb{E}[L_t | \mathcal{H}_t] = \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{H}_t}, \quad t \in [0, T].$$

Taking into account (A.2), we have that \tilde{L} solves

$$d\tilde{L}_t = -\tilde{L}_t \left(\frac{\pi_t(\beta)}{\sigma} dI_t + \pi_t(\alpha(\eta_t, \cdot)) dI_t^1 \right), \quad \tilde{L}_0 = 1,$$

and by applying the product formula to $M = \tilde{M}\tilde{L}$, we easily obtain that under \mathbb{P} for any $t \in [0, T]$,

$$\begin{aligned} dM_t &= \tilde{M}_t d\tilde{L}_t + \tilde{L}_t d\tilde{M}_t + d\langle \tilde{M}^c, \tilde{L}^c \rangle_t \\ &= \left(\tilde{L}_t \tilde{\varphi}_t - M_t \frac{\pi_t(\beta)}{\sigma} \right) dI_t + \left(\tilde{L}_t \tilde{\psi}_t - M_t \pi_t(\alpha(\eta_t, \cdot)) \right) dI_t^1 \\ &\quad + \int_{\mathbb{R}} \tilde{w}(t, q) \tilde{L}_t m^\pi(dt, dq). \end{aligned}$$

To conclude, we thus only need to set

$$\varphi_t := \tilde{L}_t \tilde{\varphi}_t - M_t \frac{\pi_t(\beta)}{\sigma}, \quad \psi_t := \tilde{L}_t \tilde{\psi}_t - M_t \pi_t(\alpha(\eta_t, \cdot)), \quad w(t, q) := \tilde{w}(t, q) \tilde{L}_t,$$

and invoke the arbitrariness of $T > 0$. □

Proof of Theorem 3.6 In order to derive the filtering equation solved by the process $(\underline{\pi}_t)_{t \geq 0} = (\pi_t(i); i \in S)_{t \geq 0}$, we apply the innovation approach (see for instance Brémaud [7, Chap. IV]). We write \widehat{R} for the \mathbb{H} -optional projection of an \mathbb{F} -progressively measurable process R such that $\mathbb{E}[|R_t|] < \infty, \forall t \geq 0$. We recall that if $(\mathbb{E}[R_t | \mathcal{H}_t])_{t \geq 0}$ admits an optional version, then $\widehat{R}_t = \mathbb{E}[R_t | \mathcal{H}_t], t \geq 0$. In this proof, we use two well-known facts:

- (i) For every \mathbb{F} -martingale m , $\widehat{m}_t = \mathbb{E}[m_t | \mathcal{H}_t], t \geq 0$, and \widehat{m} is an \mathbb{H} -martingale.
- (ii) For any \mathbb{F} -progressively measurable and integrable process Ψ , we have that

$$\widehat{\int_0^t \Psi_s ds} - \int_0^t \widehat{\Psi}_s ds$$

is an \mathbb{H} -martingale.

The first step of the innovation method consists in writing the process $\mathbb{1}_{\{Z_t=i\}}$, $i \in S, t \geq 0$, as a semimartingale. Denoting by L^Z the Markov generator of the state process Z , we have that

$$L^Z f_i(j) = \sum_{k \in S} \lambda_{ki} f_k(j), \quad i, j \in S,$$

where $f_k(j) := \mathbb{1}_{\{j=k\}}$. Hence for any $i \in S$, we can write

$$\mathbb{1}_{\{Z_t=i\}} = f_i(Z_t) = f_i(Z_0) + \int_0^t L^Z f_i(Z_s) ds + m_t(i),$$

where $(m_t(i))_{t \geq 0}$ is an \mathbb{H} -martingale. By taking the \mathbb{H} -optional projection and using (i) and (ii) above, we obtain that

$$\pi_t(i) = y_i + \int_0^t \sum_{k \in S} \lambda_{ki} \pi_s(k) ds + M_t(i), \tag{A.4}$$

where $M_t(i)$ is an \mathbb{H} -martingale null at zero. Proposition 3.5 ensures the existence of processes $\psi(i)$ and $\varphi(i)$ that are \mathbb{H} -predictable and w_i which is \mathbb{H} -predictable and indexed by \mathbb{R} such that

$$M_t(i) = \int_0^t \psi_s(i) dI_s + \int_0^t \varphi_s(i) dI_s^1 + \int_0^t \int_{\mathbb{R}} w_i(s, q) m^\pi(ds, dq). \tag{A.5}$$

To obtain (3.11), it only remains to prove that

$$\begin{aligned} \psi_s(i) &= \pi_s(i) \sigma^{-1} \left(\beta(i) - \sum_{j=1}^Q \beta(j) \pi_s(j) \right), \\ \varphi_s(i) &= \pi_s(i) \left(\alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right), \\ w_i(s, q) &= w_i^\pi(s, q) - \pi_{s-}(i) \end{aligned}$$

with w_i^π given in (3.12). Following the proof of Ceci and Colaneri [15, Theorem 3.1], we can derive the structure of the processes $\psi(i), \varphi(i)$ by exploiting the equalities

$$\widehat{f_i(Z) \widetilde{W}} = \pi(i) \widetilde{W}, \quad \widehat{f_i(Z) \widetilde{B}} = \pi(i) \widetilde{B}, \quad \forall i \in S,$$

where \widetilde{W} and \widetilde{B} are the \mathbb{H} -adapted processes defined in (A.1). To derive the expression of w_i , we consider a bounded process Γ of the form $\Gamma_t = \int_0^t \int_{\mathbb{R}} \gamma(s, q) m(ds, dq)$ with an \mathbb{H} -predictable process γ indexed by \mathbb{R} . Since Γ is \mathbb{H} -adapted, the equality

$$\widehat{f_i(Z) \Gamma} = \pi(i) \Gamma, \quad \forall i \in S, \tag{A.6}$$

holds. By applying the product rule and taking into account that there are no common jumps between Z and N , we obtain

$$d(f_i(Z_t)\Gamma_t) = f_i(Z_{t-})d\Gamma_t + \Gamma_{t-}df_i(Z_t) \\ = \Gamma_t L^Z f_i(Z_t)dt + \int_{\mathbb{R}} f_i(Z_{t-})\gamma(t, q)m^{p, \mathbb{F}}(dt, dq) + d\mathcal{M}_t^{\mathbb{F}},$$

where $m^{p, \mathbb{F}}(dt, dq)$ is given in (3.9) and $\mathcal{M}^{\mathbb{F}}$ is an \mathbb{F} -martingale. By taking the optional projection onto \mathbb{H} and denoting by $\mathcal{M}^{\mathbb{H}}$ a local \mathbb{H} -martingale, we have that

$$d(\widehat{f_i(Z_t)\Gamma_t}) = \Gamma_t \pi_t(L^Z f_i)dt \\ + \lambda^N(i)\pi_{t-}(i)\gamma(t, c(\eta_{t-}, i))\mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}}dt + d\mathcal{M}_t^{\mathbb{H}}. \tag{A.7}$$

On the other hand, the product rule and (A.4) and (A.5) yield

$$d(\pi_t(i)\Gamma_t) = \pi_{t-}(i)d\Gamma_t + \Gamma_{t-}d\pi_t(i) + d\langle \pi(i), \Gamma \rangle_t \\ = \Gamma_t \pi_t(L^Z f_i)dt + \Gamma_{t-}dM_t(i) + \int_{\mathbb{R}} \gamma(t, q)w_i(t, q)m(dt, dq).$$

Recalling that $m^{p, \mathbb{H}}(dt, dq)$ is given in (3.8), we find

$$d(\pi_t(i)\Gamma_t) = \Gamma_t \pi_t(L^Z f_i)dt + \int_{\mathbb{R}} \gamma(t, q)w_i(t, q)m^{p, \mathbb{H}}(dt, dq) + d\mathcal{M}_t^{\mathbb{H}}, \tag{A.8}$$

where again $\mathcal{M}^{\mathbb{H}}$ is a local \mathbb{H} -martingale. Gathering (A.6)–(A.8), we obtain that for a.e. $t \geq 0$,

$$\lambda^N(i)\pi_{t-}(i)\gamma(t, c(\eta_{t-}, i))\mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} \\ = \sum_{j=1}^Q \pi_{t-}(j)\lambda^N(j)\gamma(t, c(\eta_{t-}, j))\mathbb{1}_{\{c(\eta_{t-}, j) \neq 0\}}\left(\pi_{t-}(i) + w_i(t, c(\eta_{t-}, j))\right).$$

Choose now γ of the form $\gamma(t, q) = C_t \mathbb{1}_A(q) \mathbb{1}_{\{t \leq T_n\}}$, with any bounded \mathbb{H} -predictable positive process C and $A \in \mathcal{B}(\mathbb{R})$. Observe that $|\Gamma_t| \leq \int_0^{t \wedge T_n} C_s dN_s \leq Dn$ with a positive constant D so that Γ is bounded. Then we have on $\{t \leq T_n\}$ the equality

$$\int_A v_t(i, dq) = \int_A (\pi_{t-}(i) + w_i(t, q))v_t(dq), \quad \forall A \in \mathcal{B}(\mathbb{R}),$$

where we have set

$$v_t(i, dq) := \lambda^N(i)\pi_{t-}(i)\mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}}\delta_{c(\eta_{t-}, i)}(dq), \quad v_t(dq) := \sum_{i=1}^Q v_t(i, dq).$$

Thus on $\{t \leq T_n\}$,

$$w_i^\pi(t, q) = w_i(t, q) - \pi_{t-}(i) = \frac{dv_t(i, dq)}{v_t(dq)}, \quad \forall i \in S.$$

Finally, since the counting process N is nonexplosive (as it has bounded intensity and therefore finite expectation; see [7, Sect. 1]), $T_n \uparrow \infty$ a.s. for $n \uparrow \infty$, and this yields (3.12). \square

Proof of Proposition 3.10 By Proposition 3.7, (3.11) and (3.17) are equivalent to a system of recursive equations between consecutive jump times, i.e., for $t \in [T_n, T_{n+1})$, $n = 0, 1, \dots$,

$$\begin{aligned} \pi_t(i) &= \pi_{T_n}(i) + \int_{T_n}^t b^\pi(\underline{\pi}_s, \eta_s, i) ds + \int_{T_n}^t \sigma_1^\pi(\underline{\pi}_s, i) dI_s \\ &\quad + \int_{T_n}^t \sigma_2^\pi(\underline{\pi}_s, i) dI_s^1, \quad i \in S, \\ \eta_t &= \eta_{T_n} + \int_{T_n}^t b^\eta(\underline{\pi}_s, \eta_s) ds + \int_{T_n}^t \sigma_1(\eta_s) dI_s + \int_{T_n}^t \sigma_2(\eta_s) dI_s^1, \end{aligned}$$

where we have set

$$\begin{aligned} b^\pi(\underline{y}, q, i) &:= \sum_{j=1}^Q \lambda_{ji} y_j - y_i \left(\lambda^N(i) \mathbb{1}_{\{c(q,i) \neq 0\}} - \sum_{j=1}^Q \lambda^N(j) y_j \mathbb{1}_{\{c(q,j) \neq 0\}} \right), \\ b^\eta(\underline{y}, q) &:= \sum_{j=1}^Q y_j b_1(q, j), \\ \sigma_1^\pi(\underline{y}, i) &:= \sigma^{-1} y_i \left(\beta(i) - \sum_{j=1}^Q \beta(j) y_j \right), \\ \sigma_2^\pi(\underline{y}, i) &:= y_i \left(\alpha(q, i) - \sum_{j=1}^Q \alpha(q, j) y_j \right), \end{aligned}$$

with the update at time T_n given by

$$\pi_{T_n}(i) = \frac{\lambda^N(i) \pi_{T_{n-1}}(i) \mathbb{1}_{\{\zeta_n = c(\eta_{T_{n-1}}, i)\}}}{\sum_{j=1}^Q \lambda^N(j) \pi_{T_{n-1}}(j) \mathbb{1}_{\{\zeta_n = c(\eta_{T_{n-1}}, j)\}}}, \quad i \in S, \quad \eta_{T_n} = \eta_{T_{n-1}} + \zeta_n. \quad (\text{A.9})$$

Recall that by assumption, the function $\alpha(q, i)$ in (3.3) is locally Lipschitz with respect to q and satisfies a (global) sublinear growth condition with respect to $q \in \mathcal{I}$, uniformly in $i \in S$. We develop the proof of uniqueness first by considering two special cases related to the jump amplitude function c , namely $c \neq 0$ and $c \equiv 0$. In both cases, the proof relies on classical results.

If $c \neq 0$, we have

$$b^\pi(\underline{y}, q, i) = \sum_{j=1}^Q \lambda_{ji} y_j - y_i \left(\lambda^N(i) - \sum_{j=1}^Q \lambda^N(j) y_j \right),$$

and it is easy to verify that between two consecutive jump times, the pair $(\underline{\pi}, \eta)$ solves a $(Q + 1)$ -dimensional stochastic differential equation with coefficients satisfying local Lipschitz and (global) sublinear growth conditions with respect to $(\underline{y}, q) \in \mathcal{Y} \times \mathbb{R}$, uniformly in $i \in S$. As a consequence, strong uniqueness holds between two consecutive jump times, i.e., for $t \in [T_n, T_{n+1})$, $n = 0, 1, \dots$. Moreover, since the update at the jump time T_n (see (A.9)) depends on the process $(\underline{\pi}_t, \eta_t)$ for $t \in [T_{n-1}, T_n)$, we have strong uniqueness of the solution to the system (3.11) and (3.17) for all $t \geq 0$.

If $c \equiv 0$, (3.11) and (3.17) reduce to

$$\begin{aligned} d\pi_t(i) &= b^\pi(\underline{\pi}_t, \eta_t, i)dt + \sigma_1^\pi(\underline{\pi}_t, i)dI_t + \sigma_2^\pi(\underline{\pi}_t, i)dI_t^1, & i \in S, t \geq 0, \\ d\eta_t &= b^\eta(\underline{\pi}_t, \eta_t)dt + \sigma_1(\eta_t)dI_t + \sigma_2(\eta_t)dI_t^1, & t \geq 0, \end{aligned}$$

where in particular $b^\pi(\underline{y}, q, i) = \sum_{j=1}^Q \lambda_{ji} y_j$. It is easy to check that also in this case, strong uniqueness follows by the local Lipschitz-property of the coefficients and by their (global) sublinear growth condition.

In the general case where c is \mathbb{R} -valued, the jump amplitude can assume any possible real value. In particular, c can be such that η and N have not only common jumps: N might jump at a time at which $c(\eta_{t-}, Z_{t-}) = 0$, so that η does not jump at that time. The treatment of this case is more delicate and must be performed separately. Indeed, uniqueness cannot be proved by using the arguments employed in the previous two cases because of the presence of $\mathbb{1}_{\{c(q,i) \neq 0\}}$ in the coefficient b^π which makes it impossible to prove the Lipschitz-continuity of b^π with respect to q . However, one can prove uniqueness by relying on the filtered martingale problem associated to the infinitesimal generator of the triplet (Z, X^0, η) . We refer to the seminal paper by Kurtz and Ocone [48], where the notion of filtered martingale problem has been introduced and applied to prove uniqueness of the filtering equation in the case of Gaussian additive white noise, and to the more recent work by Ceci and Colaneri [15, Theorem 3.3 and Appendix B], where uniqueness for a general jump-diffusion state-observations dynamics has been considered. We can prove analogously to [15, proof of Proposition B.1] that the boundedness of the jump intensity of η and Assumption 2.1 imply uniqueness for the filtered martingale problem associated to the infinitesimal generator of the triplet (Z, X^0, η) , which in turn implies strong uniqueness of the filtering equation by applying [15, Theorem 3.3]. \square

References

1. Aronson, D.G.: Bounds for the fundamental solution of a parabolic equation. *Bull. Am. Math. Soc.* **73**, 890–896 (1967)
2. Baldursson, F.M., Karatzas, I.: Irreversible investment and industry equilibrium. *Finance Stoch.* **1**, 69–89 (1997)
3. Bäuerle, N., Rieder, U.: Portfolio optimization with jumps and unobservable intensity process. *Math. Finance* **17**, 205–224 (2007)
4. Bensoussan, A.: *Stochastic Control of Partially Observable Systems*, 1st edn. Cambridge University Press, Cambridge (1992)
5. Blanchard, O.: *Macroeconomics*, 5th edn. Prentice Hall, New York (2009)
6. Borodin, A.N., Salminen, P.: *Handbook of Brownian Motion – Facts and Formulae*, 2nd edn. Birkhäuser, Basel (2015)

7. Brémaud, P.: Point Processes and Queues: Martingale Dynamics. Springer, New York (1980)
8. Cadenillas, A., Huamán-Aguilar, R.: Explicit formula for the optimal government debt ceiling. *Ann. Oper. Res.* **247**, 415–449 (2016)
9. Cadenillas, A., Huamán-Aguilar, R.: On the failure to reach the optimal government debt ceiling. *Risks* **6**(4), 1–28 (2018)
10. Callegaro, G., Gaigi, M., Scotti, S., Sgarra, C.: Optimal investment in markets with over and under-reaction to information. *Math. Financ. Econ.* **11**, 299–322 (2017)
11. Ceci, C.: Utility maximization with intermediate consumption under restricted information for jump market models. *Int. J. Theor. Appl. Finance* **15**, 1250040-1–34 (2012)
12. Ceci, C., Gerardi, A.: Partially observed control of a Markov jump process with counting observations: equivalence with the separated problem. *Stoch. Process. Appl.* **78**, 245–260 (1998)
13. Ceci, C., Gerardi, A.: Filtering of a Markov jump process with counting observations. *Appl. Math. Optim.* **42**, 1–18 (2000)
14. Ceci, C., Gerardi, A.: A model for high frequency data under partial information: a filtering approach. *Int. J. Theor. Appl. Finance* **9**, 555–576 (2006)
15. Ceci, C., Colaneri, K.: Nonlinear filtering for jump diffusion observations. *Adv. Appl. Probab.* **44**, 678–701 (2012)
16. Ceci, C., Colaneri, K.: The Zakai equation of nonlinear filtering for jump-diffusion observation: existence and uniqueness. *Appl. Math. Optim.* **69**, 47–82 (2014)
17. Christensen, S., Croce, F., Mordecki, E., Salminen, P.: On optimal stopping of multidimensional diffusions. *Stoch. Process. Appl.* **129**, 2561–2581 (2018)
18. Colaneri, K., Eksi, Z., Frey, R., Szölgényi, M.: Optimal liquidation under partial information with price impact. *Stoch. Process. Appl.* **130**, 1913–1946 (2020)
19. De Angelis, T.: Optimal dividends with partial information and stopping of a degenerate reflecting diffusion. *Finance Stoch.* **24**, 71–123 (2020)
20. De Angelis, T., Federico, S., Ferrari, G.: Optimal boundary surface for irreversible investment with stochastic costs. *Math. Oper. Res.* **42**, 1135–1161 (2017)
21. De Angelis, T., Gensbittel, S., Villeneuve, S.: A Dynkin game on assets with incomplete information on the return. *Math. Oper. Res.* (2020), forthcoming. Available online at [arXiv:1705.07352](https://arxiv.org/abs/1705.07352)
22. De Angelis, T., Peskir, G.: Global C^1 regularity of the value function in optimal stopping problems. *Ann. Appl. Probab.* **30**, 1007–1031 (2020)
23. Décamps, J.P., Villeneuve, S.: Integrating profitability prospects and cash management. TSE Working Paper, n. 15-570 (2015). Available online at https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2015/wp_tse_570.pdf
24. Dellacherie, C., Meyer, P.A.: Probability and Potential B. Theory of Martingales. North-Holland Mathematics Studies, vol. 72 (1982)
25. Detemple, J.B.: Asset pricing in a production economy with incomplete information. *J. Finance* **41**, 383–391 (1986)
26. Dynkin, E.B.: Markov Processes, Vol. II. Springer, Berlin (1965)
27. Eksi, Z., Ku, H.: Portfolio optimization for a large investor under partial information and price impact. *Math. Methods Oper. Res.* **86**, 601–623 (2017)
28. El Karoui, N., Huu Nguyen, D., Jeanblanc-Picqué, M.: Existence of an optimal Markovian filter for control under partial observations. *SIAM J. Control Optim.* **26**, 1025–1061 (1988)
29. Elliott, R.J., Aggoun, L., Moore, J.B.: Hidden Markov Models: Estimation and Control. Springer, New York (1995)
30. European Central Bank: Growth effects of high government debt. Monthly Bulletin March 2013, 82–84 (2013)
31. Ferrari, G.: On the optimal management of public debt: a singular stochastic control problem. *SIAM J. Control Optim.* **56**, 2036–2073 (2018)
32. Ferrari, G., Rodosthenous, N.: Optimal management of debt-to-GDP ratio with regime-switching interest rate. *SIAM J. Control Optim.* **58**, 755–786 (2020)
33. Fleming, W., Pardoux, E.: Optimal control for partially observed diffusions. *SIAM J. Control Optim.* **20**, 261–285 (1982)
34. Frey, R., Runggaldier, W.: A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *Int. J. Theor. Appl. Finance* **4**, 99–210 (2001)
35. Frey, R., Schmidt, T.: Pricing and hedging of credit derivatives via the innovation approach to nonlinear filtering. *Finance Stoch.* **16**, 105–133 (2012)
36. Gennotte, G.: Optimal portfolio choice under incomplete information. *J. Finance* **41**, 733–746 (1986)

37. Gilbarg, D., Trudinger, N.S.: *Elliptic Partial Differential Equations of the Second Order*, 2nd edn. Springer, Berlin (2001)
38. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, 357–384 (1989)
39. Holmström, B., Tirole, J.: Private and public supply of liquidity. *J. Polit. Econ.* **106**, 1–40 (1998)
40. Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, 1st edn. North Holland Publishing Company, Amsterdam (1981)
41. Jacod, J.: *Calcul Stochastique et Problèmes de Martingales*. Lecture Notes in Mathematics, vol. 714. Springer, Berlin (1979)
42. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin (2003)
43. Johnson, P., Peskir, G.: Quickest detection problems for Bessel processes. *Ann. Appl. Probab.* **27**, 1003–1056 (2017)
44. Kallianpur, G.: *Stochastic Filtering Theory*. Springer, New York (1980)
45. Karatzas, I., Shreve, S.E.: *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer, New York (1991)
46. Karatzas, I., Shreve, S.E.: *Methods of Mathematical Finance*. Springer, New York (1998)
47. Kliemann, W.H., Koch, G., Marchetti, F.: On the unnormalized solution of the filtering problem with counting process observations. *IEEE Trans. Inf. Theory* **36**, 1415–1425 (1990)
48. Kurtz, T.G., Ocone, D.: Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* **16**, 80–107 (1988)
49. Liptser, R., Shiryaev, A.N.: *Statistics of Random Processes. I. General Theory*, 2nd edn. Springer, Berlin (2001)
50. Lukkezen, J., Suyker, W.: Prudent debt level: a tentative calculation. CPB background document (2013). Available online at <https://www.cpb.nl/sites/default/files/publicaties/download/cpb-background-document-prudent-debt-level-tentative-calculation.pdf>
51. Luo, Y.: Robustly strategic consumption portfolio rules with informational frictions. *Manag. Sci.* **63**, 4158–4174 (2017)
52. Mazliak, L.: Mixed control problem under partial observation. *Appl. Math. Optim.* **27**, 57–84 (1993)
53. Menaldi, J.L., Robin, M.: On optimal correction problems with partial information. *Stoch. Anal. Appl.* **3**, 63–92 (1983)
54. Øksendal, B., Sulem, A.: Singular stochastic control and optimal stopping with partial-information of Itô–Lévy processes. *SIAM J. Control Optim.* **50**, 2254–2287 (2012)
55. Ostry, J.D., Ghosh, A.R., Espinoza, R.: When should public debt be reduced? IMF Staff Discussion Note SDN/15/10 (2015). Available online at <https://www.imf.org/external/pubs/ft/sdn/2015/sdn1510.pdf>
56. Peskir, G.: Continuity of the optimal stopping boundary for two-dimensional diffusions. *Ann. Appl. Probab.* **29**, 505–530 (2019)
57. Peskir, G., Shiryaev, A.N.: *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics ETH. Birkhäuser, Basel (2006)
58. Protter, P.: *Stochastic Integration and Differential Equations*, 2nd edn. Springer, Berlin (2005)
59. Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin (1999)
60. Stock, J.H., Watson, M.W.: How did leading indicator forecasts perform during the 2001 recession? *Econ. Q.- Fed. Reserve Bank Richmond* **89**(3), 71–90 (2003)
61. Tabellini, G.: Money, debt and deficits in a dynamic game. *J. Econ. Dyn. Control* **10**, 427–442 (1986)
62. Veronesi, P.: Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Rev. Financ. Stud.* **12**, 975–1007 (1999)
63. Woo, J., Kumar, M.S.: Public debt and growth. *Economica* **82**, 705–739 (2015)
64. Xi, F., Zhu, C.: On Feller and strong Feller properties and exponential ergodicity of regime-switching jump diffusion processes with countable regimes. *SIAM J. Control Optim.* **55**, 1789–1818 (2017)
65. Zakai, M.: On the optimal filtering of diffusion processes. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **11**, 230–243 (1969)