On long term investment optimality

Anatolii A. Puhalskii *
Institute for Problems in Information Transmission

August 9, 2019

Abstract

We study the problem of optimal long term investment with a view to beat a benchmark for a diffusion model of asset prices. Two kinds of objectives are considered. One criterion concerns the probability of outperforming the benchmark and seeks either to minimise the decay rate of the probability that a portfolio exceeds the benchmark or to maximise the decay rate that the portfolio falls short. The other criterion concerns the growth rate of an expected risk–sensitised utility of wealth which has to be either minimised, for a risk–averse investor, or maximised, for a risk–seeking investor. It is assumed that the mean returns and volatilities of the securities are affected by an economic factor, possibly, in a nonlinear fashion. The economic factor and the benchmark are modelled with general Itô differential equations. The results identify asymptotically optimal portfolios and produce the decay, or growth, rates. The proportions of wealth invested in the individual securities are time–homogeneous functions of the economic factor. Furthermore, a uniform treatment is given to the out– and under– performance probability optimisation as well as to the risk–averse and risk–seeking portfolio optimisation. It is shown that there exists a portfolio that optimises the decay rates of both the outperformance probability and the underperformance probability. While earlier research on the subject has relied, for the most part, on the techniques of stochastic optimal control and dynamic programming, in this contribution the quantities of interest are studied directly by employing the methods of the large deviation theory. The key to the analysis is to recognise the setup in question as a case of coupled diffusions with time scale separation, with the economic factor representing "the fast motion".

1 Introduction

Recently, two approaches have emerged to constructing long–term optimal portfolios for diffusion models of asset prices: optimising the risk–sensitive criterion and optimising the probability of outperforming a benchmark. In the risk–sensitive framework, one is concerned with the expected risk–sensitised utility of wealth $\mathbf{E}e^{\lambda \ln Z_t}$, where Z_t represents the portfolio's wealth at time t and λ is the risk–sensitivity parameter. If $\lambda < 0$, then $\lambda \ln(Z_t/Z_0)$ is a measure of an investor's losses, whereas it measures the investor's gains if $\lambda > 0$.

*Email: puhalski@mailfrom.ru

As the risk–averse investor (respectively, the risk–seeking investor) seeks to minimise their losses (respectively, to maximise their gains), it is arguable that λ is representative of the investor's degree of risk aversion, if negative, or of risk–seeking, if positive. (One can find a more detailed discussion in Bielecki and Pliska [5].) When trying to beat a benchmark, Y_t, the expected risk–sensitised utility of wealth is given by $\mathbf{E}e^{\lambda \ln(Z_t/Y_t)}$. Since typically those expectations grow, or decay, at an exponential rate, one is led to optimise that rate, so an optimal portfolio for the risk–averse investor (respectively, for the risk–seeking investor) is defined as the one that minimises (respectively, maximises) the limit, assuming it exists, of $(1/t)\ln \mathbf{E}e^{\lambda \ln(Z_t/Y_t)}$, as $t \to \infty$. In a similar vein, there are two ways to define the criterion when the objective is to outperform the benchmark. One can either choose the limit of $(1/t)\ln \mathbf{P}(\ln(Z_t/Y_t) \le 0)$, as $t \to \infty$, as the quantity to be minimised or the limit of $(1/t)\ln \mathbf{P}(\ln(Z_t/Y_t) \ge 0)$ as the quantity to be maximised. Arguably, the former criterion is favoured by the risk–averse investor and the latter, by the risk–seeking one. More generally, one may look at the limits of $(1/t)\ln \mathbf{P}(\ln(Z_t/Y_t) \le q)$ or of $(1/t)\ln \mathbf{P}(\ln(Z_t/Y_t) \ge q)$, for some threshold q.

Risk–sensitive optimisation has received considerable attention in the literature and has been studied under various sets of hypotheses. Bielecki and Pliska [5] consider a nonbenchmarked setting with constant volatilitities and with mean returns of the securities being affine functions of an economic factor, which is modelled as a Gaussian process satisfying a linear stochastic differential equation. For the risk–averse investor, they find an asymptotically optimal portfolio and the risk–sensitised expected growth rate. Subsequent research has relaxed some of the assumptions made, such as the independence of the diffusions driving the economic factor process and the asset price process, see Kuroda and Nagai [26], Bielecki and Pliska [6]. Fleming and Sheu [17], [18] analyse both the risk–averse and the risk–seeking setups. A benchmarked setting is studied by Davis and Lleo [11], [12], [13], the latter two papers being concerned with diffusions with jumps as driving processes. Nagai [32] assumes general mean returns and volatilities and the factor process being the solution to a general stochastic differential equation and obtains an asymptotically optimal portfolio for the risk–averse investor when there is no benchmark involved. Special one–dimensional models are treated in Fleming and Sheu [16] and Bielecki, Pliska, and Sheu [7]. The methods of the aforementioned papers rely on the tools of stochastic optimal control. A Hamilton–Jacobi–Bellman equation is invoked in order to identify a portfolio that minimises the expected risk–sensitised utility of wealth on a finite horizon. Next, a limit is taken as the length of time goes to infinity. The optimal portfolio is expressed in terms of a solution to a Riccati algebraic equation in the affine case, and to an ergodic Bellman equation, in the general case.

The criterion of the probability of outperformance is considered in Pham [34], who studies a one–dimensional affine setup. The minimisation of the underperformance probability for the Bielecki and Pliska [5] model is addressed in Hata, Nagai, and Sheu [20]. Nagai [33] studies the general model with the riskless asset as the benchmark. Those authors build on the foundation laid by the work on the risk–sensitive optimisation. Stochastic control methods are applied in order to identify an optimal risk–sensitive portfolio, first, and, afterwards, duality considerations are invoked to optimise the probabilities of out/under performance. The risk–sensitive optimal portfolio for an appropriately chosen risk–sensitivity parameter is found to be optimal for the out/under performance probability criterion. The parameter is between zero and one for the outperformance case and is negative, for the underperformance case. Puhalskii [35] analyses the out/under performance probabilities directly and obtains a portfolio that is asymptotically optimal both for the outperformance and underperformance probabilities, the limitation of their study being that it is confined to a geometric Brownian motion model of the asset prices with no economic factor involved. Puhalskii and Stutzer [37], in an unpublished manuscript, study the underperformance probability for the model in Nagai [33] with a general benchmark by applying direct methods.

Whereas the cases of a negative risk–sensitivity parameter for risk–sensitive optimisation and of the underperformance probability minimisation seem to be fairly well understood, the setups of risk–sensitive optimisation for a positive parameter and of the outperformance probability optimisation lack clarity. The reason seems to be twofold. Firstly, the expected risk–sensitised utility of wealth may grow at an infinite exponential rate for certain $\lambda \in [0,1]$, see Fleming and Sheu [18]. Secondly, the analysis of the ergodic Bellman equation presents difficulty because no Lyapunov function is readily available, cf., condition (A3) in Kaise and Sheu [23]. Although Pham [34] carries out a detailed study and identifies the threshold value of λ when "the blow–up" occurs for an affine model of one security and one factor, for the multidimensional case, we are unaware of results that produce asymptotically optimal portfolios either for the risk–seeking criterion or for the outperformance probability maximisation.

The purpose of this paper is to fill in the aforementioned gaps. We study a benchmarked version of the general model in Nagai [32, 33]. Capitalising on the insights in Puhalskii and Stutzer [37], we identify an optimal portfolio for maximising the outperformance probability. For the risk–sensitive setup, we prove that there is a threshold value $\overline{\lambda} \in (0,1]$ such that for all $\lambda < \overline{\lambda}$ there exists an asymptotically optimal risk–seeking portfolio. It is obtained as an optimal outperformance portfolio for certain threshold q. If $\lambda > \overline{\lambda}$, there is a portfolio such that the expected risk–sensitised utility of wealth grows at an infinite exponential rate. Furthermore, we give a uniform treatment to the out– and under– performance probability optimisation as well as to the risk–averse and risk–seeking portfolio optimisation. We show that the same portfolio optimises both the underperformance and outperformance probabilities, in line with conclusions in Puhalskii [35]. Similarly, the same procedure can be used for finding optimal risk–sensitive portfolios both for the risk–averse investor and for the risk–seeking investor. The portfolios are expressed in terms of solutions to ergodic Bellman equations.

No stochastic control techniques are invoked and standard tools of large deviation theory are employed, such as a change of a probability measure and an exponential Markov inequality. We treat the setup as a case of coupled diffusions with time scale separation, the factor process representing "the fast motion". The empirical measure of the factor process plays a key role in the proofs. An application of the large deviation principle to the pair comprising the portfolio and the economic factor produces a heuristic derivation of the asymptotic bounds being sought. The bounds are then confirmed rigourously. Another notable feature is an extensive use of the minimax theorem and a characterisation of the optimal portfolios in terms of saddle points. Being more direct than the one based on the stochastic optimal control theory, this approach streamlines considerations, e.g., there is no need to contend with a Hamilton–Jacobi–Bellman equation on finite time, thereby enabling us both to obtain new results and relax or drop altogether a number of assumptions present in the earlier research on the subject. For instance, we do not restrict the class of portfolios under consideration to portfolios whose total wealth is a sublinear function of the economic factor, nor do we require that the limit growth rate of the expected risk–sensitised utility of wealth be an essentially smooth (or "steep") function of the risk–sensitivity parameter, which conditions are needed in Pham [34] even for a one–dimensional model.

Our results also extend and (or) complement those in Kuroda and Nagai [26], Nagai [32], Hata, Nagai, and Sheu [20], and Nagai [33] on risk–sensitive optimisation for a negative risk–sensitivity parameter and on underperformace probability optimisation on an infinite horizon. We tackle a general benchmark and dispose of a number of assumptions some of which are questionable from the modelling perspective. Besides, our proofs seem to be cleaner, see the discussion that follows the statement of the main results in Section 2 for more detail. On the other hand, we require a certain regularity condition. Another distinction is that the cited papers assume certain stability conditions which involve the coefficients of both the equations for the economic factor and the equations for the securities, whereas the model's definition has it that the economic factor is not affected by the security prices. We use a different stability condition which is along similar lines as the one in Fleming and Sheu [18] and concerns the properties of the economic factor only.

This is how this paper is organised. In Section 2, we define the model, provide the heuristics, and state the main results. More detail is given on the relation to earlier work. The proofs are provided in Section 4 whereas Section 3 and the appendix are concerned with laying the groundwork and shedding additional light on the model of Pham [34].

2 A model description and main results

We are concerned with a market of n risky securities priced $S_t^1, \ldots, S_t^n$ at time t and a safe security of price S_t^0 at time t. We assume that, for $i = 1, 2, \ldots, n$,

$$\frac{dS_t^i}{S_t^i} = a^i(X_t)\, dt + b^i(X_t)^T\, dW_t$$

and that

$$\frac{dS_t^0}{S_t^0} = r(X_t)\, dt\,,$$

where X_t represents the economic factor. It is governed by the equation

$$dX_t = \theta(X_t)\, dt + \sigma(X_t)\, dW_t\,. \tag{2.1}$$

In the equations above, the $a^i(x)$ are real-valued functions, the $b^i(x)$ are $\mathbb{R}^k$-valued functions, $\theta(x)$ is an $\mathbb{R}^l$-valued function, $\sigma(x)$ is an $l \times k$-matrix, $W = (W_t\,, t \geq 0)$ is a k-dimensional standard Wiener process, and $S_0^i > 0\,,$ ^{T} is used to denote the transpose of a matrix or a vector. Accordingly, the process $X = (X_t\,, t \geq 0)$ is l-dimensional.

Benchmark $Y = (Y_t\,, t \geq 0)$ follows an equation similar to those for the risky securities:

$$\frac{dY_t}{Y_t} = \alpha(X_t)\, dt + \beta(X_t)^T\, dW_t,$$

where $\alpha(x)$ is an $\mathbb{R}$-valued function, $\beta(x)$ is an $\mathbb{R}^k$-valued function, and $Y_0 > 0$.

All processes are defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. It is assumed, furthermore, that the processes $S^i = (S^i_t, t \geq 0)$, X, and $Y = (Y_t, t \geq 0)$ are adapted to (right–continuous, complete) filtration $\mathbf{F} = (\mathcal{F}_t, t \geq 0)$ and that W is an $\mathbf{F}$-Wiener process.

We let $a(x)$ denote the n-vector with entries $a^1(x), \ldots, a^n(x)$, let $b(x)$ denote the $n \times k$ matrix with rows $b^1(x)^T, \ldots, b^n(x)^T$ and let $\mathbf{1}$ denote the n-vector with unit entries. The matrices $b(x)b(x)^T$ and $\sigma(x)\sigma(x)^T$ are assumed to be uniformly positive definite and bounded. The functions $a(x)$, $r(x)$, $\theta(x)$, $\alpha(x)$, $b(x)$, $\sigma(x)$, and $\beta(x)$ are assumed to be continuously differentiable with bounded derivatives and the function $\sigma(x)\sigma(x)^T$ is assumed to be twice continuously differentiable. The function $|\beta(x)|^2$ is assumed to be bounded and bounded away from zero. (We will also indicate how the results change if the benchmark "is not volatile" meaning that $\beta(x) = 0$.) Under those hypotheses, the processes S^i, X, and Y are well defined, see, e.g., chapter 5 of Karatzas and Shreve [24].

For the factor process, we assume that

$$\limsup_{|x| \to \infty} \theta(x)^T \frac{x}{|x|^2} < 0 . \tag{2.2}$$

Thus, X has a unique invariant measure, see, e.g., Bogachev, Krylov, and Röckner [9]. As for the initial condition, we will assume that $\mathbf{E} e^{\gamma |X_0|^2} < \infty$, for some $\gamma > 0$. Sometimes it will be required that $|X_0|$ be, moreover, bounded.

The investor holds l^i_t shares of risky security i and l^0_t shares of the safe security at time t, so the total wealth is given by $Z_t = \sum_{i=1}^n l^i_t S^i_t + l^0_t S^0_t$. Portfolio $\pi_t = (\pi^1_t, \ldots, \pi^n_t)^T$ specifies the proportions of the total wealth invested in the risky securities so that, for $i = 1, 2, \ldots, n$, $l^i_t S^i_t = \pi^i_t Z_t$. The processes $\pi^i = (\pi^i_t, t \geq 0)$ are assumed to be $(\mathcal{B} \otimes \mathcal{F}_t,\, t \geq 0)$–progressively measurable, where $\mathcal{B}$ denotes the Borel σ–algebra on $\mathbb{R}_+$, and such that $\int_0^t {\pi^i_s}^2 \, ds < \infty$ a.s. We do not impose any other restrictions on the magnitudes of the π^i_t so that unlimited borrowing and shortselling are allowed.

Let

$$L^\pi_t = \frac{1}{t} \ln\bigl(\frac{Z_t}{Y_t}\bigr) .$$

Since the amount of wealth invested in the safe security is $(1 - \sum_{i=1}^n \pi^i_t) Z_t$, in a standard fashion by using the self–financing condition that $dZ_t = \sum_{i=1}^n l^i_t \, dS^i_t + l^0_t \, dS^0_t$, one obtains that

$$\frac{dZ_t}{Z_t} = \sum_{i=1}^n \pi^i_t \, \frac{dS^i_t}{S^i_t} + \bigl(1 - \sum_{i=1}^n \pi^i_t\bigr) \, \frac{dS^0_t}{S^0_t} .$$

Assuming that $Z_0 = Y_0$ and letting $c(x) = b(x)b(x)^T$, we have by Itô's lemma that, cf. Pham [34],

$$L^\pi_t = \frac{1}{t} \int_0^t \bigl(\pi_s^T a(X_s) + (1 - \pi_s^T \mathbf{1}) r(X_s) - \frac{1}{2} \pi_s^T c(X_s) \pi_s - \alpha(X_s) + \frac{1}{2} |\beta(X_s)|^2\bigr) \, ds$$
$$+ \frac{1}{t} \int_0^t \bigl(b(X_s)^T \pi_s - \beta(X_s)\bigr)^T \, dW_s . \tag{2.3}$$

One can see that L_t^π is "of order one" for t great. Therefore, if one embeds the probability of outperformance $\mathbf{P}(\ln(Z_t/Y_t) \ge 0)$ (respectively, the probability of underperfomance $\mathbf{P}(\ln(Z_t/Y_t) \le 0)$) into the parameterised family of probabilities $\mathbf{P}(L_t^\pi \ge q)$ (respectively, $\mathbf{P}(L_t^\pi \le q)$), one will concern themselves with large deviation probabilities.

Let, for $u \in \mathbb{R}^n$ and $x \in \mathbb{R}^l$,

$$M(u,x) = u^T(a(x) - r(x)\mathbf{1}) - \frac{1}{2}\,u^T c(x) u + r(x) - \alpha(x) + \frac{1}{2}\,|\beta(x)|^2 \tag{2.4a}$$

and

$$N(u,x) = b(x)^T u - \beta(x)\,. \tag{2.4b}$$

A change of variables brings (2.3) to the form

$$L_t^\pi = \int_0^1 M(\pi_{ts}, X_{ts})\,ds + \frac{1}{\sqrt{t}} \int_0^1 N(\pi_{ts}, X_{ts})^T\,dW_s^t\,, \tag{2.5}$$

where $W_s^t = W_{ts}/\sqrt{t}$. The righthand side of (2.5) can be viewed as a diffusion process with a small diffusion coefficient which lives in "normal time" represented by the variable s, whereas in X_{ts} and π_{ts} "time" is accelerated by a factor of t. Furthermore, on introducing $\pi_s^t = \pi_{ts}$, $X_s^t = X_{ts}$, assuming that, for suitable function $u(\cdot)$, $\pi_s^t = u(X_s^t)$, defining

$$\Psi_s^t = \int_0^s M(u(X_{\tilde{s}}^t), X_{\tilde{s}}^t)\,d\tilde{s} + \frac{1}{\sqrt{t}} \int_0^s N(u(X_{\tilde{s}}^t), X_{\tilde{s}}^t)^T\,dW_{\tilde{s}}^t\,, \tag{2.6}$$

so that $L_t^\pi = \Psi_1^t$, and writing (2.1) as

$$X_s^t = X_0^t + t\int_0^s \theta(X_{\tilde{s}}^t)\,d\tilde{s} + \sqrt{t}\int_0^s \sigma(X_{\tilde{s}}^t)\,dW_{\tilde{s}}^t\,, \tag{2.7}$$

one can see that (2.6) and (2.7) make up a similar system of equations to those studied in Liptser [28] and in Puhalskii [36]. The heuristic derivation below which is based on the Large Deviation Principle (LDP) in Theorem 2.1 in Puhalskii [36] provides insight into our results below. It is helpful to keep in mind that $W^t = (W_s^t,\, s \in [0,1])$ is a Wiener process relative to $\mathbf{F}^t = (\mathcal{F}_{ts},\, s \in [0,1])$ and that both $X^t = (X_s^t,\, s \in [0,1])$ and $\pi^t = (\pi_s^t, s \in [0,1])$ are $\mathbf{F}^t$-adapted processes.

Let us introduce additional pieces of notation first. Let $\mathbb{C}^2$ represent the set of real–valued twice continuously differentiable functions on $\mathbb{R}^l$. For $f \in \mathbb{C}^2$, we let $\nabla f(x)$ represent the gradient of f at x which is regarded as a column l–vector and we let $\nabla^2 f(x)$ represent the $l \times l$–Hessian matrix of f at x. Let $\mathbb{C}_0^1$ and $\mathbb{C}_0^2$ represent the sets of functions of compact support on $\mathbb{R}^l$ that are once and twice continuously differentiable, respectively. Let $\mathbb{P}$ denote the set of probability densities $m = (m(x)\,, x \in \mathbb{R}^l)$ on $\mathbb{R}^l$ such that $\int_{\mathbb{R}^l} |x|^2\,m(x)\,dx < \infty$ and let $\hat{\mathbb{P}}$ denote the set of probability densities m from $\mathbb{P}$ such that $m \in \mathbb{W}^{1,1}_{\text{loc}}(\mathbb{R}^l)$ and

$\sqrt{m} \in \mathbb{W}^{1,2}(\mathbb{R}^l)$, where $\mathbb{W}$ is used for denoting a Sobolev space, see, e.g., Adams and Fournier [1]. Let $\mathbb{C}([0,1],\mathbb{R})$ represent the set of continuous real–valued functions on $[0,1]$ being endowed with the uniform topology and let $\mathbb{C}_\uparrow([0,1],\mathbb{M}(\mathbb{R}^l))$ represent the set of functions μ_t on $[0,1]$ with values in the set $\mathbb{M}(\mathbb{R}^l)$ of (nonnegative) measures on $\mathbb{R}^l$ such that $\mu_t(\mathbb{R}^l)=t$ and $\mu_t-\mu_s$ is a nonnegative measure when $t\ge s$. The space $\mathbb{M}(\mathbb{R}^l)$ is assumed to be equipped with the weak topology and the space $\mathbb{C}_\uparrow([0,1],\mathbb{M}(\mathbb{R}^l))$, with the uniform topology. Let the empirical process of X^t , which is denoted by $\mu^t=(\mu^t(ds,dx))$, be defined by the equation

$$\mu^t([0,s],\Gamma)=\int\limits_0^s \chi_\Gamma(X^t_{\tilde s})\,d\tilde s\,,$$

with Γ denoting a Borel subset of $\mathbb{R}^l$ and with $\chi_\Gamma(x)$ representing the indicator function of Γ .

If one were to apply to the processes $\Psi^t=(\Psi^t_s\,,s\in[0,1])$ and μ^t Theorem 2.1 in Puhalskii [36], then the pair (Ψ^t,μ^t) would satisfy the LDP in $\mathbb{C}([0,1],\mathbb{R})\times\mathbb{C}_\uparrow([0,1],\mathbb{M}(\mathbb{R}^l))$, as $t\to\infty$, with the deviation function (usually referred to as a rate function)

$$\begin{aligned}\mathbf{J}(\Psi,\mu)=\int\limits_0^1 \sup_{\lambda\in\mathbb{R}}\Bigl(\lambda\bigl(\dot\Psi_s-\int\limits_{\mathbb{R}^l} M(u(x),x)\,m_s(x)\,dx\bigr)-\frac{1}{2}\,\lambda^2\int\limits_{\mathbb{R}^l}|N(u(x),x)|^2\,m_s(x)\,dx\\
+\sup_{f\in\mathbb{C}_0^1}\int\limits_{\mathbb{R}^l}\Bigl(\nabla f(x)^T\bigl(\frac{1}{2}\,\text{div}\,\bigl(\sigma(x)\sigma(x)^Tm_s(x)\bigr)-\bigl(\theta(x)+\lambda\sigma(x)^TN(u(x),x)\bigr)m_s(x)\bigr)\\
-\frac{1}{2}\,|\sigma(x)^T\nabla f(x)|^2\,m_s(x)\Bigr)\,dx\Bigr)\,ds\,,\end{aligned}\quad(2.8)$$

provided the function $\Psi=(\Psi_s,\ s\in[0,1])$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}$ and the function $\mu=(\mu_s(\Gamma))$, when considered as a measure on $[0,1]\times\mathbb{R}^l$, is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}\times\mathbb{R}^l$, i.e., $\mu(ds,dx)=m_s(x)\,dx\,ds$, where $m_s(x)$, as a function of x , belongs to $\hat{\mathbb{P}}$ for almost all s . If those conditions do not hold, then $\mathbf{J}(\Psi,\mu)=\infty$. (We assume that the divergence of a square matrix is evaluated rowwise.)

Integration by parts yields an alternative form:

$$\begin{aligned}\mathbf{J}(\Psi,\mu)=\int\limits_0^1 \sup_{\lambda\in\mathbb{R}}\Bigl(\lambda\bigl(\dot\Psi_s-\int\limits_{\mathbb{R}^l} M(u(x),x)\,m_s(x)\,dx\bigr)-\frac{1}{2}\,\lambda^2\int\limits_{\mathbb{R}^l}|N(u(x),x)|^2\,m_s(x)\,dx\\
+\sup_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\Bigl(-\frac{1}{2}\,\text{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr)-\nabla f(x)^T(\theta(x)+\lambda\sigma(x)^TN(u(x),x))\\
-\frac{1}{2}\,|\sigma(x)^T\nabla f(x)|^2\,\Bigr)m_s(x)\,dx\Bigr)\,ds\,,\end{aligned}\quad(2.9)$$

with $\text{tr}\,\Sigma$ standing for the trace of square matrix Σ. Since $L^\pi_t=\Psi^t_1$, by projection, L^π_t obeys the LDP in $\mathbb{R}$ for rate t with the deviation function $\mathbf{I}(L)=\inf\{\mathbf{J}(\Psi,\mu):\ \Psi_1=L\}$. Therefore,

$$\limsup_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}(L^\pi_t\ge q)\le-\inf_{L\ge q}\mathbf{I}(L)\,.\quad(2.10)$$

The integrand against ds in (2.9) being a convex function of $\dot{\Psi}_s$ and of $m_s(x)$, along with the requirements that $\int_0^1 \dot{\Psi}_s\, ds = L$ and $\int_{\mathbb{R}^l} m_s(x)\, dx = 1$ imply, by Jensen's inequality, that one may assume that $\dot{\Psi}_s = L$ and that $m_s(x)$ does not depend on s either, so that $m_s(x) = m(x)$. Hence,

$$\inf_{L\ge q} \mathbf{I}(L) = \inf_{L\ge q} \inf_{m\in\hat{\mathbb{P}}} \sup_{\lambda\in\mathbb{R}} \Bigl(\lambda\bigl(L - \int\limits_{\mathbb{R}^l} M(u(x),x)\, m(x)\, dx\bigr) - \frac{1}{2}\lambda^2 \int\limits_{\mathbb{R}^l} |N(u(x),x)|^2\, m(x)\, dx$$

$$+ \sup_{f\in\mathbb{C}_0^2} \int\limits_{\mathbb{R}^l} \Bigl(-\frac{1}{2}\operatorname{tr}\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr) - \nabla f(x)^T(\theta(x) + \lambda\sigma(x)^T N(u(x),x))$$

$$- \frac{1}{2}|\sigma(x)^T\nabla f(x)|^2\Bigr)\, m(x)\, dx\Bigr) .$$

On noting that the expression on the righthand side is convex in (L, m) and is concave in (λ, f) , one hopes to be able to apply a minimax theorem to change the order of taking inf and sup so that

$$\inf_{L\ge q} \mathbf{I}(L) = \sup_{\lambda\in\mathbb{R}} \sup_{f\in\mathbb{C}_0^2} \inf_{L\ge q} \inf_{m\in\hat{\mathbb{P}}} \Bigl(\lambda\bigl(L - \int\limits_{\mathbb{R}^l} M(u(x),x)\, m(x)\, dx\bigr) - \frac{1}{2}\lambda^2 \int\limits_{\mathbb{R}^l} |N(u(x),x)|^2\, m(x)\, dx$$

$$+ \int\limits_{\mathbb{R}^l} \Bigl(-\frac{1}{2}\operatorname{tr}\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr) - \nabla f(x)^T(\theta(x) + \lambda\sigma(x)^T N(u(x),x))$$

$$- \frac{1}{2}|\sigma(x)^T\nabla f(x)|^2\Bigr)\, m(x)\, dx\Bigr) . \quad (2.11)$$

If $\lambda < 0$, then the infimum over $L \ge q$ equals $-\infty$. If $\lambda \ge 0$, it is attained at $L = q$ and $\inf_{m\in\hat{\mathbb{P}}}$ "is attained at a δ–density" so that (2.11) results in

$$\inf_{L\ge q} \mathbf{I}(L) = \sup_{\lambda\in\mathbb{R}_+} \sup_{f\in\mathbb{C}_0^2} \Bigl(\lambda q - \sup_{x\in\mathbb{R}^l} \bigl(\lambda M(u(x),x) + \frac{1}{2}\lambda^2 |N(u(x),x)|^2$$

$$+ \frac{1}{2}\operatorname{tr}\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr) + \nabla f(x)^T(\theta(x) + \lambda\sigma(x)^T N(u(x),x)) + \frac{1}{2}|\sigma(x)^T\nabla f(x)|^2\bigr)\Bigr) . \quad (2.12)$$

For an optimal outperforming portfolio, one wants to maximise the righthand side of (2.10) over functions $u(x)$, so the righthand side of (2.12) has to be minimised. Assuming one can apply minimax considerations once again yields

$$\inf_{u(\cdot)} \inf_{L\ge q} \mathbf{I}(L) = \sup_{\lambda\in\mathbb{R}_+} \sup_{f\in\mathbb{C}_0^2} \Bigl(\lambda q - \sup_{x\in\mathbb{R}^l} \sup_{u\in\mathbb{R}^n} \bigl(\lambda M(u,x) + \frac{1}{2}\lambda^2 |N(u,x)|^2$$

$$+ \nabla f(x)^T(\theta(x) + \lambda\sigma(x)^T N(u,x))\bigr) + \frac{1}{2}\operatorname{tr}\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr) + \frac{1}{2}|\sigma(x)^T\nabla f(x)|^2\Bigr) .$$

By (2.4a) and (2.4b), the $\sup_{u\in\mathbb{R}^n} = \infty$ if $\lambda > 1$ so, on recalling (2.10), it is reasonable to conjecture that

$$\sup_{\pi}\limsup_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}(L_t^\pi\geq q)=-\sup_{\lambda\in[0,1]}\sup_{f\in\mathbb{C}_0^2}\Bigl(\lambda q-\sup_{x\in\mathbb{R}^l}\sup_{u\in\mathbb{R}^n}\bigl(\lambda M(u,x)+\frac{1}{2}\,\lambda^2|N(u,x)|^2$$
$$+\nabla f(x)^T(\theta(x)+\lambda\sigma(x)^TN(u,x))\bigr)+\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2f(x)\bigr)+\frac{1}{2}\,|\sigma(x)^T\nabla f(x)|^2\Bigr)\quad(2.13)$$

and an optimal portfolio is of the form $u(X_t)$, with $u(x)$ attaining the supremum with respect to u on the righthand side of (2.13) for λ and f that deliver their respective suprema. Similar arguments may be applied to finding $\inf_\pi\liminf_{t\to\infty}(1/t)\,\ln\mathbf{P}(L_t^\pi<q)$. Unfortunately, we are unable to fill in the gaps in the above derivation, e.g., in order for the results of Puhalskii [36] to apply, the function $u(x)$ has to be bounded in x, while the optimal portfolio typically is not. Besides, it is not at all obvious that the optimal portfolio should be expressed as a function of the economic factor. Nevertheless, the above line of reasoning is essentially correct, as our results show. Besides, there is a special case which we analyse at the final stages of our proofs that allows a direct application of Theorem 2.1 in Puhalskii [36]. We now proceed to stating the results. That requires introducing more pieces of notation and foreshadowing certain properties to be proved later.

The following nondegeneracy condition is needed. Let I_k denote the $k\times k$–identity matrix and let

$$Q_1(x)=I_k-b(x)^Tc(x)^{-1}b(x)\,.$$

The matrix $Q_1(x)$ represents the orthogonal projection operator onto the null space of $b(x)$ in $\mathbb{R}^k$. We will assume the following "general position" condition:

(N) 1. The matrix $\sigma(x)Q_1(x)\sigma(x)^T$ is uniformly positive definite.

2. The quantity $\beta(x)^TQ_2(x)\beta(x)$ is bounded away from zero, where

$$Q_2(x)=Q_1(x)\bigl(I_k-\sigma(x)^T(\sigma(x)Q_1(x)\sigma(x)^T)^{-1}\sigma(x)\bigr)Q_1(x)\,.\quad(2.14)$$

It admits the following geometric interpretation.

Lemma 2.1. *The matrix $\sigma(x)Q_1(x)\sigma(x)^T$ is uniformly positive definite if and only if arbitrary nonzero vectors from the ranges of $\sigma(x)^T$ and $b(x)^T$, respectively, are at angles bounded away from zero if and only if the matrix $c(x)-b(x)\sigma(x)^T(\sigma(x)\sigma(x)^T)^{-1}\sigma(x)b(x)^T$ is uniformly positive definite. Also, $\beta(x)^TQ_2(x)\beta(x)$ is bounded away from zero if and only if the projection of $\beta(x)$ onto the null space of $b(x)$ is of length bounded away from zero and is at angles bounded away from zero to all projections onto that null space of nonzero vectors from the range of $\sigma(x)^T$.*

The proof of the lemma is provided in the appendix. Under part 1 of condition (N), we have that $k\geq n+l$ and the rows of the matrices $\sigma(x)$ and $b(x)$ are linearly independent. Part 2 of condition (N) implies that $\beta(x)$ does not belong to the sum of the ranges of $b(x)^T$ and of $\sigma(x)^T$. (Indeed, if that were the case, then $Q_1(x)\beta(x)$, which is the projection of $\beta(x)$ onto the null space of $b(x)$, would also be the projection of a vector from the range of $\sigma(x)^T$ onto the null space of $b(x)$.) Thus, $k>n+l$.

The righthand side of (2.13) motivates the following definitions. Let, given $x \in \mathbb{R}^l$, $\lambda \in \mathbb{R}$, and $p \in \mathbb{R}^l$,

$$\breve{H}(x;\lambda,p) = \lambda \sup_{u\in\mathbb{R}^n} \big(M(u,x)+\frac{1}{2}\,\lambda|N(u,x)|^2+p^T\sigma(x)N(u,x)\big)+p^T\theta(x)+\frac{1}{2}\,|\sigma(x)^Tp|^2\,. \quad (2.15)$$

By (2.4a) and (2.4b), the latter righthand side is finite if $\lambda < 1$, with the supremum being attained at

$$u(x) = \frac{1}{1-\lambda}\,c(x)^{-1}\big(a(x)-r(x)\mathbf{1}-\lambda b(x)\beta(x)+b(x)\sigma(x)^Tp\big)\,. \quad (2.16)$$

Furthermore,

$$\sup_{u\in\mathbb{R}^n}\big(M(u,x)+\frac{1}{2}\,\lambda|N(u,x)|^2+p^T\sigma(x)N(u,x)\big) \\ = \frac{1}{2}\,\frac{1}{1-\lambda}\,\|a(x)-r(x)\mathbf{1}-\lambda b(x)\beta(x)+b(x)\sigma(x)^Tp\|^2_{c(x)^{-1}} \\ +\frac{1}{2}\,\lambda|\beta(x)|^2+r(x)-\alpha(x)+\frac{1}{2}\,|\beta(x)|^2-\beta(x)^T\sigma(x)^Tp\,, \quad (2.17)$$

where, for $y \in \mathbb{R}^n$ and positive definite symmetric $n\times n$–matrix Σ, we denote $\|y\|^2_\Sigma = y^T\Sigma y$.

Therefore, on introducing

$$T_\lambda(x) = \sigma(x)\sigma(x)^T+\frac{\lambda}{1-\lambda}\,\sigma(x)b(x)^Tc(x)^{-1}b(x)\sigma(x)^T, \quad (2.18a)$$

$$S_\lambda(x) = \frac{\lambda}{1-\lambda}\,(a(x)-r(x)\mathbf{1}-\lambda b(x)\beta(x))^Tc(x)^{-1}b(x)\sigma(x)^T-\lambda\beta(x)^T\sigma(x)^T+\theta(x)^T\,, \quad (2.18b)$$

and

$$R_\lambda(x) = \frac{\lambda}{2(1-\lambda)}\,\|a(x)-r(x)\mathbf{1}-\lambda b(x)\beta(x)\|^2_{c(x)^{-1}}+\lambda(r(x)-\alpha(x)+\frac{1}{2}\,|\beta(x)|^2) \\ +\frac{1}{2}\,\lambda^2|\beta(x)|^2\,, \quad (2.18c)$$

we have that

$$\breve{H}(x;\lambda,p) = \frac{1}{2}\,p^TT_\lambda(x)p+S_\lambda(x)p+R_\lambda(x)\,. \quad (2.19)$$

Let us note that, by condition (N), $T_\lambda(x)$ is a uniformly positive definite matrix.

If $\lambda = 1$, then, on noting that

$$M(u,x)+\frac{1}{2}\,|N(u,x)|^2+p^T\sigma(x)N(u,x) = u^T(a(x)-r(x)\mathbf{1}-b(x)\beta(x)+b(x)\sigma(x)^Tp) \\ +r(x)-\alpha(x)+|\beta(x)|^2-p^T\sigma(x)\beta(x)\,, \quad (2.20)$$

we have that $\breve{H}(x;1,p) < \infty$ if and only if

$$a(x)-r(x)\mathbf{1}-b(x)\beta(x)+b(x)\sigma(x)^Tp = 0\,, \quad (2.21)$$

in which case

$$\breve{H}(x;1,p) = r(x) - \alpha(x) + |\beta(x)|^2 - p^T\sigma(x)\beta(x) + p^T\theta(x) + \frac{1}{2}\,|\sigma(x)^T p|^2\,. \tag{2.22}$$

As mentioned, if $\lambda > 1$, then the righthand side of (2.15) equals infinity. Consequently, $\breve{H}(x;\lambda,p)$ is a lower semicontinuous function of (λ,p) with values in $\mathbb{R}\cup\{+\infty\}$. By Lemma 3.5 below, $\breve{H}(x;\lambda,p)$ is convex in (λ,p).

We define, given $f \in \mathbb{C}^2$,

$$H(x;\lambda,f) = \breve{H}(x;\lambda,\nabla f(x)) + \frac{1}{2}\,\mathrm{tr}\left(\sigma(x)\sigma(x)^T\nabla^2 f(x)\right). \tag{2.23}$$

By the convexity of $\breve{H}$, the function $H(x;\lambda,f)$ is convex in (λ,f).

Let

$$F(\lambda) = \inf_{f\in\mathbb{C}^2}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f) \text{ if } \lambda < 1\,, \tag{2.24}$$

$F(1) = \lim_{\lambda\uparrow 1} F(\lambda)$, $F(\lambda) = \infty$ if $\lambda > 1$, and $\overline{\lambda} = \sup\{\lambda\in\mathbb{R}:\ F(\lambda) < \infty\}$. By $H(x;\lambda,f)$ being convex in (λ,f), $F(\lambda)$ is convex for $\lambda < 1$, so $F(1)$ is well defined, see, e.g., Theorem 7.5 on p.57 in Rockafellar [38]. As a function on $\mathbb{R}$, $F(\lambda)$ is seen to be convex and proper (i.e., $F(\lambda) > -\infty$, see Remark 3.3). It is finite when $\lambda < \lambda_0$, for some $\lambda_0 \in (0,1]$, which is obtained by taking $f(x) = \kappa|x|^2$, $\kappa > 0$ being small enough (see Lemma 3.1 for more detail). Therefore $\overline{\lambda} \in (0,1]$. Lemma 3.4 below establishes that $F(0) = 0$, that $F(\lambda)$ is lower semicontinuous on $\mathbb{R}$ and that if $F(\lambda)$ is finite, with $\lambda < 1$, then the infimum in (2.24) is attained at function f^λ which satisfies the equation

$$H(x;\lambda,f^\lambda) = F(\lambda)\,,\ \text{for all } x\in\mathbb{R}^l\,. \tag{2.25}$$

Furthermore, $f^\lambda \in \mathbb{C}^1_\ell$, with $\mathbb{C}^1_\ell$ representing the set of real–valued continuously differentiable functions on $\mathbb{R}^l$ whose gradients grow no faster than linearly. Thus, the infimum in (2.24) can be taken over $\mathbb{C}^2\cap\mathbb{C}^1_\ell$ when $\lambda < 1$. Equation (2.25) is referred to as an ergodic Bellman equation, see, e.g., Fleming and Sheu [18], Kaise and Sheu [23], Hata, Nagai, and Sheu [20], Ichihara [21].

Let

$$J^{\mathrm{o}}_q = \sup_{\lambda\in[0,1]}\left(\lambda q - F(\lambda)\right) \tag{2.26a}$$

and

$$J^{\mathrm{s}}_q = \sup_{\lambda\le 0}(\lambda q - F(\lambda))\,. \tag{2.26b}$$

We note that $-J^{\mathrm{o}}_q$ is the righthand side of (2.13).

Theorem 2.1. *1. If $|X_0|$ is bounded and, for all $0 < \lambda < \overline{\lambda}$, there exist minimisers $f^\lambda(x)$ of the righthand sides of (2.24) which are bounded below by affine functions of x, then*

$$\limsup_{t\to\infty}\frac{1}{t}\ln\mathbf{P}(L^\pi_t \ge q) \le -J^o_q\,. \tag{2.27}$$

2.

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^\pi < q) \ge -J_q^s\,. \tag{2.28}$$

Remark 2.1. The requirement that $f^\lambda(x)$ be bounded below by affine functions when $0 < \lambda < \overline{\lambda}$ is fulfilled for the affine model, as we discuss below.

Our next aim is to produce a portfolio that will attain the bounds. Let $\mathcal{P}$ represent the set of probability measures ν on $\mathbb{R}^l$ such that $\int_{\mathbb{R}^l} |x|^2\,\nu(dx) < \infty$. For $\nu \in \mathcal{P}$, we let $\mathbb{L}^2(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$ represent the Hilbert space (of the equivalence classes) of $\mathbb{R}^l$-valued functions $h(x)$ on $\mathbb{R}^l$ that are square integrable with respect to $\nu(dx)$ equipped with the norm $\bigl(\int_{\mathbb{R}^l} |h(x)|^2\,\nu(dx)\bigr)^{1/2}$ and we let $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$ represent the closure in $\mathbb{L}^2(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$ of the set of gradients of $\mathbb{C}_0^1$-functions. We will retain the notation ∇f for the elements of $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$, although those functions might not be proper gradients. Let $\mathcal{U}_\lambda$ denote the set of functions $f \in \mathbb{C}^2 \cap \mathbb{C}^1_\ell$ such that $\sup_{x\in\mathbb{R}^l} H(x;\lambda,f) < \infty$. The set $\mathcal{U}_\lambda$ is nonempty if and only if $F(\lambda) < \infty$. It is convenient to write (2.24) in the form, cf. (2.11),

$$F(\lambda) = \inf_{f\in\mathcal{U}_\lambda} \sup_{\nu\in\mathcal{P}} \int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,, \quad \text{if } \lambda < 1, \tag{2.29}$$

the latter integral possibly being equal to $-\infty$. We adopt the convention that $\inf_\emptyset = \infty$, so that (2.29) holds when $\mathcal{U}_\lambda = \emptyset$ too. Let $\mathbb{C}_b^2$ represent the subset of $\mathbb{C}^2$ of functions with bounded second derivatives. Let, for $f \in \mathbb{C}_b^2$ and $m \in \mathbb{P}$,

$$G(\lambda,f,m) = \int_{\mathbb{R}^l} H(x;\lambda,f)\,m(x)\,dx\,. \tag{2.30}$$

This function is well defined, is convex in (λ,f) and is concave in m. By Lemma 3.5 and Lemma 3.6 below, for $\lambda < \overline{\lambda}$, $F(\lambda) = \sup_{m\in\hat{\mathbb{P}}} \inf_{f\in\mathbb{C}_0^2} G(\lambda,f,m)$. One can replace $\hat{\mathbb{P}}$ with $\mathbb{P}$ in the preceding sup and replace $\mathbb{C}_0^2$ with $\mathbb{C}_b^2$ in the preceding inf. If $m \in \hat{\mathbb{P}}$, then integration by parts in (2.30) obtains that, for $f \in \mathbb{C}_b^2$,

$$G(\lambda,f,m) = \breve{G}(\lambda,\nabla f,m)\,, \tag{2.31}$$

where

$$\breve{G}(\lambda,\nabla f,m) = \int_{\mathbb{R}^l} \Bigl(\breve{H}(x;\lambda,\nabla f(x)) - \frac{1}{2}\,\nabla f(x)^T\,\frac{\operatorname{div}(\sigma(x)\sigma(x)^T\,m(x))}{m(x)}\Bigr)\,m(x)\,dx\,. \tag{2.32}$$

(Unless specifically mentioned otherwise, it is assumed throughout that $0/0 = 0$. More detail on the integration by parts is given in the proof of Lemma 3.4.) The function $\breve{G}(\lambda,\nabla f,m)$ is convex in (λ,f) and is concave in m. The righthand side of (2.32) being well defined for $\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$, we adopt (2.32) as the definition of $\breve{G}(\lambda,\nabla f,m)$ for $(\lambda,\nabla f,m) \in \mathbb{R}\times\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)\times\hat{\mathbb{P}}$.

Let, for $m \in \hat{\mathbb{P}}$,

$$\breve{F}(\lambda,m) = \inf_{\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \breve{G}(\lambda,\nabla f,m)\,, \tag{2.33}$$

when $\lambda \le 1$ and let $\breve{F}(\lambda,m)=\infty$, for $\lambda>1$. By Lemma 3.5 below, the infimum in (2.33) is attained uniquely, if finite, the latter always being the case for $\lambda<1$. Furthermore, if $\lambda<1$, then $\breve{F}(\lambda,m)=\inf_{f\in\mathbb{C}_0^2} G(\lambda,f,m)$. By (2.32), the function $\breve{F}(\lambda,m)$ is convex in λ and is concave in m. It is lower semicontinuous in λ and is strictly convex on $(-\infty,1)$ by Lemma 3.5, so, by convexity, see Corollary 7.5.1 on p.57 in Rockafellar [38], $\breve{F}(1,m)=\lim_{\lambda\uparrow 1}\inf_{f\in\mathbb{C}_0^2} G(\lambda,f,m)$. By Lemma 3.6 below, $\lambda q-\breve{F}(\lambda,m)$ has saddle point $(\hat{\lambda},\hat{m})$ in $(-\infty,\overline{\lambda}]\times\hat{\mathbb{P}}$, with $\hat{\lambda}$ being specified uniquely, and the supremum of $\lambda q-F(\lambda)$ over $\mathbb{R}$ is attained at $\hat{\lambda}$. It is noteworthy that if $\hat{\lambda}<0$, then $J_q^{\mathrm{s}}>0$ and $J_q^0=0$, while if $\hat{\lambda}>0$, then $J_q^{\mathrm{o}}>0$ and $J_q^{\mathrm{s}}=0$, if $\hat{\lambda}=0$, then $J_q^{\mathrm{o}}=J_q^{\mathrm{s}}=0$. Consequently, $J_q^{\mathrm{o}}\vee J_q^{\mathrm{s}}=\hat{\lambda}q-F(\hat{\lambda})$, where $a\vee b=\max(a,b)$.

If $\hat{\lambda}<1$, which is "the regular case", then $\hat{m}$ is specified uniquely and there exists $\hat{f}\in\mathbb{C}^2\cap\mathbb{C}_\ell^1$ such that $(\hat{\lambda},\hat{f},\hat{m})$ is a saddle point of the function $\lambda q-\breve{G}(\lambda,\nabla f,m)$ in $\mathbb{R}\times(\mathbb{C}^2\cap\mathbb{C}_\ell^1)\times\hat{\mathbb{P}}$, with $\nabla\hat{f}$ being specified uniquely, see Lemma 3.6. As a matter of fact, $\hat{f}$ is a minimizer for the righthand side of (2.24), so, it satisfies the ergodic Bellman equation

$$H(x;\hat{\lambda},\hat{f})=F(\hat{\lambda})\,,\ \text{for all } x\in\mathbb{R}^l\,. \tag{2.34}$$

We define $\hat{u}(x)$ as the u that attains supremum in (2.15) for $\lambda=\hat{\lambda}$ and $p=\nabla\hat{f}(x)$ so that, by (2.16),

$$\hat{u}(x)=\frac{1}{1-\hat{\lambda}}\,c(x)^{-1}\big(a(x)-r(x)\mathbf{1}-\hat{\lambda}b(x)\beta(x)+b(x)\sigma(x)^T\nabla\hat{f}(x)\big)\,. \tag{2.35}$$

The density $\hat{m}$ is the invariant density of a diffusion process in that

$$\int_{\mathbb{R}^l}\big(\nabla h(x)^T(\hat{\lambda}\sigma(x)N(\hat{u}(x),x)+\theta(x)+\sigma(x)\sigma(x)^T\nabla\hat{f}(x))+\frac{1}{2}\operatorname{tr}\left(\sigma(x)\sigma(x)^T\,\nabla^2 h(x)\right)\big)\,\hat{m}(x)\,dx=0\,, \tag{2.36}$$

for all $h\in\mathbb{C}_0^2$. Essentially, equations (2.34) and (2.36) represent Euler–Lagrange equations for $\breve{G}(\hat{\lambda},\nabla f,m)$ at $(\hat{f},\hat{m})$. They specify $\nabla\hat{f}$ and $\hat{m}$ uniquely and imply that $(\hat{f},\hat{m})$ is a saddle point of $\breve{G}(\hat{\lambda},\nabla f,m)$, cf., Proposition 1.6 on p.169 in Ekeland and Temam [14].

Suppose that $\hat{\lambda}=1$, which is "the degenerate case". Necessarily, $\overline{\lambda}=1$, so, the infimum on the righthand side of (2.33) for $\lambda=1$ and $m=\hat{m}$ is finite and is attained at unique $\nabla\hat{f}$ (see Lemma 3.5). Consequently, $F(1)<\infty$. According to Lemma 3.6 below, cf., (2.21) and (2.36),

$$a(x)-r(x)\mathbf{1}-b(x)\beta(x)+b(x)\sigma(x)^T\nabla\hat{f}(x)=0\quad \hat{m}(x)dx\text{–a.e.} \tag{2.37}$$

and

$$\int_{\mathbb{R}^l}\big(\nabla h(x)^T\big(-\sigma(x)\beta(x)+\theta(x)+\sigma(x)\sigma(x)^T\nabla\hat{f}(x)\big)+\frac{1}{2}\operatorname{tr}\big(\sigma(x)\sigma(x)^T\nabla^2 h(x)\big)\big)\hat{m}(x)\,dx=0\,,$$

provided that $h\in\mathbb{C}_0^2$ and $b(x)\sigma(x)^T\nabla h(x)=0$ $\hat{m}(x)\,dx$–a.e. By (2.20), the value of the expression in the supremum in (2.15) does not depend on the choice of u

when $\lambda = 1$ and $p = \nabla\hat{f}(x)$, so, there is some leeway as to the choice of an optimal control. As the concave function $\lambda q - \breve{F}(\lambda, \hat{m})$ attains maximum at $\lambda = 1$, $d/d\lambda\, \breve{F}(\lambda, \hat{m})\Big|_{1-} \le q$, with $d/d\lambda\, \breve{F}(\lambda, \hat{m})\Big|_{1-}$ standing for the lefthand derivative of $\breve{F}(\lambda, \hat{m})$ at $\lambda = 1$. Hence, there exists bounded continuous function $\hat{v}(x)$ with values in the range of $b(x)^T$ such that $|\hat{v}(x)|^2/2 = q - d/d\lambda\, \breve{F}(\lambda, \hat{m})\Big|_{1-}$. (For instance, one can take $\hat{v}(x) = b(x)^T c(x)^{-1/2}\, z \sqrt{2(q - d/d\lambda\, \breve{F}(\lambda, \hat{m})\Big|_{1-})}$, where z represents an element of $\mathbb{R}^n$ of length one.) We let $\hat{u}(x) = c(x)^{-1} b(x)(\beta(x) + \hat{v}(x))$.

In either case, we define $\hat{\pi}_t = \hat{u}(X_t)$. The next theorem provides conditions for $\hat{\pi} = (\hat{\pi}_t\,, t \ge 0)$ to be an asymptotically optimal portfolio.

Theorem 2.2. *1.*

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\hat{\pi}} > q) \ge -J_q^o\,. \tag{2.38}$$

2. If

$$\lim_{|x|\to\infty} \frac{1}{|x|} \left(\|b(x)\sigma(x)^T \nabla\hat{f}(x)\|^2_{c(x)^{-1}} - \|a(x) - r(x)\mathbf{1}\|^2_{c(x)^{-1}} \right) = -\infty\,, \tag{2.39}$$

then

$$\limsup_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\hat{\pi}} \le q) \le -J_q^s\,. \tag{2.40}$$

Remark 2.2. As a consequence, if conditions of part 1 of Theorem 2.1 hold, then

$$\lim_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\hat{\pi}} > q) = \sup_{\pi} \limsup_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\pi} \ge q) = -J_q^{\mathrm{o}}\,.$$

If conditions of part 2 of Theorem 2.2 hold, then

$$\lim_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\hat{\pi}} \le q) = \inf_{\pi} \liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\pi} < q) = -J_q^{\mathrm{s}}\,.$$

Remark 2.3. The upper bounds in (2.27) and in (2.40) are of interest only if $\hat{\lambda} > 0$ and $\hat{\lambda} < 0$, respectively.

Remark 2.4. When $\beta(x) = 0$, the proofs of Theorems 2.1 and 2.2 go through and their assertions are maintained provided part 1 of condition (N) is satisfied and $\inf_{x\in\mathbb{R}^l}(r(x) - \alpha(x)) < q$. If $\inf_{x\in\mathbb{R}^l}(r(x) - \alpha(x)) \ge q$, then investing in the safe security only is obviously optimal.

Remark 2.5. One can relax condition (2.2) and require that there exist a positive definite symmetric $l \times l$–matrix Φ such that

$$\limsup_{|x|\to\infty} \theta(x)^T \frac{\Phi x}{|x|^2} < 0\,.$$

The following theorem states risk–sensitive optimality properties of $\hat{\pi}$. More specifically, it shows that, given risk–sensitivity parameter λ, the portfolio $\hat{\pi}$ is risk–sensitive optimal for appropriately chosen q. If F is subdifferentiable at λ, we let $\hat{u}^\lambda(x)$ represent the function $\hat{u}(x)$ for a value of q that is a subgradient of F at λ. We also let $\hat{\pi}_t^\lambda = \hat{u}^\lambda(X_t)$, and $\hat{\pi}^\lambda = (\hat{\pi}_t^\lambda,\, t \ge 0)$. The function F is subdifferentiable at $\lambda < \overline{\lambda}$. It might not be subdifferentiable at $\overline{\lambda}$.

Theorem 2.3. *1. If* $0 < \lambda < \overline{\lambda}$, *if, for all small enough* $\epsilon > 0$, *there exist* $f^{\lambda(1+\epsilon)}(x)$ *(as in* (2.25)*) which are bounded below by affine functions of* x *and if* $|X_0|$ *is bounded, then, for any portfolio* π,

$$\limsup_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\pi}} \le F(\lambda)\,.$$

If either $0 < \lambda < \overline{\lambda}$ *or* $\lambda = \overline{\lambda}$ *and* F *is subdifferentiable at* $\overline{\lambda}$, *then*

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^\lambda}} \ge F(\lambda)\,.$$

If either $\lambda = \overline{\lambda}$ *and* F *is not subdifferentiable at* $\overline{\lambda}$ *or* $\lambda > \overline{\lambda}$, *then there exists portfolio* π^λ *such that*

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\pi^\lambda}} \ge F(\lambda)\,.$$

2. If $\lambda < 0$, *then, for any portfolio* π,

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\pi}} \ge F(\lambda)$$

and, provided (2.39) *holds for* $\hat{\lambda} = \lambda$,

$$\limsup_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^\lambda}} \le F(\lambda)\,.$$

Remark 2.6. The inequalities in parts 1 and 2 imply that, if $\lambda < \overline{\lambda}$, then, in a fairly general situation,

$$\lim_{t\to\infty} \frac{1}{t} \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^\lambda}} = F(\lambda)\,.$$

Remark 2.7. We recall that $F(\lambda) = \infty$ if $\lambda > \overline{\lambda}$. For a one–dimensional model, $\overline{\lambda}$ is found explicitly in Pham [34], also, see the appendix below. We conjecture that F is differentiable and strictly convex for $\lambda < \overline{\lambda}$, which would imply that π^λ is specified uniquely. This is provably the case for the model of Pham [34] and provided $\lambda < 0$, see Pham [34] and Lemma 3.7 (or Puhalskii and Stutzer [37]), respectively.

If we assume that the functions $a(x)$, $r(x)$, $\alpha(x)$ and $\theta(x)$ are affine functions of x and that the diffusion coefficients are constant so that X is a Gaussian process, then fairly explicit formulas are available. More specifically, let

$$a(x) = A_1 x + a_2\,, \tag{2.41a}$$
$$r(x) = r_1^T x + r_2\,, \tag{2.41b}$$
$$\alpha(x) = \alpha_1^T x + \alpha_2\,, \tag{2.41c}$$
$$\theta(x) = \Theta_1 x + \theta_2\,, \tag{2.41d}$$

and

$$b(x) = b,\ \beta(x) = \beta,\ \sigma(x) = \sigma\,, \tag{2.41e}$$

where $A_1 \in \mathbb{R}^{n\times l}$, $a_2 \in \mathbb{R}^n$, $r_1 \in \mathbb{R}^l$, $r_2 \in \mathbb{R}$, $\alpha_1 \in \mathbb{R}^l$, $\alpha_2 \in \mathbb{R}$, Θ_1 is a negative definite $l\times l$-matrix (in fact, owing to Remark 2.5, one can only require that Θ_1 be stable), $\theta_1 \in \mathbb{R}^l$, b is an $n \times k$-matrix such that the matrix bb^T is positive definite, β is a non-zero k-vector, and σ is an $l \times k$-matrix such that the matrix $\sigma\sigma^T$ is positive definite. Condition (N) expresses the requirement that the ranges of σ^T and b^T have the trivial intersection and that β is not an element of the sum of those ranges.

Finding the optimal portfolio $\hat{\pi}_t$ may be reduced to solving an algebraic Riccati equation. We introduce, for $\lambda < 1$,

$$A(\lambda) = \Theta_1 + \frac{\lambda}{1-\lambda}\,\sigma b^T c^{-1}(A_1 - \mathbf{1}r_1^T), \tag{2.42a}$$

$$B(\lambda) = T_\lambda(x) = \sigma\sigma^T + \frac{\lambda}{1-\lambda}\,\sigma b^T c^{-1} b\sigma^T\,, \tag{2.42b}$$

and

$$C = \|A_1 - \mathbf{1}r_1^T\|^2_{c^{-1}}\,. \tag{2.42c}$$

Let us suppose that there exists symmetric $l \times l$–matrix $P_1(\lambda)$ that satisfies the algebraic Riccati equation

$$P_1(\lambda)B(\lambda)P_1(\lambda) + A(\lambda)^T P_1(\lambda) + P_1(\lambda)A(\lambda) + \frac{\lambda}{1-\lambda}\,C = 0\,. \tag{2.43}$$

Conditions for the existence of solutions can be found in Fleming and Sheu [18], see also Willems [42] and Wonham [43]. According to Lemma 3.3 in Fleming and Sheu [18], provided that $\lambda < 0$, there exists unique $P_1(\lambda)$ solving (2.43) such that $P_1(\lambda)$ is negative semidefinite. Furthermore, the matrix

$$D(\lambda) = A(\lambda) + B(\lambda)P_1(\lambda) \tag{2.44}$$

is stable. If $0 < \lambda < 1$ and $F(\lambda) < \infty$, then, by Lemma 4.3 in Fleming and Sheu [18], there exists unique $P_1(\lambda)$ solving (2.43) such that $P_1(\lambda)$ is positive semidefinite and $D(\lambda)$ is semistable. By Theorem 4.6 in Fleming and Sheu [18], the matrix $D(\lambda)$ is stable if λ is small enough.

With $D(\lambda)$ being stable, the equation

$$D(\lambda)^T p_2(\lambda) + E(\lambda) = 0 \tag{2.45}$$

has a unique solution for $p_2(\lambda)$, where

$$E(\lambda) = \frac{\lambda}{1-\lambda}\,(A_1 - \mathbf{1}r_1^T + b\sigma^T P_1(\lambda))^T c^{-1}(a_2 - r_2\mathbf{1} - \lambda b\beta) + \lambda(r_1 - \alpha_1 - P_1(\lambda)\sigma\beta) + P_1(\lambda)\theta_2\,. \tag{2.46}$$

Substitution shows that $H(x;\lambda,\tilde{f}^\lambda)$, with $\tilde{f}^\lambda(x) = x^T P_1(\lambda)x/2 + p_2(\lambda)^T x$, does not depend on x. Let m^λ denote the invariant distribution of the linear diffusion

$$dY_t = D(\lambda)Y_t\,dt + \big(\frac{\lambda}{1-\lambda}\,\sigma b^T c^{-1}(a_2 - r_2\mathbf{1} - \lambda b\beta + b\sigma^T p_2(\lambda)) - \lambda\sigma\beta + \sigma\sigma^T p_2(\lambda) + \theta_2\big)\,dt + \sigma\,dW_t\,. \tag{2.47}$$

Then the pair $(\tilde{f}^\lambda, m^\lambda)$ is a saddle point of $\breve{G}(\lambda, \nabla f, m)$ as well as of $G(\lambda, f, m)$ considered as functions of $(f, m) \in \mathcal{U}_\lambda \times \hat{\mathbb{P}}$. Hence,

$$H(x; \lambda, f^\lambda) = \breve{G}(\lambda, \nabla f^\lambda, m^\lambda) = \inf_{f \in \mathcal{U}_\lambda} \sup_{m \in \hat{\mathbb{P}}} \breve{G}(\lambda, \nabla f, m) = \inf_{f \in \mathcal{U}_\lambda} \sup_{m \in \mathbb{P}} G(\lambda, f, m)$$
$$= \inf_{f \in \mathcal{U}_\lambda} \sup_{x \in \mathbb{R}^l} H(x; \lambda, f) = F(\lambda),$$

so $\tilde{f}^\lambda$ satisfies the Bellman equation (2.25). As a result, under the hypotheses of Fleming and Sheu [18], $\tilde{f}^\lambda$ is bounded below by an affine function when $\lambda \in (0,1)$ and $F(\lambda) < \infty$. Condition (2.39) is implied by the condition that the matrix $(b\sigma^T P_1(\hat{\lambda}))^T c^{-1} b\sigma^T P_1(\hat{\lambda}) - (A_1 - \mathbf{1}r_1^T)^T c^{-1}(A_1 - \mathbf{1}r_1^T)$ is negative definite.

Furthermore, one can see that

$$F(\lambda) = \frac{1}{2}\,\|p_2(\lambda)\|^2_{\sigma\sigma^T} + \frac{1}{2}\,\frac{\lambda}{1-\lambda}\,\|a_2 - r_2\mathbf{1} - \lambda b\beta + b\sigma^T p_2(\lambda)\|^2_{c^{-1}}$$
$$+ (-\lambda\beta^T\sigma^T + \theta_2^T)p_2(\lambda) + \lambda(r_2 - \alpha_2 + \frac{1}{2}\,|\beta|^2) + \frac{1}{2}\,\lambda^2|\beta|^2 + \frac{1}{2}\,\mathrm{tr}\,(\sigma\sigma^T P_1(\lambda)). \quad (2.48)$$

If $\hat{\lambda} < 1$, equation (2.35) is as follows

$$\hat{u}(x) = \frac{1}{1-\hat{\lambda}}\,c^{-1}\big(A_1 - \mathbf{1}r_1^T + b\sigma^T P_1(\hat{\lambda}))x + \frac{1}{1-\hat{\lambda}}\,c^{-1}\big(a_2 - r_2\mathbf{1} - \hat{\lambda}b\beta + b\sigma^T p_2(\hat{\lambda})\big). \quad (2.49)$$

If $\hat{\lambda} = 1$, then one may look, once again, for $\hat{f}(x) = x^T P_1(1)x/2 + p_2(1)^T x$. Substitution in (2.37) yields

$$A_1 - \mathbf{1}r_1^T + b\sigma^T P_1(1) = 0, \quad (2.50a)$$
$$a_2 - r_2\mathbf{1} - b\beta + b\sigma^T p_2(1) = 0. \quad (2.50b)$$

(One can also obtain (2.50a) by multiplying (2.43) through with $1-\lambda$ and taking a formal limit as $\lambda \uparrow 1$.) If those conditions hold, choosing $\hat{f}(x)$ quadratic is justified. An optimal control is $\hat{u}(x) = c^{-1}(b\beta + \hat{v})$, with $\hat{v}$ coming from the range of b^T and with $|\hat{v}|^2/2 = q - d/d\lambda\,\breve{F}(\lambda, \hat{m})\Big|_{1-}$.

With $\tilde{\lambda}$ representing the supremum of λ such that $P_1(\lambda)$ exists and $D(\lambda)$ is stable, one has that $\tilde{\lambda} \le \overline{\lambda}$. Pham [34] shows that, in the one–dimensional case, under broad assumptions, $\tilde{\lambda} = \overline{\lambda}$ and $F(\lambda)$ is differentiable on $(-\infty, \overline{\lambda})$, both cases that $\overline{\lambda} < 1$ and $\overline{\lambda} = 1$ being realisable. The hypotheses in Pham [34], however, rule out the possibility that $\hat{\lambda} = 1$. In the appendix, we complete the analysis of Pham [34] so that the case where $\hat{\lambda} = 1$ is realised too.

For the case where $\alpha(x) = r(x)$ and $\beta(x) = 0$, the control in (2.35) appears in Theorem 2.5 in Nagai [33], which obtains the limit in part 2 of Theorem 2.2. The stability condition is of the form $\limsup_{|x|\to\infty} \big(\theta(x) - \sigma(x)b(x)^T c(x)^{-1}(a(x) - r(x)\mathbf{1})\big)^T x/|x|^2 < 0$. Instead of condition (2.39), it is required in Nagai [33] that $\|b(x)\sigma(x)^T \nabla \hat{f}(x)\|^2_{c(x)^{-1}} - \|a(x) - r(x)\mathbf{1}\|^2_{c(x)^{-1}} < 0$, for all x (see (2.25) in Nagai [33]). We have our doubts, though, as to the proof of Theorem

2.5 in Nagai [33] being sound: the last display of the proof on p.660 doesn't seem to be substantiated in that it is not clear how the term $\check{E}\int_0^T e^{-w(X_s)}(-\chi)\,ds$ on the preceding line is tackled, $-\chi$ being a positive number, e.g., why should $\check{E}e^{-w(X_s)} < \infty$, given that $-w(x)$ grows no slower than quadratically with $|x|$? Theorem 2.4 in Nagai [33], which produces an optimal portfolio on a finite time horizon, has no such condition. Besides, additional assumptions are introduced both in Theorem 2.4 and in Theorem 2.5 in Nagai [33] (see (2.20) and (2.21) there, the latter condition being characterised as "crucial") along with the requirement that $0 < q < F'(0)$. Puhalskii and Stutzer [37] obtain the bound in (2.28) under a relaxed stability condition. Theorems 2.1 and 2.2 improve on the results in Puhalskii [35] by doing away with a certain growth requirement on $|\pi_t|$ (see (2.12) in Puhalskii [35]).

The portfolio in (2.49) generalises the one in Hata, Nagai, and Sheu [20] (see (2.39) and Theorem 2.2 there whose proof is omitted) who assume that $r_1 = 0$, $\alpha_1 = 0$, $\alpha_2 = 0$, and $\beta = 0$. For the optimality of $\hat{\pi}$, those authors, in addition to requiring that the matrix $\Theta_1 - \sigma b^T c^{-1} A_1$ be stable and that the matrix $(b\sigma^T \hat{P}_1)^T c^{-1} b\sigma^T \hat{P}_1 - (A_1 - \mathbf{1} r_1^T)^T c^{-1} (A_1 - \mathbf{1} r_1^T)$ be negative definite, need that (Θ_1, σ) be controllable and that $q < F'(0)$. Hata, Nagai, and Sheu [20] produce optimal finite–horizon portfolios as well.

Maximising the probability of outperformance for a one-dimensional affine model is studied in Pham [34], who, however, stops short of proving the asymptotic optimality of $\hat{\pi}$ and produces nearly optimal portfolios instead. Besides, the requirements in Pham [34] amount to $F(\lambda)$ being essentially smooth, the portfolio's wealth growing no faster than linearly with the economic factor (see condition in (2.5) in Pham [34]) and $\theta_2 = 0$.

Most of the results on the risk–sensitive optimisation concern the case of a negative risk–sensitivity parameter. Theorem 4.1 in Nagai [32] obtains asymptotic optimality of $\hat{\pi}^\lambda$ in part 2 of Theorem 2.3 for a nonbenchmarked setup under a number of additional conditions. The stability condition is the same as in Nagai [33]. (Unfortunately, there seem to be pieces of undefined notation in Nagai [32] such as $u(0, x; T)$.) Fleming and Sheu [17], [18], who treat an affine model, allow λ to assume either sign. Their stability condition is similar to ours. The authors prove that $F(\lambda)$ can be obtained as the limit of the optimal growth rates associated with bounded portfolios as the bound constraint is being relaxed. It is also required that λ be sufficiently small, if positive. The assertion of part 1 of Theorem 2.3 has not been available in this generality even for the affine model, Theorem 4.1 in Pham [34] tackling a case of one security.

On the other hand, it has to be mentioned that none of the cited papers requires condition (N).

3 Technical preliminaries

In this section, we lay the groundwork for the proofs of the main results. Drawing on Bonnans and Shapiro [10] (see p.14 there), we will say that function $h : \mathbb{T} \to \mathbb{R}$, with $\mathbb{T}$ representing a topological space, is inf–compact (respectively, sup–compact) if the sets $\{x \in \mathbb{T} : h(x) \le \delta\}$ (respectively, the sets $\{x \in \mathbb{T} : h(x) \ge \delta\}$) are compact for all $\delta \in \mathbb{R}$. (It is worth noting that Aubin [3] and Aubin and Ekeland [4] adopt a slightly different terminology by requiring only that the sets $\{x \in \mathbb{T} : h(x) \le \delta\}$ be relatively compact in order for h to be inf–compact. Both definitions are equivalent if h is, in addition, lower semicontinuous.)

We endow the set $\mathcal{P}$ of probability measures ν on $\mathbb{R}^l$ such that $\int_{\mathbb{R}^l}|x|^2\,\nu(dx)<\infty$ with the Kantorovich–Rubinstein distance

$$d_1(\mu,\nu)=\sup\{|\int\limits_{\mathbb{R}^l} g(x)\,\mu(dx)-\int\limits_{\mathbb{R}^l} g(x)\,\nu(dx)|:\ \frac{|g(x)-g(y)|}{|x-y|}\le 1 \text{ for all } x\ne y\}\,.$$

Convergence with respect to d_1 is equivalent to weak convergence coupled with convergence of first moments, see, e.g., Villani [41]. For $\kappa>0$, let $f_\kappa(x)=\kappa|x|^2/2$, where $\kappa>0$ and $x\in\mathbb{R}^l$, and let $\mathcal{A}_\kappa$ represent the convex hull of $\mathbb{C}_0^2$ and of the function f_κ.

Lemma 3.1. *There exist $\kappa_0>0$ and $\lambda_0>0$ such that if $\kappa\le\kappa_0$ and $\lambda\le\lambda_0$, then the functions $\int_{\mathbb{R}^l}H(x;\lambda,f_\kappa)\,\nu(dx)$ and $\inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)$ are* sup*–compact in $\nu\in\mathcal{P}$ for the Kantorovich–Rubinstein distance d_1. Furthermore, given $\tilde\lambda\le\lambda_0$, the set $\bigcup_{\lambda\in[\tilde\lambda,\lambda_0]}\{\nu\in\mathcal{P}:\ \inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)\ge\delta\}$ is relatively compact, where $\delta\in\mathbb{R}$.*

Proof. By (2.19) and (2.23), for $\lambda<1$,

$$H(x;\lambda,f_\kappa)=\frac{\kappa^2}{2}\,x^TT_\lambda(x)x+\kappa S_\lambda(x)x+R_\lambda(x)+\operatorname{tr}(\sigma(x)\sigma(x)^T)\,.$$

By (2.2), (2.18a), (2.18b), and (2.18c), as $|x|\to\infty$, if κ is small, then the dominating term in $(\kappa^2/2)\,x^TT_\lambda(x)x$ is of order $\kappa^2|x|^2$, the dominating terms in $\kappa S_\lambda(x)x$ are of orders $(\lambda/(1-\lambda))\,\kappa|x|^2$ and $-\kappa|x|^2$, and the dominating term in $R_\lambda(x)$ is of order $(\lambda/(1-\lambda))\,|x|^2$. If κ is small enough, then $-\kappa|x|^2$ dominates $\kappa^2|x|^2$. For those κ, $(\lambda/(1-\lambda))\,|x|^2$ is dominated by $-\kappa|x|^2$ if λ is small relative to κ. We conclude that, provided κ is small enough, there exist $\lambda_0\in(0,1)$, K_1, and $K_2>0$, such that

$$H(x;\lambda,f_\kappa)\le K_1-K_2|x|^2\,, \tag{3.1}$$

for all $\lambda\le\lambda_0$. Therefore, given $\delta\in\mathbb{R}$, $\sup_{\nu\in\Gamma_\delta}\int_{\mathbb{R}^l}|x|^2\,\nu(dx)<\infty$, where $\Gamma_\delta=\big\{\nu:\ \int_{\mathbb{R}^l}H(x;\lambda,f_\kappa)\,\nu(dx)\ge\delta\big\}$. In addition, by $H(x;\lambda,f_\kappa)$ being continuous in x and Fatou's lemma, $\int_{\mathbb{R}^l}H(x;\lambda,f_\kappa)\,\nu(dx)$ is an upper semicontinuous function of ν, so Γ_δ is a closed set. Thus, by Prohorov's theorem, Γ_δ is compact.

By (2.19) and (2.23), the function $H(x;\lambda,f)$ is convex in f. Therefore, if $f\in\mathcal{A}_\kappa$, then, by (2.19), (2.23), and (3.1), $H(x;\lambda,f)$ is bounded above by an affine function of x. Since $H(x;\lambda,f)$ is continuous in x, the function $\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)$ is upper semicontinuous in ν. Since $f_\kappa\in\mathcal{A}_\kappa$, we obtain that $\inf_{f\in\mathcal{A}_\kappa}\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)$ is sup–compact, provided $\lambda\le\lambda_0$. Since $\inf_{f\in\mathcal{A}_\kappa}\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)=\inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)$, the latter function is sup–compact.

An examination of the reasoning that led to (3.1) reveals that there exist $\overline{K}_1$ and $\overline{K}_2>0$ such that $H(x;\lambda,f_\kappa)\le\overline{K}_1-\overline{K}_2|x|^2$ if $\lambda\in[\tilde\lambda,\lambda_0]$. Therefore,

$$\begin{aligned}\bigcup_{\lambda\in[\tilde\lambda,\lambda_0]}\{\nu\in\mathcal{P}:\ \inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}H(x;\lambda,f)\,\nu(dx)\ge\delta\}&\subset\bigcup_{\lambda\in[\tilde\lambda,\lambda_0]}\{\nu\in\mathcal{P}:\ \int\limits_{\mathbb{R}^l}H(x;\lambda,f_\kappa)\,\nu(dx)\ge\delta\}\\&\subset\{\nu\in\mathcal{P}:\ \overline{K}_2\int\limits_{\mathbb{R}^l}|x|^2\,\nu(dx)\le\overline{K}_1-\delta\}\,.\end{aligned}$$

□

Lemma 3.2. *If $\lambda < 1$ and $\mathcal{U}_\lambda \neq \emptyset$, then, for $\nu \in \mathcal{P}$,*

$$\inf_{f\in\mathcal{U}_\lambda} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) = \inf_{f\in\mathbb{C}_0^2} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,. \tag{3.2}$$

Proof. Let η be a cut–off function, i.e., a $[0,1]$–valued smooth nonincreasing function on $\mathbb{R}_+$ such that $\eta(y)=1$ when $y\in[0,1]$ and $\eta(y)=0$ when $y\geq 2$. Let us assume, in addition, that the derivative η' does not exceed 2 in absolute value and let $R>0$. Let $\eta_R(x)=\eta(|x|/R)$. Given $\psi\in\mathbb{C}_0^2$ and $\varphi\in\mathcal{U}_\lambda$, by (2.19) and (2.23),

$$\begin{aligned} H(x;\lambda,\eta_R\psi+(1-\eta_R)\varphi) &= \frac{1}{2}\,\nabla\psi(x)^T T_\lambda(x)\nabla\psi(x)\,\eta_R(x)^2 + S_\lambda(x)\nabla\psi(x)\,\eta_R(x)\\ +\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2\psi(x)\bigr)\eta_R(x) &+\frac{1}{2}\,\nabla\varphi(x)^T T_\lambda(x)\nabla\varphi(x)\,(1-\eta_R(x))^2+S_\lambda(x)\nabla\varphi(x)\,(1-\eta_R(x))\\ &+\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2\varphi(x)\bigr)(1-\eta_R(x))+\epsilon_R(x)+R_\lambda(x)\,, \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} \epsilon_R(x) &= \frac{1}{2}\,\nabla\eta_R(x)^T T_\lambda(x)\nabla\eta_R(x)\,(\psi(x)-\varphi(x))^2+\nabla\psi(x)^T T_\lambda(x)\nabla\eta_R(x)\,(\psi(x)-\varphi(x))\eta_R(x)\\ &+\nabla\psi(x)^T T_\lambda(x)\nabla\varphi(x)\,(1-\eta_R(x))\eta_R(x)+\nabla\varphi(x)^T T_\lambda(x)\nabla\eta_R(x)\,(\psi(x)-\varphi(x))(1-\eta_R(x))\\ &+S_\lambda(x)(\psi(x)-\varphi(x))\nabla\eta_R(x)+\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\bigl((\psi(x)-\varphi(x))\nabla^2\eta_R(x)\\ &+(\nabla\psi(x)-\nabla\varphi(x))\nabla\eta_R(x)^T\bigr)\bigr)\,. \end{aligned} \tag{3.4}$$

Replacing on the righthand side of (3.3) $\eta_R(x)^2$ and $(1-\eta_R(x))^2$ with $\eta_R(x)$ and $1-\eta_R(x)$, respectively, obtains that

$$H(x;\lambda,\eta_R\psi+(1-\eta_R)\varphi) \leq \eta_R(x)H(x;\lambda,\psi)+(1-\eta_R(x))H(x;\lambda,\varphi)+\epsilon_R(x)\,. \tag{3.5}$$

Therefore,

$$\begin{aligned} \int\limits_{\mathbb{R}^l} H(x;\lambda,\eta_R\psi+(1-\eta_R)\varphi)\,\nu(dx) &\leq \int\limits_{\mathbb{R}^l}\eta_R(x)H(x;\lambda,\psi)\,\nu(dx)\\ &+\sup_{x\in\mathbb{R}^l}(H(x;\lambda,\varphi)\vee 0)\nu(x:\,|x|>R)+\int\limits_{\mathbb{R}^l}\epsilon_R(x)\,\nu(dx)\,. \end{aligned}$$

By dominated convergence, the first integral on the righthand side converges to $\int_{\mathbb{R}^l} H(x;\lambda,\psi)\,\nu(dx)$, as $R\to\infty$. Since $|\nabla\eta_R(x)|\leq 4\chi_{\{|x|\geq R\}}(x)/|x|$, $|\nabla\varphi(x)|$ is, at most, of linear growth, by φ being a member of $\mathbb{C}_\ell^1$, so that $\varphi(x)$ grows, at most, quadratically, and since $\int_{\mathbb{R}^l}|x|^2\,\nu(dx)<\infty$, by (3.4), one has that

$$\lim_{R\to\infty}\int\limits_{\mathbb{R}^l}\epsilon_R(x)\,\nu(dx)=0\,. \tag{3.6}$$

Since $\psi\eta_R + \varphi(1-\eta_R) \in \mathcal{U}_\lambda$, agreeing with φ if $|x| > 2R$,

$$\inf_{f\in\mathcal{U}_\lambda} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) \le \inf_{f\in\mathbb{C}_0^2} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,.$$

Conversely, let $\varphi \in \mathcal{U}_\lambda$ and $\psi_R(x) = \eta_R(x)\varphi(x)$. One can see that ψ_R is a $\mathbb{C}_0^2$–function. By (2.31), in analogy with (3.5) and (3.6),

$$\int\limits_{\mathbb{R}^l} H(x;\lambda,\psi_R)\,\nu(dx) \le \int\limits_{\mathbb{R}^l} \big(\eta_R(x)H(x;\lambda,\varphi) + (1-\eta_R(x))H(x;\lambda,\mathbf{0})\big)\,\nu(dx) + \hat{\epsilon}_R\,,$$

where $\lim_{R\to\infty}\hat{\epsilon}_R = 0$, with $\mathbf{0}$ representing the function that is equal to zero identically. By Fatou's lemma, $H(x;\lambda,\varphi)$ being bounded from above,

$$\limsup_{R\to\infty} \int\limits_{\mathbb{R}^l} \eta_R(x)H(x;\lambda,\varphi)\,\nu(dx) \le \int\limits_{\mathbb{R}^l} H(x;\lambda,\varphi)\,\nu(dx)\,. \tag{3.7}$$

By dominated convergence,

$$\lim_{R\to\infty} \int\limits_{\mathbb{R}^l} (1-\eta_R(x))H(x;\lambda,\mathbf{0})\,\nu(dx) = 0\,.$$

Hence,

$$\inf_{f\in\mathbb{C}_0^2} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) \le \inf_{f\in\mathcal{U}_\lambda} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,,$$

which concludes the proof of (3.2).

□

Remark 3.1. Similarly, it can be shown that, if $\lambda < 1$, then

$$\inf_{f\in\mathbb{C}_b^2} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) = \inf_{f\in\mathbb{C}_0^2} \int\limits_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,.$$

(The analogue of (3.7) holds with equality by bounded convergence.)

The following lemma appears in Puhalskii and Stutzer [37].

Lemma 3.3. *Let $\lambda < 1$ and $\nu \in \mathcal{P}$. The integrals $\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)$ are bounded below uniformly over $f \in \mathbb{C}_0^2$ if and only if ν admits density which belongs to $\hat{\mathbb{P}}$.*

Proof. We start with necessity. The reasoning follows that of Puhalskii [36], cf. Lemma 6.1, Lemma 6.4, and Theorem 6.1 there. If there exists $\kappa \in \mathbb{R}$ such that $\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) \ge \kappa$ for all $f \in \mathbb{C}_0^2$, then by (2.23), for arbitrary $\delta > 0$,

$$\delta \int\limits_{\mathbb{R}^l} \frac{1}{2}\,\mathrm{tr}\,\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\,\nu(dx) \ge \kappa - \int\limits_{\mathbb{R}^l} \breve{H}(x;\lambda,\delta\nabla f(x))\,\nu(dx)\,.$$

On letting

$$\delta=\kappa^{1/2}\Big(\int\limits_{\mathbb{R}^l}\nabla f(x)^T T_\lambda(x)\nabla f(x)\,\nu(dx)\Big)^{-1/2},$$

we obtain with the aid of (2.19) and the Cauchy–Schwarz inequality that there exists constant $K_1>0$ such that, for all $f\in\mathbb{C}_0^2$,

$$\int\limits_{\mathbb{R}^l}\operatorname{tr}\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\;\nu(dx)\le K_1\Big(\int\limits_{\mathbb{R}^l}|\nabla f(x)|^2\,\nu(dx)\Big)^{1/2}.$$

It follows that the lefthand side extends to a bounded linear functional on $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$, hence, by the Riesz representation theorem, there exists $\nabla h\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\nu(dx))$ such that

$$\int\limits_{\mathbb{R}^l}\operatorname{tr}\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\;\nu(dx)=\int\limits_{\mathbb{R}^l}\nabla h(x)^T\nabla f(x)\,\nu(dx) \tag{3.8}$$

and $\int_{\mathbb{R}^l}|\nabla h(x)|^2\nu(dx)\le K_1$. Theorem 2.1 in Bogachev, Krylov, and Röckner [8] implies that the measure $\nu(dx)$ has density $m(x)$ with respect to Lebesgue measure which belongs to $L^\xi_{\text{loc}}(\mathbb{R}^l)$ for all $\xi\in(1,l/(l-1))$. It follows that, for arbitrary open ball S in $\mathbb{R}^l$, there exists $K_2>0$ such that for all $f\in\mathbb{C}_0^2$ with support in S,

$$|\int\limits_S\operatorname{tr}\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\,m(x)\,dx|\le K_2\Big(\int\limits_S|\nabla f(x)|^{2\xi/(\xi-1)}\,dx\Big)^{(\xi-1)/(2\xi)}.$$

By Theorem 6.1 in Agmon [2], the density m belongs to $\mathbb{W}^{1,\zeta}_{\text{loc}}(S)$ for all $\zeta\in(1,2l/(2l-1))$. Furthermore, $\nabla h(x)=-\operatorname{div}\big(\sigma(x)\sigma(x)^T m(x)\big)/m(x)$ so that $\sqrt{m}\in\mathbb{W}^{1,2}(\mathbb{R}^l)$.

The sufficiency follows by (2.18a), (2.18b), (2.18c), (2.19), (2.30), (2.31), and (2.32) via integration by parts. □

Remark 3.2. Essentially, (3.8) signifies that one can integrate by parts on the lefthand side, so $m(x)$ needs to be differentiable.

Remark 3.3. By (2.29), Lemma 3.2 and Lemma 3.3, $F(\lambda)>-\infty$.

Lemma 3.4. *If $\lambda<1$ and $F(\lambda)<\infty$, then the infimum in* (2.24) *is attained at $\mathbb{C}^2$–function f^λ that satisfies the ergodic Bellman equation* (2.25) *and belongs to $\mathbb{C}^1_\ell$. In addition, the function $F(\lambda)$ is lower semicontinuous and $F(0)=0$.*

Proof. Applying the reasoning on pp.289–294 in Kaise and Sheu [23], one can see that, for arbitrary $\epsilon>0$, there exists $\mathbb{C}^2$–function f_ϵ such that, for all $x\in\mathbb{R}^l$, $H(x;\lambda,f_\epsilon)=F(\lambda)+\epsilon$. Considering that some details are omitted in Kaise and Sheu [23], we give an outline of the proof, following the lead of Ichihara [21]. As $F(\lambda)<\infty$, by (2.24), there exists function $f^{(1)}_\epsilon\in\mathbb{C}^2$ such that $H(x;\lambda,f^{(1)}_\epsilon)<F(\lambda)+\epsilon$ for all x. Given open ball S, centred at the origin, by Theorem 6.14 on p.107 in Gilbarg and Trudinger [19], there exists $\mathbb{C}^2$–solution $f^{(2)}_\epsilon$ to the linear elliptic boundary value problem $H(x;\lambda,f)-(1/2)\nabla f(x)^T T_\lambda(x)\nabla f(x)=F(\lambda)+2\epsilon$ when $x\in S$ and $f(x)=f_\kappa(x)$ when $x\in\partial S$, with ∂S standing for the boundary

of S. Therefore, $H(x;\lambda, f_\epsilon^{(2)}) > F(\lambda)+\epsilon$ in S. By Theorem 8.4 on p.302 of Chapter 4 in Ladyzhenskaya and Uraltseva [27], for any ball S' contained in S and centred at the origin, there exists $\mathbb{C}^2$–solution $f^{(3)}_{\epsilon,S'}$ to the boundary value problem $H(x;\lambda,f) = F(\lambda)+\epsilon$ in S' and $f(x)=f_\kappa(x)$ on $\partial S'$. Since $f^{(3)}_{\epsilon,S'}$ solves the boundary value problem $(1/2)\text{tr}\,(\sigma(x)\sigma(x)^T\nabla^2 f(x)) = -\breve{H}(x;\lambda,\nabla f^{(3)}_{\epsilon,S'}(x))+F(\lambda)+\epsilon$ when $x\in S'$ and $f(x)=f_\kappa(x)$ when $x\in\partial S'$, we have by Theorem 6.17 on p.109 of Gilbarg and Trudinger [19] that $f^{(3)}_{\epsilon,S'}(x)$ is thrice continuously differentiable. Letting the radius of S' (and that of S) go to infinity, we have, by p.294 in Kaise and Sheu [23], see also Proposition 3.2 in Ichihara [21], that the $f^{(3)}_{\epsilon,S'}$ converge locally uniformly and in $\mathbb{W}^{1,2}_{\text{loc}}(\mathbb{R}^l)$ to f_ϵ which is a weak solution to $H(x;\lambda,f)=F(\lambda)+\epsilon$. Furthermore, by Lemma 2.4 in Kaise and Sheu [23], the $\mathbb{W}^{1,\infty}(S'')$–norms of the $f^{(3)}_{\epsilon,S'}$ are uniformly bounded over balls S' for any fixed ball S'' contained in the S'. Therefore, f_ϵ belongs to $\mathbb{W}^{1,\infty}_{\text{loc}}(\mathbb{R}^l)$. By Theorem 6.4 on p.284 in Ladyzhenskaya and Uraltseva [27], f_ϵ is thrice continuously differentiable.

As in Theorem 4.2 in Kaise and Sheu [23], by using the gradient bound in Lemma 2.4 there (which proof does require f_ϵ to be thrice continuously differentiable), we have that the f_ϵ converge along a subsequence uniformly on compact sets as $\epsilon\to 0$ to a $\mathbb{C}^2$–solution of $H(x;\lambda,f)=F(\lambda)$. That solution, which we denote by f^λ, delivers the infimum in (2.24) and satisfies the ergodic Bellman equation, with $\nabla f^\lambda(x)$ being, at most, of linear growth, see Remark 2.5 in Kaise and Sheu [23].

We prove that F is a lower semicontinuous function. Let $\lambda_i\to\lambda<1$, and the $F(\lambda_i)$ converge to a finite quantity, as $i\to\infty$. By the part just proved, there exist $\tilde f_i\in\mathbb{C}^2$ such that $H(x;\lambda_i,\tilde f_i)=F(\lambda_i)$, for all x. Furthermore, by a similar reasoning to the one used above, the sequence $\tilde f_i$ is relatively compact in $\mathbb{L}^\infty_{\text{loc}}(\mathbb{R}^l)\cap\mathbb{W}^{1,2}_{\text{loc}}(\mathbb{R}^l)$ with limit points being in $\mathbb{W}^{1,\infty}_{\text{loc}}(\mathbb{R}^l)$ as well. A subsequential limit $\tilde f$ is a $\mathbb{C}^2$-function such that $H(x;\lambda,\tilde f)=\lim_{i\to\infty}F(\lambda_i)$. By (2.24), $\lim_{i\to\infty}F(\lambda_i)\ge F(\lambda)$.

We prove that $F(0)=0$. Taking $f(x)=0$ in (2.24) yields $F(0)\le 0$. Suppose that $F(0)<0$ and let $f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell$ be such that, for all $x\in\mathbb{R}^l$,

$$\nabla f(x)^T\theta(x)+\frac{1}{2}\,|\sigma(x)^T\nabla f(x)|^2+\frac{1}{2}\,\text{tr}\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)<0\,. \tag{3.9}$$

By (2.2), there exists density $m\in\hat{\mathbb{P}}$ such that

$$\int_{\mathbb{R}^l}\big(\nabla h(x)^T\theta(x)+\frac{1}{2}\,\text{tr}\,\big(\sigma(x)\sigma(x)^T\nabla^2 h(x)\big)\big)\,m(x)\,dx=0\,, \tag{3.10}$$

for all $h\in\mathbb{C}^2_0$, see, e.g., Corollary 1.4.2 in Bogachev, Krylov, and Rëckner [9]. By (3.9), $\int_{\mathbb{R}^l}\big(\nabla f(x)^T\theta(x)+(1/2)\text{tr}\,\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\big)\,m(x)\,dx$ is well defined, being possibly equal to $-\infty$ and, by monotone convergence,

$$\int_{\mathbb{R}^l}\big(\nabla f(x)^T\theta(x)+\frac{1}{2}\,\text{tr}\,\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\big)\,m(x)\,dx$$
$$=\lim_{R\to\infty}\int_{x\in\mathbb{R}^l:\,|x|\le R}\big(\nabla f(x)^T\theta(x)+\frac{1}{2}\,\text{tr}\,\big(\sigma(x)\sigma(x)^T\nabla^2 f(x)\big)\big)\,m(x)\,dx\,.$$

By integration by parts,

$$\int_{x\in\mathbb{R}^l:\,|x|\le R} \bigl(\nabla f(x)^T\,\theta(x)+\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr)\bigr)\,m(x)\,dx$$

$$=\int_{x\in\mathbb{R}^l:\,|x|\le R} \bigl(\nabla f(x)^T\,\theta(x)-\frac{1}{2}\,\nabla f(x)^T\,\frac{\mathrm{div}\,\bigl(\sigma(x)\sigma(x)^T m(x)\bigr)}{m(x)}\bigr)\,m(x)\,dx$$

$$+\frac{1}{2}\int_{x\in\mathbb{R}^l:\,|x|=R} \nabla f(x)^T\sigma(x)\sigma(x)^T d(x)m(x)\,d\tau,$$

with $d(x)$ denoting the unit outward normal to the sphere $\{x\in\mathbb{R}^l : |x|=R\}$ at point x and with the latter integral being a surface integral. As $\int_{\mathbb{R}^l}|\nabla f(x)|m(x)\,dx<\infty$,

$$\liminf_{R\to\infty}\int_{x\in\mathbb{R}^l:\,|x|=R} |\nabla f(x)^T\sigma(x)\sigma(x)^T d(x)|m(x)\,d\tau=0\,,$$

so letting $R\to\infty$ appropriately yields the identity

$$\int_{\mathbb{R}^l}\bigl(\nabla f(x)^T\,\theta(x)+\frac{1}{2}\,\mathrm{tr}\,\bigl(\sigma(x)\sigma(x)^T\nabla^2 f(x)\bigr)\bigr)\,m(x)\,dx$$

$$=\int_{\mathbb{R}^l}\bigl(\nabla f(x)^T\,\theta(x)-\frac{1}{2}\,\nabla f(x)^T\,\frac{\mathrm{div}\,\bigl(\sigma(x)\sigma(x)^T m(x)\bigr)}{m(x)}\bigr)\,m(x)\,dx\,,\quad (3.11)$$

implying that the lefthand side is finite. A similar integration by parts in (3.10) yields

$$\int_{\mathbb{R}^l}\bigl(\nabla h(x)^T\,\theta(x)-\frac{1}{2}\,\nabla h(x)^T\,\frac{\mathrm{div}\,\bigl(\sigma(x)\sigma(x)^T m(x)\bigr)}{m(x)}\bigr)\,m(x)\,dx=0\,.$$

Since $m\in\hat{\mathbb{P}}$, this identity extends to $h\in\mathbb{C}^2\cap\mathbb{C}^1_\ell$, so the lefthand side of (3.11) equals zero, which contradicts (3.9). Thus, $F(0)=0$.

□

Remark 3.4. As a byproduct, for $\lambda<1$,

$$\inf_{f\in\mathbb{C}^2}\sup_{x\in\mathbb{R}^l}H(x;\lambda,f)=\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\sup_{x\in\mathbb{R}^l}H(x;\lambda,f)\,.$$

Remark 3.5. Since one can write (3.9) as $H(x;0,f)<0$, a similar line of reasoning to the one used for (3.11) yields the identity $G(\lambda;f,m)=\breve{G}(\lambda;\nabla f,m)$, the lefthand side being well defined, provided $\lambda<1$ and $f\in\mathcal{U}_\lambda$.

Lemma 3.5. *1. The function $\breve{H}(x,\lambda,p)$ is strictly convex in (λ,p) on $(-\infty,1)\times\mathbb{R}^l$ and is convex on $\mathbb{R}\times\mathbb{R}^l$. The function $H(x;\lambda,f)$ is convex in (λ,f) on $\mathbb{R}\times\mathbb{C}^2$. For $m\in\mathbb{P}$, the function $G(\lambda,f,m)$ is convex in (λ,f) on $\mathbb{R}\times\mathbb{C}^2_b$.*

2. *Let $m \in \hat{\mathbb{P}}$. Then the function $\breve{G}(\lambda,\nabla f,m)$ is convex and lower semicontinuous in $(\lambda,\nabla f)$ on $\mathbb{R}\times\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ and is strictly convex on $(-\infty,1)\times\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$. If $\lambda<1$, then the infimum in (2.33) is attained at unique ∇f. If $\lambda=1$ and the infimum in (2.33) is finite, then it is attained at unique ∇f too. The function $\breve{F}(\lambda,m)$ is convex and lower semicontinuous with respect to λ, it is strictly convex on $(-\infty,1)$, and tends to ∞ superlinearly, as $\lambda\to-\infty$. If $\lambda<1$, then*

$$\breve{F}(\lambda,m)=\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\breve{G}(\lambda,\nabla f,m)=\inf_{f\in\mathbb{C}^2_0}G(\lambda,f,m)\,. \tag{3.12}$$

If $\lambda<1$ and $\mathcal{U}_\lambda\neq\emptyset$, then

$$\breve{F}(\lambda,m)=\inf_{f\in\mathcal{U}_\lambda}\breve{G}(\lambda,\nabla f,m)=\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,m)\,. \tag{3.13}$$

If $\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$, then $\breve{G}(\lambda,\nabla f,m)$ is differentiable in $\lambda\in(-\infty,1)$ and

$$\frac{d}{d\lambda}\breve{G}(\lambda,\nabla f,m)=\int_{\mathbb{R}^l}\big(M(u^{\lambda,\nabla f}(x),x)+\lambda|N(u^{\lambda,\nabla f}(x),x)|^2 + \nabla f(x)^T\sigma(x)N(u^{\lambda,\nabla f}(x),x)\big)m(x)\,dx\,, \tag{3.14}$$

where $u^{\lambda,\nabla f}(x)$ is defined by (2.16) with $\nabla f(x)$ as p. Furthermore, $\breve{F}(\lambda,m)$ is differentiable with respect to λ and

$$\frac{d}{d\lambda}\breve{F}(\lambda,m)=\frac{d}{d\lambda}\breve{G}(\lambda,\nabla f^{\lambda,m},m)\,, \tag{3.15}$$

with $\nabla f^{\lambda,m}$ attaining the infimum on the righthand side of (2.33). In addition, if $\breve{F}(1,m)<\infty$, then the lefthand derivatives at 1 equal each other as well:

$$\frac{d}{d\lambda}\breve{F}(\lambda,m)\big|_{1-}=\frac{d}{d\lambda}\breve{G}(\lambda,\nabla f^{1,m},m)\big|_{1-}\,. \tag{3.16}$$

3. *The function $F(\lambda)$ is convex, is continuous for $\lambda<\overline{\lambda}$, and $F(\lambda)\to\infty$ superlinearly, as $\lambda\to-\infty$. The functions J^o_q and J^s_q are continuous.*

Proof. If $\lambda<1$, then, by (2.15), (2.17), and (2.19), the Hessian matrix of $\breve{H}(x;\lambda,p)$ with respect to (λ,p) is given by

$$\begin{aligned}
\breve{H}_{pp}(x;\lambda,p)&=T_\lambda(x)\,,\\
\breve{H}_{\lambda\lambda}(x;\lambda,p)&=\frac{1}{(1-\lambda)^3}\,\|a(x)-r(x)\mathbf{1}+b(x)\sigma(x)^Tp-b(x)\beta(x)\|^2_{c(x)^{-1}}+\beta(x)^TQ_1(x)\beta(x)\,,\\
\breve{H}_{\lambda p}(x;\lambda,p)&=-\frac{1}{(1-\lambda)^2}\,\big(a(x)-r(x)\mathbf{1}+b(x)\sigma(x)^Tp-b(x)\beta(x)\big)^Tc(x)^{-1}b(x)\sigma(x)^T\\
&\quad+\beta(x)^TQ_1(x)\sigma(x)^T\,.
\end{aligned}$$

We show that it is positive definite. More specifically, we prove that for all $\tau \in \mathbb{R}$ and $y \in \mathbb{R}^l$ such that $\tau^2 + |y|^2 \neq 0$,

$$\tau^2 \breve{H}_{\lambda\lambda}(x;\lambda,p) + y^T T_\lambda(x) y + 2\tau \breve{H}_{\lambda p}(x;\lambda,p) y > 0 .$$

Since $T_\lambda(x)$ is a positive definite matrix by condition (N), the latter inequality holds when $\tau = 0$. Assuming $\tau \neq 0$, we need to show that

$$\breve{H}_{\lambda\lambda}(x;\lambda,p) + y^T T_\lambda(x) y + 2 \breve{H}_{\lambda p}(x;\lambda,p) y > 0 . \tag{3.17}$$

Let, for $d_1 = (v_1(x), w_1(x))$ and $d_2 = (v_2(x), w_2(x))$, where $v_1(x) \in \mathbb{R}^n$, $w_1(x) \in \mathbb{R}^k$, $v_2(x) \in \mathbb{R}^n$, $w_2(x) \in \mathbb{R}^k$, and $x \in \mathbb{R}^l$, the inner product be defined by $d_1 \cdot d_2 = v_1(x)^T c(x)^{-1} v_2(x) + w_1(x)^T w_2(x)$. By the Cauchy–Schwarz inequality, applied to $d_1 = \big((1-\lambda)^{-3/2}(a(x) - r(x)\mathbf{1} + b(x)\sigma(x)^T p - b(x)\beta(x)), Q_1(x)\beta(x)\big)$ and $d_2 = ((1-\lambda)^{-1/2} b(x)\sigma(x)^T y, Q_1(x)\sigma(x)^T y)$, we have that $(\breve{H}_{\lambda p}(x;\lambda,p)y)^2 < y^T T_\lambda(x) y \breve{H}_{\lambda\lambda}(x;\lambda,p)$, with the inequality being strict because, by part 2 of condition (N), $Q_1(x)\beta(x)$ is not a scalar multiple of $Q_1(x)\sigma(x)^T y$. Thus, (3.17) holds, so the function $\breve{H}(x;\lambda,p)$ is strictly convex in (λ,p) on $(-\infty,1) \times \mathbb{R}^l$, for all $x \in \mathbb{R}^l$.

Since by (2.15) and (2.17), $\breve{H}(x;\lambda_n,p_n) \to \breve{H}(x;1,p) \le \infty$ as $\lambda_n \uparrow 1$ and $p_n \to p$, and $\breve{H}(x;\lambda,p) = \infty$ if $\lambda > 1$, the function $\breve{H}(x;\lambda,p)$ is convex in (λ,p) on $\mathbb{R} \times \mathbb{R}^l$. By (2.23), the function $H(x;\lambda,f)$ is convex in (λ,f) on $\mathbb{R} \times \mathbb{C}^2$. By (2.30), for any $m \in \mathbb{P}$, $G(\lambda,f,m)$ is convex in (λ,f) on $\mathbb{R} \times \mathbb{C}_b^2$.

Let $m \in \hat{\mathbb{P}}$. By (2.19), by $\breve{H}(x;\lambda,p)$ being a lower semicontinuous function of (λ,p) with values in $\mathbb{R} \cup \{+\infty\}$, by (2.32) and Fatou's lemma, $\breve{G}(\lambda,\nabla f,m)$ is lower semicontinuous in $(\lambda,\nabla f)$ on $\mathbb{R} \times \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$. By (2.32) and the strict convexity of $\breve{H}(x;\lambda,p)$, $\breve{G}(\lambda,\nabla f,m)$ is strictly convex in $(\lambda,\nabla f) \in (-\infty,1) \times \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$. By (2.19), (2.32), and by the facts that $\int_{\mathbb{R}^l}|x|^2 m(x)\,dx < \infty$ and $\int_{\mathbb{R}^l}|\nabla m(x)|^2/m(x)\,dx < \infty$, $\breve{G}(\lambda,\nabla f,m)$ tends to infinity as the $\mathbb{L}^2(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$–norm of ∇f tends to infinity, locally uniformly over λ. Hence, the infimum on the righthand side of (2.33) is attained at unique ∇f, if finite, see, e.g., Proposition 1.2 on p.35 in Ekeland and Temam [14]. (If $\lambda < 1$, then $\breve{G}(\lambda,\nabla f,m) < \infty$, for all $\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$, by (2.19) and (2.32).) Hence, the righthand side of (2.33) is strictly convex in λ on $(-\infty,1)$. (For, let $\inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \breve{G}(\lambda_i,\nabla f,m) = \breve{G}(\lambda_i,\nabla f_i,m)$, for $i = 1,2$. Then $\inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \breve{G}((\lambda_1+\lambda_2)/2,\nabla f,m) \le \breve{G}((\lambda_1+\lambda_2)/2,(\nabla f_1 + \nabla f_2)/2,m) < (\breve{G}(\lambda_1,\nabla f_1,m) + \breve{G}(\lambda_2,\nabla f_2,m))/2 = (\inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \breve{G}(\lambda_1,\nabla f,m) + \inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \breve{G}(\lambda_2,\nabla f,m))/2$.)

By a similar argument to that in Proposition 1.7 on p.14 in Aubin [3] or Proposition 5 on p.12 in Aubin and Ekeland [4], the function $\breve{F}(\lambda,m)$ is lower semicontinuous in λ. More specifically, let $\lambda_i \to \lambda$ and let $K_1 = \liminf_{i\to\infty} \breve{F}(\lambda_i,m)$. Assuming that $K_1 < \infty$, by (2.33), for all i great enough,

$$\breve{F}(\lambda_i,m) = \inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx):\, \breve{G}(\lambda_i,\nabla f,m) \le K_1+1} \breve{G}(\lambda_i,\nabla f,m) .$$

By (2.19) and (2.32), there exists K_2 such that, for all i, if $\breve{G}(\lambda_i,\nabla f,m) \le K_1 + 1$, then $\int_{\mathbb{R}^l}|\nabla f(x)|^2\,m(x)\,dx \le K_2$. The set of the latter ∇f being weakly compact in

$\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ and the function $\breve{G}(\lambda,\nabla f,m)$ being convex and lower semicontinuous in ∇f, there exist ∇f_i such that $\breve{F}(\lambda_i,m)=\breve{G}(\lambda_i,\nabla f_i,m)$. Extracting a suitable subsequence of ∇f_i that weakly converges to some $\nabla\tilde{f}$ and invoking the lower semicontinuity of $\breve{G}(\lambda,\nabla f,m)$ in $(\lambda,\nabla f)$ yields

$$\begin{aligned}\liminf_{i\to\infty}\breve{F}(\lambda_i,m)=\liminf_{i\to\infty}&\inf_{\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx):\,\breve{G}(\lambda_i,\nabla f,m)\le K_1+1}\breve{G}(\lambda_i,\nabla f,m)\\ &\ge\liminf_{i\to\infty}\inf_{\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx):\,\int_{\mathbb{R}^l}|\nabla f(x)|^2\,m(x)\,dx\le K_2}\breve{G}(\lambda_i,\nabla f,m)\\ &=\liminf_{i\to\infty}\breve{G}(\lambda_i,\nabla f_i,m)\ge\breve{G}(\lambda,\nabla\tilde{f},m)\ge\breve{F}(\lambda,m)\,.\end{aligned}$$

We have proved that the function $\breve{F}(\lambda,m)$ is lower semicontinuous in λ. It follows that the function $\sup_{m\in\hat{\mathbb{P}}}\breve{F}(\lambda,m)$ is lower semicontinuous.

Let us show that the gradients of functions from $\mathbb{C}^2\cap\mathbb{C}^1_\ell$ make up a dense subset of $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$. Let $f\in\mathbb{C}^1_\ell$ and let $\eta(y)$ represent a cut–off function, i.e., a $[0,1]$–valued smooth nonincreasing function on $\mathbb{R}_+$ such that $\eta(y)=1$ when $y\in[0,1]$ and $\eta(y)=0$ when $y\ge 2$. Let $R>0$. The function $f(x)\eta(|x|/R)$ belongs to $\mathbb{C}^1_0$. In addition,

$$\begin{aligned}\int_{\mathbb{R}^l}|\nabla f(x)-\nabla\big(f(x)\eta(\frac{|x|}{R})\big)|^2m(x)\,dx\le 2\int_{\mathbb{R}^l}|\nabla f(x)|^2\big(1-\eta(\frac{|x|}{R})\big)^2m(x)\,dx\\ +\frac{2}{R^2}\int_{\mathbb{R}^l}f(x)^2\eta'\big(\frac{|x|}{R}\big)^2m(x)\,dx\,,\end{aligned}$$

where η' stands for the derivative of η. Since $\int_{\mathbb{R}^l}|x|^2\,m(x)\,dx$ converges, the righthand side of the latter inequality tends to 0 as $R\to\infty$. Hence, $\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$. On the other hand, the gradients of $\mathbb{C}^1_0$–functions can be approximated with the gradients of $\mathbb{C}^2\cap\mathbb{C}^1_\ell$–functions in $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$, which ends the proof.

On recalling (2.33), we obtain the leftmost equality in (3.12). Similarly, since $G(\lambda,f,m)=\breve{G}(\lambda,\nabla f,m)$ when $f\in\mathbb{C}^2_0$ and the gradients of $\mathbb{C}^2_0$–functions are dense in $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$, the rightmost side of (3.12) equals the leftmost side. For (3.13), we recall Lemma 3.2 and Remark 3.5.

By (2.14), (2.18a), (2.18b), (2.18c), and (2.19),

$$\begin{aligned}\inf_{p\in\mathbb{R}^l}\big(\breve{H}(x;\lambda,p)-\frac{1}{2}\,p^T\,\frac{\operatorname{div}(\sigma(x)\sigma(x)^Tm(x))}{m(x)}\big)=-\frac{1}{2}\,\|S_\lambda(x)-\frac{1}{2}\,\frac{\operatorname{div}(\sigma(x)\sigma(x)^Tm(x))}{m(x)}\|^2_{T_\lambda(x)^{-1}}\\ +R_\lambda(x)\,,\end{aligned}$$

so, the lefthand sides are bounded below uniformly over $\lambda\le 0$ by an integrable function and

$$\lim_{\lambda\to-\infty}\frac{1}{\lambda^2}\,\inf_{p\in\mathbb{R}^l}\big(\breve{H}(x;\lambda,p)-\frac{1}{2}\,p^T\,\frac{\operatorname{div}(\sigma(x)\sigma(x)^Tm(x))}{m(x)}\big)=\frac{1}{2}\,\|\beta(x)\|^2_{Q_2(x)}\,.$$

The latter quantity being positive by the second part of condition (N) implies, by (2.33) and Fatou's lemma, that $\liminf_{\lambda\to-\infty}(1/\lambda^2)\breve{F}(\lambda,m)>0$, so,

$\liminf_{\lambda\to-\infty}(1/\lambda^2)\inf_{f\in\mathbb{C}_0^2} G(\lambda,f,m)>0$. By (2.24) and (2.30), $\liminf_{\lambda\to-\infty} F(\lambda)/\lambda^2>0$. Therefore, for all q from a bounded set, the supremum in (2.26b) can be taken over λ from the same compact set, which implies that J^{s}_q is continuous. The function J^{o}_q is continuous for a similar reason. Since $\sup_{x\in\mathbb{R}^l} H(x;\lambda,f)$ is a convex function of (λ,f) , by (2.24), $F(\lambda)$ is convex. Being finite, it is continuous for $\lambda<\overline{\lambda}$.

We prove the differentiability properties. The equality in (3.14) follows by Theorem 4.13 on p.273 in Bonnans and Shapiro [10] and dominated convergence, once we recall (2.19) and (2.32). Equation (3.15) is obtained similarly, with $\breve{G}(\cdot,\cdot,m)$ as $f(\cdot,\cdot)$, with λ as u , and with ∇f as x , respectively, in the hypotheses of Theorem 4.13 on p.273 in Bonnans and Shapiro [10]. In some more detail, $\breve{G}(\lambda,\nabla f,m)$ and $d\breve{G}(\lambda,\nabla f,m)/d\lambda$ are continuous functions of $(\lambda,\nabla f)$ by (2.15), (2.16), and (2.32). The inf–compactness condition on p.272 in Bonnans and Shapiro [10] holds because, as it has been shown in the proof of the lower semicontinuity of $\breve{F}(\lambda,m)$, the infimum on the righthand side of (2.33) can be taken over the same weakly compact subset of $\mathbb{L}^{1,2}_0(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ for all λ from a compact subset of $(-\infty,1)$. For (3.16), one can also apply the reasoning of the proof of Theorem 4.13 on p.273 in Bonnans and Shapiro [10]. Although the hypotheses of the theorem are not satisfied, the proof on pp.274,275 goes through, the key being that the function $\breve{G}(\lambda,\nabla f,m)$ tends to infinity uniformly over λ close enough to 1 on the left, as the $\mathbb{L}^2(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$–norm of ∇f tends to infinity.

□

Remark 3.6. If condition (N) is not assumed, then strict convexity in the statement has to be replaced with convexity.

Remark 3.7. If $\beta(x)=0$, then $F(\lambda)/\lambda^2$ tends to zero as $\lambda\to-\infty$. Furthermore,

$$\liminf_{\lambda\to-\infty}\frac{1}{|\lambda|}\inf_{f\in\mathbb{C}_0^2} G(\lambda,f,m)\ge-\int_{\mathbb{R}^l}(r(x)-\alpha(x))m(x)\,dx\,,$$

so that

$$\liminf_{\lambda\to-\infty}\frac{F(\lambda)}{|\lambda|}\ge\sup_{m\in\hat{\mathbb{P}}}\bigl(-\int_{\mathbb{R}^l}(r(x)-\alpha(x))m(x)\,dx\bigr)=-\inf_{x\in\mathbb{R}^l}(r(x)-\alpha(x))\,.$$

Consequently, if $\inf_{x\in\mathbb{R}^l}(r(x)-\alpha(x))<q$, then $\lambda q-F(\lambda)$ tends to $-\infty$ as $\lambda\to-\infty$, so $\sup_{\lambda\in\mathbb{R}}(\lambda q-F(\lambda))$ is attained. That might not be the case if $\inf_{x\in\mathbb{R}^l}(r(x)-\alpha(x)\ge q$. For instance, if the functions $a(x)$, $r(x)$, $\alpha(x)$, $b(x)$, and $\sigma(x)$ are constant and q is small enough, then the derivative of $\lambda q-F(\lambda)$ is positive for all $\lambda<0$. In particular, either J^{s}_q or J^{o}_q might not be continuous at $\inf_{x\in\mathbb{R}^l}(r(x)-\alpha(x))$, J^{s}_q being right continuous and J^{o}_q being left continuous regardless.

Lemma 3.6. *1. The function $\lambda q-\breve{F}(\lambda,m)$ has saddle point $(\hat{\lambda},\hat{m})$ in $(-\infty,\overline{\lambda}]\times\hat{\mathbb{P}}$, with $\hat{\lambda}$ being specified uniquely. In addition, $\hat{\lambda}q-F(\hat{\lambda})=\sup_{\lambda\in\mathbb{R}}(\lambda q-F(\lambda))$. If $\lambda\le\overline{\lambda}$, then $F(\lambda)=\sup_{m\in\hat{\mathbb{P}}}\breve{F}(\lambda,m)$.*

2. Suppose that $\hat{\lambda}<1$. Then the function $\lambda q-\breve{G}(\lambda,\nabla f,m)$, being concave in (λ,f) and convex in m , has saddle point $(\hat{\lambda},\hat{f},\hat{m})$ in $(-\infty,\overline{\lambda}]\times(\mathbb{C}^2\cap\mathbb{C}^1_\ell)\times\hat{\mathbb{P}}$, with $\nabla\hat{f}$ and

$\hat m$ being specified uniquely. Equations (2.34) *and* (2.36) *hold. The density $\hat m$ may be chosen positive, bounded and of class $\mathbb{C}^1$.*

3. *Suppose that $\hat\lambda = 1$. Then there exists unique $\nabla \hat f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l, \mathbb{R}^l, \hat m(x)\,dx)$ such that $\breve F(1,\hat m) = \breve G(1, \nabla \hat f, \hat m)$, $a(x) - r(x)\mathbf{1} - b(x)\beta(x) + b(x)\sigma(x)^T \nabla \hat f(x) = 0$ $\hat m(x)\,dx$–a.e. and*

$$\int_{\mathbb{R}^l} \bigl(\nabla h(x)^T\bigl(-\sigma(x)\beta(x)+\theta(x)+\sigma(x)\sigma(x)^T\nabla \hat f(x)\bigr)+\frac{1}{2}\,tr\bigl(\sigma(x)\sigma(x)^T\nabla^2 h(x)\bigr)\bigr)\hat m(x)\,dx = 0\,,$$

for all $h \in \mathbb{C}_0^2$ such that $b(x)\sigma(x)^T\nabla h(x) = 0$ $\hat m(x)\,dx$–a.e.

Proof. Let $\mathcal{U} = \{(\lambda, f) :\ f \in \mathcal{U}_\lambda\}$. It is a convex set by $H(x;\lambda,f)$ being convex in (λ, f). Let $\tilde q \in \mathbb{R}$. When $(\lambda, f) \in \mathcal{U}$ and $\nu \in \mathcal{P}$, the function $\lambda\tilde q - \int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)$ is well defined, being possibly equal to $+\infty$, is concave in (λ, f), is convex and lower semicontinuous in ν, and is inf–compact in ν, provided $\lambda < 0$ and $f = f_\kappa$, κ being small enough, the latter property holding by Lemma 3.1. Theorem 7 on p.319 in Aubin and Ekeland [4], whose proof applies to the case of the function $f(x,y)$ in the statement of the theorem taking values in $\mathbb{R}\cup\{+\infty\}$ yields the identity

$$\inf_{\nu\in\mathcal{P}}\sup_{(\lambda,f)\in\mathcal{U}}\bigl(\lambda\tilde q - \int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\bigr) = \sup_{(\lambda,f)\in\mathcal{U}}\inf_{\nu\in\mathcal{P}}\bigl(\lambda\tilde q - \int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\bigr)\,,$$

with the infimum on the lefthand side being attained. We denote that ν by $\hat\nu$ when $\tilde q = q$. If ν has no density with respect to Lebesgue measure that belongs to $\hat{\mathbb{P}}$, then, by Lemma 3.2 and Lemma 3.3, the supremum on the lefthand side equals $+\infty$. (We recall that if $\mathcal{U}_\lambda = \emptyset$ then $\inf_{f\in\mathcal{U}_\lambda} = \infty$.) Hence, the infimum on the lefthand side may be taken over ν with densities from $\hat{\mathbb{P}}$, in particular, it may be assumed that $\hat\nu(dx) = \hat m(x)\,dx$, where $\hat m \in \hat{\mathbb{P}}$. We thus have that

$$\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m)) = \sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f))\,. \tag{3.18}$$

By part 3 of Lemma 3.5, $F(\lambda) \to \infty$ superlinearly, as $\lambda \to -\infty$, so, the righthand side of (3.18) is finite. We have that

$$\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m)) \ge \sup_{\lambda\in\mathbb{R}}\inf_{m\in\hat{\mathbb{P}}}(\lambda\tilde q - \inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m)) \ge \sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \inf_{f\in\mathcal{U}_\lambda}\sup_{m\in\hat{\mathbb{P}}} G(\lambda,f,m))\,.$$

The latter rightmost side being equal to the rightmost side of (3.18) implies that the inequalities on the preceding are, in fact, equalities. Besides, by the definition of $F(\lambda)$, $\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\sup_{x\in\mathbb{R}^l} H(x;1,f) \ge F(1)$ and $F(\lambda)$ is continuous on the left at $\lambda = 1$. On recalling Lemma 3.4 and Remark 3.4, we obtain that

$$\sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m)) = \sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - \inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f)) = \sup_{\lambda\in\mathbb{R}}(\lambda\tilde q - F(\lambda))\,. \tag{3.19}$$

Therefore, for arbitrary $\lambda \in \mathbb{R}$ and $\tilde q \in \mathbb{R}$,

$$\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m) \ge \lambda\tilde q - \sup_{\tilde\lambda\in\mathbb{R}}(\tilde\lambda\tilde q - F(\tilde\lambda))\,. \tag{3.20}$$

Since F is a lower semicontinuous and convex function, it equals its bidual, so, taking supremum over $\tilde{q}$ in (3.20) yields the inequality $\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m) \geq F(\lambda)$. The opposite inequality being true for $\lambda<1$ by the definition of $F(\lambda)$ (see (2.24)) implies that, if $\lambda<1$, then

$$F(\lambda)=\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathcal{U}_\lambda} G(\lambda,f,m)\,.$$

Owing to Lemma 3.5, if $\lambda<\overline{\lambda}$, then

$$F(\lambda)=\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\breve{G}(\lambda,\nabla f,m)=\sup_{m\in\hat{\mathbb{P}}}\breve{F}(\lambda,m)\,. \tag{3.21}$$

By convexity and lower semicontinuity, the leftmost side equals the rightmost side for $\lambda=\overline{\lambda}$ too.

Since the infimum on the lefthand side of (3.18) is attained at $\hat{m}$ when $\tilde{q}=q$, by (3.19),

$$\sup_{\lambda\in\mathbb{R}}\bigl(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,\hat{m})\bigr)=\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\in\mathbb{R}}\bigl(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,m)\bigr)$$
$$=\sup_{\lambda\in\mathbb{R}}\inf_{m\in\hat{\mathbb{P}}}\bigl(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,m)\bigr)\,. \tag{3.22}$$

By the convexity of $\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,\hat{m})$ and of $\breve{F}(\lambda,\hat{m})$ in λ, we have that $\inf_{f\in\mathcal{U}_{\overline{\lambda}}}G(\overline{\lambda},f,\hat{m})$ and $\breve{F}(\overline{\lambda},\hat{m})$ are greater than or equal to their respective lefthand limits at $\overline{\lambda}$, so, by the fact that $\mathcal{U}_\lambda=\emptyset$ if $\lambda>\overline{\lambda}$ and part 2 of Lemma 3.5,

$$\sup_{\lambda\in\mathbb{R}}(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,\hat{m}))=\sup_{\lambda<\overline{\lambda}}(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,\hat{m}))=\sup_{\lambda<\overline{\lambda}}(\lambda q-\breve{F}(\lambda,\hat{m}))=\sup_{\lambda\leq\overline{\lambda}}(\lambda q-\breve{F}(\lambda,\hat{m}))\,.$$

Similarly,

$$\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\in\mathbb{R}}\bigl(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,m)\bigr)=\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\leq\overline{\lambda}}\bigl(\lambda q-\breve{F}(\lambda,m)\bigr)$$

and

$$\sup_{\lambda\in\mathbb{R}}\inf_{m\in\hat{\mathbb{P}}}\bigl(\lambda q-\inf_{f\in\mathcal{U}_\lambda}G(\lambda,f,m)\bigr)=\sup_{\lambda\leq\overline{\lambda}}\inf_{m\in\hat{\mathbb{P}}}\bigl(\lambda q-\breve{F}(\lambda,m)\bigr)\,,$$

so, by (3.22),

$$\sup_{\lambda\leq\overline{\lambda}}(\lambda q-\breve{F}(\lambda,\hat{m}))=\inf_{m\in\hat{\mathbb{P}}}\sup_{\lambda\leq\overline{\lambda}}\bigl(\lambda q-\breve{F}(\lambda,m)\bigr)=\sup_{\lambda\leq\overline{\lambda}}\inf_{m\in\hat{\mathbb{P}}}\bigl(\lambda q-\breve{F}(\lambda,m)\bigr)\,.$$

Since, by Lemma 3.5, $\breve{F}(\lambda,\hat{m})$ is a lower semicontinuous function of λ and $\breve{F}(\lambda,\hat{m})\to\infty$ superlinearly as $\lambda\to-\infty$, the supremum on the leftmost side is attained at some $\hat{\lambda}$. It follows that $(\hat{\lambda},\hat{m})$ is a saddle point of $\lambda q-\breve{F}(\lambda,m)$ in $(-\infty,\overline{\lambda}]\times\hat{\mathbb{P}}$. By Lemma 3.5, $\lambda q-\breve{F}(\lambda,m)$ is a strictly concave function of λ on $(-\infty,1)$ for all m, so $\hat{\lambda}$ is specified uniquely, see Proposition 1.5 on p.169 in Ekeland and Temam [14].

On recalling (3.21), we obtain that

$$\sup_{\lambda\in\mathbb{R}}(\lambda q-F(\lambda))=\sup_{\lambda\leq\overline{\lambda}}(\lambda q-F(\lambda))=\sup_{\lambda\leq\overline{\lambda}}(\lambda q-\sup_{m\in\hat{\mathbb{P}}}\breve{F}(\lambda,m))=\hat{\lambda}q-\breve{F}(\hat{\lambda},\hat{m})$$
$$=\hat{\lambda}q-\sup_{m\in\hat{\mathbb{P}}}\breve{F}(\hat{\lambda},m)=\hat{\lambda}q-F(\hat{\lambda})\,.$$

Part 1 has been proved.

Suppose that $\hat{\lambda} < 1$ and let $\hat{f} = f^{\hat{\lambda}}$, f^{λ} being defined in Lemma 3.4. Since $H(x;\hat{\lambda},\hat{f}) = F(\hat{\lambda})$ for all $x \in \mathbb{R}^l$, we have, by (2.30) and Remark 3.5, that $F(\hat{\lambda}) = G(\hat{\lambda},\hat{f},m) = \breve{G}(\hat{\lambda},\nabla\hat{f},m)$, for all $m \in \hat{\mathbb{P}}$. It follows that

$$\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\sup_{m\in\hat{\mathbb{P}}}\breve{G}(\hat{\lambda},\nabla f,m) \le \sup_{m\in\hat{\mathbb{P}}}\breve{G}(\hat{\lambda},\nabla\hat{f},m) = F(\hat{\lambda}) = \breve{G}(\hat{\lambda},\nabla\hat{f},\hat{m})\,. \tag{3.23}$$

By (3.21), the latter inequality is actually equality, so, $(\hat{f},\hat{m})$ is a saddle point of $\breve{G}(\hat{\lambda},\nabla f,m)$ in $(\mathbb{C}^2\cap\mathbb{C}^1_\ell)\times\hat{\mathbb{P}}$, see, e.g., Proposition 2.156 on p.104 in Bonnans and Shapiro [10] or Proposition 1.2 on p.167 in Ekeland and Temam [14]. As a result,

$$\inf_{f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell}\breve{G}(\hat{\lambda},\nabla f,\hat{m}) = \breve{G}(\hat{\lambda},\nabla\hat{f},\hat{m})\,. \tag{3.24}$$

By (2.32), (2.33) and the gradients of the functions from $\mathbb{C}^2\cap\mathbb{C}^1_\ell$ being dense in $\mathbb{L}^{1,2}_0(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$, the lefthand side of (3.24) equals $\breve{F}(\lambda,\hat{m})$, so, the infimum on the righthand side of (2.33) for $m=\hat{m}$ is attained at the gradient of the $\mathbb{C}^2\cap\mathbb{C}^1_\ell$–function $\hat{f}$.

The following reasoning shows that $(\hat{\lambda},\hat{f},\hat{m})$ is a saddle point of $\lambda q - \breve{G}(\lambda,\nabla f,m)$ in $(-\infty,\overline{\lambda}]\times(\mathbb{C}^2\cap\mathbb{C}^1_\ell)\times\hat{\mathbb{P}}$. Let $\lambda\le\overline{\lambda}$, $f\in\mathbb{C}^2\cap\mathbb{C}^1_\ell$, and $m\in\hat{\mathbb{P}}$. Since $\breve{G}(\hat{\lambda},\nabla\hat{f},\hat{m})\ge\breve{G}(\hat{\lambda},\nabla\hat{f},m)$ by $(\hat{f},\hat{m})$ being a saddle point of $\breve{G}(\hat{\lambda},\nabla f,m)$, we have that

$$\hat{\lambda}q - \breve{G}(\hat{\lambda},\nabla\hat{f},\hat{m}) \le \hat{\lambda}q - \breve{G}(\hat{\lambda},\nabla\hat{f},m)\,. \tag{3.25}$$

By (3.23), by $(\hat{\lambda},\hat{m})$ being a saddle point of $\lambda q - \breve{F}(\lambda,m)$, and by (2.33),

$$\hat{\lambda}q - \breve{G}(\hat{\lambda},\nabla\hat{f},\hat{m}) = \hat{\lambda}q - \breve{F}(\hat{\lambda},\hat{m}) \ge \lambda q - \breve{F}(\lambda,\hat{m}) \ge \lambda q - \breve{G}(\lambda,\nabla f,\hat{m})\,. \tag{3.26}$$

Putting together (3.25) and (3.26) yields the required property.

Since $(\hat{\lambda},\hat{f},\hat{m})$ is a saddle point of $\lambda q - \breve{G}(\lambda,\nabla f,m)$ in $(-\infty,\overline{\lambda}]\times(\mathbb{C}^2\cap\mathbb{C}^1_\ell)\times\hat{\mathbb{P}}$ and $\lambda q - \breve{G}(\lambda,\nabla f,m)$ is strictly concave in $(\lambda,\nabla f)$ for all m by Lemma 3.5, the pair $(\hat{\lambda},\nabla\hat{f})$ is specified uniquely, see Proposition 1.5 on p.169 of Ekeland and Temam [14]. Equation (2.34) follows by Lemma 3.4. Since $\hat{f}$ is a stationary point of $\breve{G}(\hat{\lambda},\nabla f,\hat{m})$, the directional derivatives of $\breve{G}(\hat{\lambda},\nabla f,\hat{m})$ at $\hat{f}$ are equal to zero, cf. Proposition 1.6 on p.169 in Ekeland and Temam [14]. By (2.32),

$$\int\limits_{\mathbb{R}^l}\Bigl(\breve{H}_p(x;\hat{\lambda},\nabla\hat{f}(x)) - \frac{1}{2}\,\frac{\bigl(\operatorname{div}(\sigma(x)\sigma(x)^T\,\hat{m}(x))\bigr)^T}{\hat{m}(x)}\Bigr)\nabla h(x)\,\hat{m}(x)\,dx = 0\,, \tag{3.27}$$

for all $h\in\mathbb{C}^2_0$. Integration by parts yields (2.36). In more detail, by Theorem 4.17 on p.276 in Bonnans and Shapiro [10], if $\lambda<1$, then the function $\sup_{u\in\mathbb{R}^n}\bigl(M(u,x)+\lambda|N(u,x)|^2/2 + p^T\sigma(x)N(u,x)\bigr)$, with the supremum being attained at unique point $\tilde{u}(x)$, has a derivative with respect to p given by $(\sigma(x)N(\tilde{u}(x),x))^T$, which, when combined with (2.15), (2.16) and (2.32), yields (3.27). According to Example 1.7.11 (or Example 1.7.14) in Bogachev, Krylov, and Röckner [9], $\hat{m}$ is specified uniquely by (2.36). The function $\hat{m}(x)$ is positive, bounded and is of class $\mathbb{C}^1$ by Corollaries 2.10 and 2.11 in Bogachev, Krylov, and Röckner [8] and by Agmon [2], see also Theorem 4.1(ii) and p.413 in Metafune, Pallardi, and Rhandi [31], and Proposition 1.2.18 in Bogachev, Krylov, and Röckner [9]. Part 2 has been proved.

If $\hat{\lambda}=1$, then $\breve{F}(1,\hat{m})<\infty$. By part 2 of Lemma 3.5, $\nabla\hat{f}$ exists and is specified uniquely. The other properties in part 3 follow by (2.21), (2.22), and (2.32). □

Remark 3.8. By the proof of Lemma 3.1, if $\hat{\lambda} \le \lambda_0$, then $H(x;\hat{\lambda}, f_\kappa) \to -\infty$ as $|x| \to \infty$, where $\kappa > 0$ and is small enough. In that case, the theory in Keise and Sheu [23] and Ichihara [21] yields an alternative approach to the existence of solution $\hat{m}$ to (2.34). If $\hat{\lambda} > 0$, however, those results do not seem to apply.

Remark 3.9. If the suprema in (3.21) were attained, then $F(\lambda)$ would be strictly convex.

Lemma 3.7. *Suppose that* $\lambda < \lambda_0$. *Then there exists* $m^\lambda \in \hat{\mathbb{P}}$ *such that* (f^λ, m^λ) *is a saddle point of* $\breve{G}(\lambda, \nabla f, m)$ *as a function of* (f, m) *in* $(\mathbb{C}^1_\ell \cap \mathbb{C}^2) \times \hat{\mathbb{P}}$, *so,*

$$\inf_{f\in\mathbb{C}^1_\ell\cap\mathbb{C}^2}\sup_{m\in\hat{\mathbb{P}}}\breve{G}(\lambda,\nabla f,m) = \sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathbb{C}^1_\ell\cap\mathbb{C}^2}\breve{G}(\lambda,\nabla f,m) = F(\lambda)\,, \tag{3.28}$$

with the infimum and the supremum being attained at f^λ *and* m^λ, *respectively. The density* $m^\lambda(x)$ *is the invariant density of a diffusion:*

$$\int_{\mathbb{R}^l} \bigl(\nabla h(x)^T(\lambda\sigma(x)N(u^\lambda(x),x) + \theta(x) + \sigma(x)\sigma(x)^T\nabla f^\lambda(x)) + \frac{1}{2}\,tr\,(\sigma(x)\sigma(x)^T\,\nabla^2 h(x))\bigr)\,m^\lambda(x)\,dx = 0\,, \tag{3.29}$$

for all $h \in \mathbb{C}^2_0$, *where*

$$u^\lambda(x) = \frac{1}{1-\lambda}\,c(x)^{-1}\bigl(a(x) - r(x)\mathbf{1} - \lambda b(x)\beta(x) + b(x)\sigma(x)^T\nabla f^\lambda(x)\bigr)\,.$$

The density $m^\lambda(x)$ *may be chosen positive, bounded and of class* $\mathbb{C}^1$. *The functions* $\nabla f^\lambda(x)$ *and* $m^\lambda(x)$ *are specified uniquely.*

In addition, the function $F(\lambda)$ *is strictly convex and continuously differentiable and*

$$\frac{d}{d\lambda}\,F(\lambda) = \int_{\mathbb{R}^l}\bigl(M(u^\lambda(x),x) + \lambda|N(u^\lambda(x),x)|^2 + \nabla f^\lambda(x)^T\sigma(x)N(u^\lambda(x),x)\bigr)m^\lambda(x)\,dx\,.$$

Proof. Let us begin by noting that, since, by (3.1) in the proof of Lemma 3.1, $H(x;\lambda,f_\kappa) \to -\infty$ as $|x| \to \infty$, provided $\kappa > 0$ and is small enough, we have that $F(\lambda) < \infty$, so f^λ is well defined by Lemma 3.4. Since $\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)$ is an upper semicontinuous and concave function of $\nu \in \mathcal{P}$, for all $f \in \mathcal{A}_\kappa$, is convex in $f \in \mathcal{A}_\kappa$, and $\int_{\mathbb{R}^l} H(x;\lambda,f_\kappa)\,\nu(dx)$ is sup–compact in ν by Lemma 3.1, an application of Theorem 7 on p.319 in Aubin and Ekeland [4] yields

$$\sup_{\nu\in\mathcal{P}}\inf_{f\in\mathbb{C}^2_b}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) = \sup_{\nu\in\mathcal{P}}\inf_{f\in\mathcal{A}_\kappa}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) = \inf_{f\in\mathcal{A}_\kappa}\sup_{\nu\in\mathcal{P}}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx) \ge \inf_{f\in\mathbb{C}^2_b}\sup_{\nu\in\mathcal{P}}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,,$$

the supremum on the leftmost side being attained at some ν^λ. It follows that

$$\sup_{\nu\in\mathcal{P}}\inf_{f\in\mathbb{C}_b^2}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)=\inf_{f\in\mathbb{C}_b^2}\sup_{\nu\in\mathcal{P}}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,.$$

By Remark 3.1 and Lemma 3.3,

$$\sup_{\nu\in\mathcal{P}}\inf_{f\in\mathbb{C}_b^2}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)=\sup_{m\in\mathbb{P}}\inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l} H(x;\lambda,f)\,m(x)\,dx$$

and $\nu^\lambda(dx)=m^\lambda(x)\,dx$, where $m^\lambda\in\hat{\mathbb{P}}$, and, by an approximation argument,

$$\sup_{\nu\in\mathcal{P}}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)=\sup_{m\in\mathbb{P}}\int_{\mathbb{R}^l} H(x;\lambda,f)\,m(x)\,dx\,.$$

We obtain that

$$\inf_{f\in\mathbb{C}_b^2}\sup_{m\in\mathbb{P}}G(\lambda,f,m)=\sup_{m\in\mathbb{P}}\inf_{f\in\mathbb{C}_0^2}G(\lambda,f,m)=\inf_{f\in\mathbb{C}_0^2}G(\lambda,f,m^\lambda)\,.$$

Therefore, on applying Lemma 3.3 and recalling (2.31),

$$\inf_{f\in\mathbb{C}_\ell^1\cap\mathbb{C}^2}\sup_{m\in\hat{\mathbb{P}}}\breve{G}(\lambda,\nabla f,m)\le\inf_{f\in\mathbb{C}_b^2}\sup_{m\in\mathbb{P}}\breve{G}(\lambda,\nabla f,m)=\inf_{f\in\mathbb{C}_b^2}\sup_{m\in\mathbb{P}}G(\lambda,f,m)$$
$$=\sup_{m\in\mathbb{P}}\inf_{f\in\mathbb{C}_0^2}G(\lambda,f,m)=\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathbb{C}_0^2}G(\lambda,f,m)=\sup_{m\in\hat{\mathbb{P}}}\inf_{f\in\mathbb{C}_\ell^1\cap\mathbb{C}^2}\breve{G}(\lambda,\nabla f,m)\,.$$

The leftmost side not being less than the rightmost side and (3.21) in the proof of Lemma 3.6 imply (3.28).

Since f^λ delivers infimum on the leftmost side of (3.28) and m^λ delivers supremum on the rightmost side, by Proposition 2.156 on p.104 in Bonnans and Shapiro [10] or by Proposition 1.2 on p.167 in Ekeland and Temam [14], the pair (f^λ,m^λ) is a saddle point of $\breve{G}(\lambda,\nabla f,m)$ as a function of (f,m). Equation (3.29) expresses the requirement of the directional derivative of $\breve{G}(\lambda,\nabla f,m)$ with respect to f in the direction h being equal to zero at (f^λ,m^λ) and is established similarly to (2.36) in the proof of Lemma 3.6.

The function ∇f^λ is specified uniquely because $\breve{G}(\lambda,\nabla f,m)$ is a strictly convex function of ∇f by Lemma 3.5, cf. Proposition 1.5 on p.169 in Ekeland and Temam [14]. By an argument of the proof of Lemma 3.6, the density m^λ is specified uniquely, is positive, bounded and is of class $\mathbb{C}^1$. Since m^λ is specified uniquely, the suprema in (3.28) are attained at unique ν which is ν^λ. Both the infimum and supremum in (3.28) being attained and the function $\hat{G}(\lambda,\nabla f,m)$ being strictly convex in $(\lambda,\nabla f)$ on $(-\infty,1)\times\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ by Lemma 3.5 imply that the function $F(\lambda)$ is strictly convex for $\lambda<\lambda_0$.

We address the differentiability of $F(\lambda)$. Given $\tilde{\lambda}<\lambda_0$, if λ is close enough to $\tilde{\lambda}$, then, by F being continuous at $\tilde{\lambda}$,

$$F(\lambda)=\sup_{\nu\in\mathcal{P}}\inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)=\sup_{\nu\in\mathcal{P}_{\tilde{\lambda}}}\inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\,,\qquad(3.30)$$

where $\mathcal{P}_{\tilde{\lambda}} = \cup_{\{\breve{\lambda}:\, |\breve{\lambda}-\tilde{\lambda}|\le \lambda_0-\tilde{\lambda}\}}\{\nu \in \mathcal{P} :\, \inf_{f\in \mathbb{C}_0^2}\int_{\mathbb{R}^l} H(x;\breve{\lambda},f)\,\nu(dx) \ge F(\tilde{\lambda})-1\}$. By Lemma 3.3, the measures from $\mathcal{P}_{\tilde{\lambda}}$ possess densities m which belong to $\hat{\mathbb{P}}$. By Lemma 3.5, for those ν, the function in the supremum on the rightmost side of (3.30) can be written as $\breve{F}(\lambda,m)$ and is differentiable in λ. It is also convex in λ and upper semicontinuous in ν. By Lemma 3.1, the set $\mathcal{P}_{\tilde{\lambda}}$ is relatively compact. In addition, $\nu(dx) = m^{\tilde{\lambda}}(x)\,dx$ is the only point at which the supremum in the middle term of (3.30) is attained for $\lambda = \tilde{\lambda}$. Theorem 3 on p.201 in Ioffe and Tihomirov [22] enables us to conclude that the rightmost side of (3.30) is differentiable in λ at $\tilde{\lambda}$, with the derivative being equal to

$$\int\limits_{\mathbb{R}^l}\bigl(M(u^{\tilde{\lambda}}(x),x)+\tilde{\lambda}|N(u^{\tilde{\lambda}}(x),x)|^2+\nabla f^{\tilde{\lambda}}(x)^T\sigma(x)N(u^{\tilde{\lambda}}(x),x)\bigr)m^{\tilde{\lambda}}(x)\,dx\,.$$

□

4 Proofs of Theorems 2.1 and 2.2

We prove Theorems 2.1 and 2.2 together by proving, firstly, the upper bounds and, afterwards, the lower bounds.

4.1 The upper bounds

This subsection contains the proofs of (2.27) and (2.40). Let us note that, by (2.5),

$$L_t^\pi = \int\limits_0^1 M(\pi_s^t,X_s^t)\,ds+\frac{1}{\sqrt{t}}\int\limits_0^1 N(\pi_s^t,X_s^t)^T\,dW_s^t = \int\limits_0^1\int\limits_{\mathbb{R}^l} M(\pi_s^t,x)\,\mu^t(ds,dx)+\frac{1}{\sqrt{t}}\int\limits_0^1 N(\pi_s^t,X_s^t)^T\,dW_s^t\,. \quad (4.1)$$

4.1.1 The proof of (2.27).

By (2.1) and Itô's lemma, for $\mathbb{C}^2$–function f,

$$f(X_t) = f(X_0)+\int\limits_0^t \nabla f(X_s)^T\theta(X_s)\,ds+\frac{1}{2}\int\limits_0^t \operatorname{tr}\bigl(\sigma(X_s)\sigma(X_s)^T\nabla^2 f(X_s)\bigr)\,ds + \int\limits_0^t \nabla f(X_s)^T\sigma(X_s)\,dW_s\,.$$

Since the process $\exp\bigl(\int_0^t(\lambda N(\pi_s,X_s)^T + \nabla f(X_s)^T\sigma(X_s))\,dW_s - (1/2)\int_0^t|\lambda N(\pi_s,X_s) + \sigma(X_s)^T\nabla f(X_s)|^2\,ds\bigr)$ is a local martingale, where $\lambda\in\mathbb{R}$, by (2.1) and (4.1),

$$\mathbf{E}\exp\bigl(t\lambda L_t^\pi + f(X_t) - f(X_0) - t\int_0^1 \lambda M(\pi_s^t,X_s^t)\,ds - t\int_0^1 \nabla f(X_s^t)^T\,\theta(X_s^t)\,ds$$
$$-\frac{t}{2}\int_0^1 \operatorname{tr}\bigl(\sigma(X_s^t)\sigma(X_s^t)^T\,\nabla^2 f(X_s^t)\bigr)\,ds - \frac{t}{2}\int_0^1 |\lambda N(\pi_s^t,X_s^t) + \sigma(X_s^t)^T\nabla f(X_s^t)|^2\,ds\bigr) \le 1\,.$$

Let $\nu^t(dx) = \mu^t([0,1],dx)$. By (2.15) and (2.23), for $\lambda\in(0,1)$,

$$\mathbf{E}\exp\bigl(t\lambda L_t^\pi + f(X_t) - f(X_0) - t\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu^t(dx)\bigr) \le 1\,. \tag{4.2}$$

Consequently,

$$\mathbf{E}\chi_{\{L_t^\pi\ge q\}}\exp\bigl(t\lambda L_t^\pi + f(X_t) - f(X_0) - t\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu^t(dx)\bigr) \le 1$$

Thus,

$$\ln\mathbf{E}\chi_{\{L_t^\pi\ge q\}}e^{f(X_t)-f(X_0)} \le \sup_{\nu\in\mathcal{P}}\bigl(-\lambda qt + t\int_{\mathbb{R}^l} H(x;\lambda,f)\,\nu(dx)\bigr) = -\lambda qt + t\sup_{x\in\mathbb{R}^l} H(x;\lambda,f)\,.$$

By the reverse Hölder inequality, for arbitrary $\epsilon>0$,

$$\mathbf{E}\chi_{\{L_t^\pi\ge q\}}e^{f(X_t)-f(X_0)} \ge \mathbf{P}(L_t^\pi\ge q)^{1+\epsilon}\bigl(\mathbf{E}e^{-(f(X_t)-f(X_0))/\epsilon}\bigr)^{-\epsilon}\,,$$

so,

$$\frac{1+\epsilon}{t}\,\ln\mathbf{P}(L_t^\pi\ge q) \le -\lambda q + \sup_{x\in\mathbb{R}^l} H(x;\lambda,f) + \frac{\epsilon}{t}\,\ln\mathbf{E}e^{-(f(X_t)-f(X_0))/\epsilon}\,.$$

We may assume that $\inf_{f\in\mathbb{C}^2}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f) < \infty$. By Lemma 3.4 and the hypotheses, the latter infimum is attained at function f^λ such that $f^\lambda(x)\ge -C_1|x| - C_2$ for some positive C_1 and C_2. Since, in addition, $|X_0|$ is bounded, we have that

$$\limsup_{t\to\infty}\frac{1+\epsilon}{t}\,\ln\mathbf{P}(L_t^\pi\ge q) \le -\lambda q + \inf_{f\in\mathbb{C}^2}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f) + \limsup_{t\to\infty}\frac{\epsilon}{t}\,\ln\mathbf{E}e^{C_1|X_t|/\epsilon}\,.$$

Consequently, by $\mathbf{E}e^{C_1|X_t|/\epsilon}$ being bounded in t according to Lemma C.2 of the appendix and by ϵ being arbitrarily small,

$$\limsup_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}(L_t^\pi\ge q) \le -\bigl(\lambda q - \inf_{f\in\mathbb{C}^2}\sup_{x\in\mathbb{R}^l} H(x;\lambda,f)\bigr)$$

yielding (2.27), if one recalls (2.24), (2.26a), and F being convex so that the supremum in (2.26a) can be taken over $(0,1)$.

4.1.2 The proof of (2.40)

We draw on Koncz [25] and Nagai [32]. Since $J_q^{\mathrm{s}} = 0$ when $\hat\lambda \geq 0$, we may assume that $\hat\lambda < 0$. By Lemma 3.6, $\hat f \in \mathbb{C}^2 \cap \mathbb{C}^1_\ell$. Given $t > 0$, let us introduce a Girsanov change of measure on $(\Omega, \mathcal{F}_t, \mathbf{P})$ by

$$\frac{d\hat{\mathbf{P}}^t}{d\mathbf{P}}\Big|_{\mathcal{F}_t} = \exp\Big(\int_0^t (\hat\lambda N(\hat u(X_s), X_s) + \sigma(X_s)^T \nabla \hat f(X_s))^T \, dW_s \\ - \frac{1}{2}\int_0^t |\hat\lambda N(\hat u(X_s), X_s) + \sigma(X_s)^T \nabla \hat f(X_s)|^2 \, ds\Big) . \quad (4.3)$$

A multidimensional extension of Theorem 4.7 on p.137 in Liptser and Shiryayev [29], which is proved similarly, obtains that there exists $\gamma' > 0$ such that $\sup_{s \leq t} \mathbf{E} e^{\gamma' |X_s|^2} < \infty$. By Example 3 on pp.220,221 in Liptser and Shiryayev [29] and the linear growth condition on $\nabla \hat f(x)$, the expectation of the righthand side of (4.3) with respect to $\mathbf{P}$ equals unity. Therefore, $\hat{\mathbf{P}}^t$ is a valid probability measure and, by Lemma 6.4 on p.216 in Liptser and Shiryayev [29] and Theorem 5.1 on p.191 in Karatzas and Shreve [24], the process $(\hat W_s\,, 0 \leq s \leq t)$ is a standard Wiener process with respect to $\hat{\mathbf{P}}^t$, where

$$\hat W_s = W_s - \int_0^s \big(\hat\lambda N(\hat u(X_{\tilde s}), X_{\tilde s}) + \sigma(X_{\tilde s})^T \nabla \hat f(X_{\tilde s})\big) \, d\tilde s \, . \quad (4.4)$$

By (2.1) and Itô's lemma,

$$dX_s = \big(\theta(X_s) + \sigma(X_s)(\hat\lambda N(\hat u(X_s), X_s) + \sigma(X_s)^T \nabla \hat f(X_s))\big) \, ds + \sigma(X_s) \, d\hat W_s \quad (4.5)$$

and

$$d\hat f(X_s) = \big(\nabla \hat f(X_s)^T \big(\theta(X_s) + \sigma(X_s)(\hat\lambda N(\hat u(X_s), X_s) + \sigma(X_s)^T \nabla \hat f(X_s))\big) \\ + \frac{1}{2}\,\mathrm{tr}(\sigma(X_s)\sigma(X_s)^T \nabla^2 \hat f(X_s))\big) \, ds + \nabla \hat f(X_s)^T \sigma(X_s) \, d\hat W_s \, . \quad (4.6)$$

Let, given $\lambda \in \mathbb{R}$, $f \in \mathbb{C}^2$, and measurable $\mathbb{R}^n$–valued function $v = (v(x)\,, x \in \mathbb{R}^l)$,

$$\overline{H}(x; \lambda, f, v) = \lambda M(v(x), x) + \frac{1}{2}\,|\lambda N(v(x), x) + \sigma(x)^T \nabla f(x)|^2 + \nabla f(x)^T \theta(x) \\ + \frac{1}{2}\,\mathrm{tr}\,\big(\sigma(x)\sigma(x)^T \nabla^2 f(x)\big) . \quad (4.7)$$

By (2.15), (2.23), (2.34), and (2.35), $\overline{H}(x; \hat\lambda, \hat f, \hat u) = H(x; \hat\lambda, \hat f) = F(\hat\lambda)$. On taking into account (2.3), (2.4a), (2.4b), (2.15), (2.23), and (2.34), with $\hat\pi_s$ being the optimal u in (2.15),

$$\mathbf{E} e^{t\hat\lambda L_t^{\hat\pi}} = e^{tF(\hat\lambda)} \hat{\mathbf{E}}^t e^{\hat f(X_0) - \hat f(X_t)} \, . \quad (4.8)$$

By Itô's lemma and (4.6), noting that $|x^Ty|^2 = \mathrm{tr}(xx^Tyy^T)$,

$$e^{\hat f(X_0)-\hat f(X_t)} = 1 + \int\limits_0^t e^{\hat f(X_0)-\hat f(X_s)}\bigl(\overline{H}(X_s;\hat\lambda,\mathbf{0},\hat u) - F(\hat\lambda)\bigr)\,ds - \int\limits_0^t e^{\hat f(X_0)-\hat f(X_s)}\nabla \hat f(X_s)^T\sigma(X_s)\,d\hat W_s\,.$$

Let

$$\hat\tau_R = \inf\{t\ge 0:\ |X_t|>R\}\,,$$

where $R>0$. Since $\bigl(\int_0^{s\wedge\hat\tau_R} e^{\hat f(X_0)-\hat f(X_{\tilde s})}\nabla\hat f(X_{\tilde s})^T\sigma(X_{\tilde s})\,d\hat W_{\tilde s}\,, 0\le s\le t\bigr)$ is a martingale with respect to $\hat{\mathbf{P}}^t$,

$$\hat{\mathbf{E}}^t e^{\hat f(X_0)-\hat f(X_{t\wedge\hat\tau_R})} = 1 + \hat{\mathbf{E}}^t \int\limits_0^{t\wedge\hat\tau_R} e^{\hat f(X_0)-\hat f(X_s)}\bigl(\overline{H}(X_s;\hat\lambda,\mathbf{0},\hat u) - F(\hat\lambda)\bigr)\,ds\,.$$

Since, by (2.4a), (2.4b), and (2.35),

$$\overline{H}(x;\hat\lambda,\mathbf{0},\hat u) = -\frac{\hat\lambda}{2(1-\hat\lambda)}\bigl(\|b(x)\sigma(x)^T\nabla\hat f(x)\|^2_{c(x)^{-1}} - \|a(x)-r(x)\mathbf{1}\|^2_{c(x)^{-1}}\bigr) + \hat\lambda(r(x)-\alpha(x)+\frac{1}{2}\,|\beta(x)|^2) + \frac{1}{2}\,\lambda^2|\beta(x)|^2 + \frac{\hat\lambda}{2(1-\hat\lambda)}\,\|\hat\lambda b(x)\beta(x)\|^2_{c(x)^{-1}} - \frac{\hat\lambda}{1-\hat\lambda}\,\bigl(a(x)-r(x)\mathbf{1}\bigr)^T c(x)^{-1}b(x)\hat\lambda\beta(x)\,,$$

by (2.39), there exists $K>0$ such that $\overline{H}(x;\hat\lambda,\mathbf{0},\hat u) - F(\hat\lambda) < 0$ if $|x|>K$. Therefore,

$$\hat{\mathbf{E}}^t e^{\hat f(X_0)-\hat f(X_{t\wedge\hat\tau_R})} \le 1 + \sup_{|x|\le K} e^{2|\hat f(x)|}\sup_{|x|\le K}|\overline{H}(x;\hat\lambda,\mathbf{0},\hat u) - F(\hat\lambda)|t\,,$$

so, on letting $R\to\infty$, by Fatou's lemma,

$$\hat{\mathbf{E}}^t e^{\hat f(X_0)-\hat f(X_t)} \le 1 + \sup_{|x|\le K} e^{2|\hat f(x)|}\sup_{|x|\le K}|\overline{H}(x;\hat\lambda,\mathbf{0},\hat u) - F(\hat\lambda)|t\,,$$

which implies, by (4.8), that

$$\limsup_{t\to\infty}\frac{1}{t}\,\ln\mathbf{E}e^{t\hat\lambda L_t^{\hat\pi}} \le F(\hat\lambda)\,, \tag{4.9}$$

so, (2.40) follows by Jensen's inequality and the supremum in (2.26b) being attained at $\hat\lambda$.

4.2 The lower bounds

In this subsection, we prove (2.28) and (2.38). Let us assume that $\hat{\lambda} < \overline{\lambda}$. We prove that, if $q' > q$, then

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\pi} < q') \ge -\big(\hat{\lambda} q - G(\hat{\lambda}, \hat{f}, \hat{m})\big) \tag{4.10a}$$

and that, if $q'' < q$, then

$$\liminf_{t\to\infty} \frac{1}{t} \ln \mathbf{P}(L_t^{\hat{\pi}} > q'') \ge -\big(\hat{\lambda} q - G(\hat{\lambda}, \hat{f}, \hat{m})\big)\,. \tag{4.10b}$$

We begin with showing that

$$\hat{\lambda} q - G(\hat{\lambda}, \hat{f}, \hat{m}) = \frac{1}{2} \int\limits_{\mathbb{R}^l} |\hat{\lambda} N(\hat{u}(x), x) + \sigma(x)^T \nabla \hat{f}(x)|^2 \hat{m}(x)\, dx\,. \tag{4.11}$$

Since $(\hat{\lambda}, \hat{f}, \hat{m})$ is a saddle point of $\lambda q - \breve{G}(\lambda, \nabla f, m)$ in $(-\infty, \overline{\lambda}] \times (\mathbb{C}^2 \cap \mathbb{C}^1_\ell) \times \mathbb{P}$ by Lemma 3.6, $\hat{\lambda}$ is the point of the maximum of the strictly concave function $\lambda q - \breve{G}(\lambda, \nabla \hat{f}, \hat{m})$ on $(-\infty, \overline{\lambda}]$. Since $\hat{\lambda} < \overline{\lambda}$ and $\breve{G}(\lambda, \nabla \hat{f}, \hat{m})$ is differentiable on $(-\infty, \overline{\lambda})$, the λ–derivative of $\breve{G}(\lambda, \nabla \hat{f}, \hat{m})$ at $\hat{\lambda}$ equals zero. By (3.14) of Lemma 3.5,

$$\frac{d}{d\lambda} \breve{G}(\lambda, \nabla \hat{f}, \hat{m})\Big|_{\lambda=\hat{\lambda}} = \int\limits_{\mathbb{R}^l} \big(M(\hat{u}(x), x) + \hat{\lambda} |N(\hat{u}(x), x)|^2 + \nabla \hat{f}(x)^T \sigma(x) N(\hat{u}(x), x)\big) \hat{m}(x)\, dx\,, \tag{4.12}$$

so,

$$\int\limits_{\mathbb{R}^l} \big(M(\hat{u}(x), x) + \hat{\lambda} |N(\hat{u}(x), x)|^2 + \nabla \hat{f}(x)^T \sigma(x) N(\hat{u}(x), x)\big) \hat{m}(x)\, dx = q\,. \tag{4.13}$$

Therefore, by (2.15), (2.23), and (2.30),

$$\begin{aligned}
\hat{\lambda} q - G(\hat{\lambda}, \hat{f}, \hat{m}) &= \hat{\lambda} \int\limits_{\mathbb{R}^l} \big(M(\hat{u}(x), x) + \hat{\lambda} |N(\hat{u}(x), x)|^2 + \nabla \hat{f}(x)^T \sigma(x) N(\hat{u}(x), x)\big) \hat{m}(x)\, dx \\
&\quad - \int\limits_{\mathbb{R}^l} \big(\hat{\lambda} M(\hat{u}(x), x) + \frac{1}{2} \hat{\lambda}^2 |N(\hat{u}(x), x)|^2 + \hat{\lambda}\, \nabla \hat{f}(x)^T \sigma(x) N(\hat{u}(x), x) \\
&\quad + \frac{1}{2} |\sigma(x)^T \nabla \hat{f}(x)|^2 + \nabla \hat{f}(x)^T \theta(x) + \frac{1}{2} \operatorname{tr}\big(\sigma(x)\sigma(x)^T \nabla^2 \hat{f}(x)\,\big)\big) \hat{m}(x)\, dx \\
&= \int\limits_{\mathbb{R}^l} \frac{1}{2} \hat{\lambda}^2 |N(\hat{u}(x), x)|^2 \hat{m}(x)\, dx - \int\limits_{\mathbb{R}^l} \big(\frac{1}{2} |\sigma(x)^T \nabla \hat{f}(x)|^2 + \nabla \hat{f}(x)^T \theta(x) \\
&\qquad + \frac{1}{2} \operatorname{tr}\big(\sigma(x)\sigma(x)^T \nabla^2 \hat{f}(x)\,\big)\big) \hat{m}(x)\, dx\,.
\end{aligned} \tag{4.14}$$

Integration by parts in (2.36) combined with the facts that $|\nabla \hat{f}(x)|$ grows at most linearly with $|x|$, that $\hat{u}(x)$ is a linear function of $\nabla \hat{f}(x)$ by (2.35), that $\int_{\mathbb{R}^l} |x|^2\, \hat{m}(x)\, dx < \infty$, and that $\int_{\mathbb{R}^l}|\nabla\hat{m}(x)|^2/\hat{m}(x)\,dx<\infty$, shows that (2.36) holds with $\hat{f}(x)$ as $h(x)$. Substitution on the rightmost side of (4.14) yields (4.11).

By (2.4b) and (2.35),

$$a(x)-r(x)\mathbf{1}+b(x)(\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x))=c(x)\hat{u}(x)\,.$$

By (4.4), with $\tilde{W}^t_s=\hat{W}_{st}/\sqrt{t}$,

$$\begin{aligned}L^\pi_t&=\int_0^1 M(\pi^t_s,X^t_s)\,ds+\frac{1}{\sqrt{t}}\int_0^1 N(\pi^t_s,X^t_s)^T\,dW^t_s=\int_0^1 M(\pi^t_s,X^t_s)\,ds\\&\quad+\int_0^1 N(\pi^t_s,X^t_s)^T(\hat{\lambda}N(\hat{u}(X^t_s),X^t_s)+\sigma(X^t_s)^T\nabla\hat{f}(X^t_s))\,ds+\frac{1}{\sqrt{t}}\int_0^1 N(\pi^t_s,X^t_s)^T\,d\tilde{W}^t_s\\&=\frac{1}{t}\ln\mathcal{E}^t_1+\int_0^1 M(\hat{u}(X^t_s),X^t_s)\,ds+\int_0^1 N(\hat{u}(X^t_s),X^t_s)^T(\hat{\lambda}N(\hat{u}(X^t_s),X^t_s)+\sigma(X^t_s)^T\nabla\hat{f}(X^t_s))\,ds\\&\quad+\frac{1}{\sqrt{t}}\int_0^1 N(\hat{u}(X^t_s),X^t_s)^T\,d\tilde{W}^t_s\,,\end{aligned}\tag{4.15}$$

where $\mathcal{E}^t_s$ represents the stochastic exponential defined by

$$\mathcal{E}^t_s=\exp\bigl(\sqrt{t}\int_0^s(\pi^t_{\tilde{s}}-\hat{u}(X^t_{\tilde{s}}))^Tb(X^t_{\tilde{s}})\,d\tilde{W}^t_{\tilde{s}}-\frac{t}{2}\int_0^s\|\pi^t_{\tilde{s}}-\hat{u}(X^t_{\tilde{s}})\|^2_{c(X^t_{\tilde{s}})}d\tilde{s}\bigr)\,.$$

By (4.3) and (4.15), for $\delta>0$,

$$\begin{aligned}\mathbf{P}\bigl(L^\pi_t<q+3\delta\bigr)&=\hat{\mathbf{E}}^t\chi_{\{\int_0^1 M(\pi^t_s,X^t_s)\,ds+\frac{1}{\sqrt{t}}\int_0^1 N(\pi^t_s,X^t_s)^T\,dW^t_s<q+3\delta\}}\\&\quad\exp\bigl(-\sqrt{t}\int_0^1(\hat{\lambda}N(\hat{u}(X^t_s),X^t_s)+\sigma(X^t_s)^T\nabla\hat{f}(X^t_s))^T\,d\tilde{W}^t_s\\&\quad-\frac{t}{2}\int_0^1|\hat{\lambda}N(\hat{u}(X^t_s),X^t_s)+\sigma(X^t_s)^T\nabla\hat{f}(X^t_s)|^2\,ds\bigr)\\&\geq\hat{\mathbf{E}}^t\chi_{\left\{\frac{1}{t}\ln\mathcal{E}^t_1<\delta\right\}}\chi_{\left\{\frac{1}{\sqrt{t}}|\int_0^1 N(\hat{u}(X^t_s),X^t_s)^T\,d\tilde{W}^t_s|<\delta\right\}}\chi_{\left\{\int_{\mathbb{R}^l}M(\hat{u}(x),x)\,\nu^t(dx)\right.}\\&\quad{}_{\left.+\int_{\mathbb{R}^l}N(\hat{u}(x),x)^T(\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x))\,\nu^t(dx)<q+\delta\right\}}\end{aligned}$$

$$\chi_{\left\{\frac{1}{\sqrt{t}}|\int_0^1 (\hat\lambda N(\hat u(X_s^t),X_s^t)+\sigma(X_s^t)^T\nabla \hat f(X_s^t))^T\,d\tilde W_s^t|<\delta\right\}}$$

$$\chi_{\left\{\int_{\mathbb{R}^l}|\hat\lambda N(\hat u(x),x)+\sigma(x)^T\nabla\hat f(x)|^2\,\nu^t(dx)-\int_{\mathbb{R}^l}|\hat\lambda N(\hat u(x),x)+\sigma(x)^T\nabla\hat f(x)|^2\hat m(x)\,dx<2\delta\right\}}$$

$$\exp\bigl(-2\delta t-\frac{t}{2}\int_{\mathbb{R}^l}|\hat\lambda N(\hat u(x),x)+\sigma(x)^T\nabla\hat f(x)|^2\hat m(x)\,dx\bigr)\,. \quad (4.16)$$

We will work with the terms on the righthand side in order. Since $\hat{\mathbf{E}}^t\mathcal{E}_1^t\le 1$, Markov's inequality yields the convergence

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\bigl(\frac{1}{t}\ln\mathcal{E}_1^t<\delta\bigr)=1\,. \quad (4.17)$$

We show that if $g(x)$ is a continuous function such that $|g(x)|\le K(1+|x|^2)$, for all $x\in\mathbb{R}^l$ and some $K>0$, then

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\Bigl(|\int_{\mathbb{R}^l}g(x)\nu^t(dx)-\int_{\mathbb{R}^l}g(x)\hat m(x)\,dx|>\epsilon\Bigr)=0\,. \quad (4.18)$$

By (4.5), by $(\hat W_s\,,0\le s\le t)$ being a standard Wiener process under $\hat{\mathbf{P}}^t$, and Theorem 10.1.3 on p.251 in Stroock and Varadhan [40], the distribution of $(X_s\,,0\le s\le t)$ under $\hat{\mathbf{P}}^t$ is the same as the distribution of $(\overline{X}_s\,,0\le s\le t)$, with $(\overline{X}_s\,,s\ge 0)$ being the unique strong solution to

$$d\overline{X}_s=\bigl(\theta(\overline{X}_s)+\sigma(\overline{X}_s)(\hat\lambda N(\hat u(\overline{X}_s),\overline{X}_s)+\sigma(\overline{X}_s)^T\nabla\hat f(\overline{X}_s))\bigr)\,ds+\sigma(\overline{X}_s)\,d\overline{W}_s\,,\ \overline{X}_0=X_0\,,$$

and with $\overline{W}=(\overline{W}_s\,,s\ge 0)$ being a standard Wiener process. Assuming that $\overline{X}$ and $\overline{W}$ are defined on $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbf{P}})$, we have that

$$\hat{\mathbf{P}}^t\Bigl(|\int_{\mathbb{R}^l}g(x)\nu^t(dx)-\int_{\mathbb{R}^l}g(x)\hat m(x)\,dx|>\epsilon\Bigr)=\overline{\mathbf{P}}\Bigl(|\int_{\mathbb{R}^l}g(x)\overline{\nu}^t(dx)-\int_{\mathbb{R}^l}g(x)\hat m(x)\,dx|>\epsilon\Bigr)\,,$$

where $\overline{\nu}^t(dx)=(1/t)\int_0^t\chi_{dx}(\overline{X}_s)\,ds$. Since $\hat m(x)$ is a unique solution to (2.36), by Theorem 1.7.5 in Bogachev, Krylov, and Röckner [9], $\hat m(x)\,dx$ is a unique invariant measure of $\overline{X}$, see also Proposition 9.2 on p.239 in Ethier and Kurtz [15]. It is thus an ergodic measure. We recall that $\hat m\in\hat{\mathbb{P}}$, so $\int_{\mathbb{R}^l}|x|^2\hat m(x)\,dx<\infty$. Let P^* denote the probability measure on the space $\mathbb{C}(\mathbb{R}_+,\mathbb{R}^l)$ of continuous $\mathbb{R}^l$–valued functions equipped with the locally uniform topology that is defined by $P^*(B)=\int_{\mathbb{R}^l}P_x(B)\,\hat m(x)\,dx$, where P_x is the distribution in $\mathbb{C}(\mathbb{R}_+,\mathbb{R}^l)$ of the process $\overline{X}$ started at x. Since $\hat m(x)\,dx$ is ergodic, so is P^*, see Corollary on p.12 in Skorokhod [39]. Hence, P^*–a.s.,

$$\lim_{s\to\infty}\frac{1}{s}\int_0^s g(\tilde X_{\tilde s})\,d\tilde s=\int_{\mathbb{R}^l}g(x)\hat m(x)\,dx\,, \quad (4.19)$$

see, e.g., Theorem 3 on p.9 in Skorokhod [39], with $\tilde{X}$ representing a generic element of $\mathbb{C}(\mathbb{R}_+,\mathbb{R}^l)$. Let $\mathcal{C}$ denote the complement of the set of elements of $\mathbb{C}(\mathbb{R}_+,\mathbb{R}^l)$ such that (4.19) holds. By Lemma 3.6, $\hat{m}(x)$ is continuous and strictly positive. Since $P^*(\mathcal{C})=0$, we have that $P_x(\mathcal{C})=0$ for almost all $x\in\mathbb{R}^l$ with respect to Lebesgue measure. It follows that if X_0 has an absolutely continuous distribution $n(x)\,dx$, then $\int_{\mathbb{R}^l} P_x(\mathcal{C})n(x)\,dx=0$, which means that (4.19) holds a.s. w.r.t. $\overline{P}^*$, the latter symbol denoting the distribution of $\overline{X}$ on the space of trajectories. If the distribution of X_0 is not absolutely continuous, then the distribution of $\overline{X}_1$ is because the transition probability has a density, see pp. 220–226 in Stroock and Varadhan [40]. Hence, (4.19) holds $\overline{P}^*$–a.s. for that case too. The limit in (4.18) has been proved. (A different proof can be found in Puhalskii and Stutzer [37].)

By (2.35), the linear growth condition on $\nabla\hat{f}(x)$, and (4.18),

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\bigl(\bigl|\int_{\mathbb{R}^l}|\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x)|^2\,\nu^t(dx) - \int_{\mathbb{R}^l}|\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x)|^2\hat{m}(x)\,dx\bigr|<2\delta\bigr)=1\,. \quad (4.20)$$

Since, for $\eta>0$, by the Lenglart–Rebolledo inequality, see Theorem 3 on p.66 in Liptser and Shiryayev [30],

$$\hat{\mathbf{P}}^t\bigl(\bigl|\frac{1}{\sqrt{t}}\int_0^1(\hat{\lambda}N(\hat{u}(X_s^t),X_s^t)+\sigma(x)^T\nabla\hat{f}(X_s^t))\,d\tilde{W}_s^t\bigr|\ge\delta\bigr) \le\frac{\eta}{\delta^2}+\hat{\mathbf{P}}^t\bigl(\int_0^1|\hat{\lambda}N(\hat{u}(X_s^t),X_s^t)+\sigma(x)^T\nabla\hat{f}(X_s^t)|^2\,ds\ge\eta t\bigr)\,,$$

we conclude that

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\bigl(\frac{1}{\sqrt{t}}\bigl|\int_0^1(\hat{\lambda}N(\hat{u}(X_s^t),X_s^t)+\sigma(X_s^t)^T\nabla\hat{f}(X_s^t))\,d\tilde{W}_s^t\bigr|<\delta\bigr)=1\,. \quad (4.21)$$

Similarly,

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\bigl(\frac{1}{\sqrt{t}}\bigl|\int_0^1 N(\hat{u}(X_s^t),X_s^t)^T\,d\tilde{W}_s^t<\delta\bigr)=1\,. \quad (4.22)$$

By (4.13) and (4.18),

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\bigl(\int_{\mathbb{R}^l}\bigl(M(\hat{u}(x),x)+N(\hat{u}(x),x)^T(\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x))\bigr)\,\nu^t(dx)<q+\delta\bigr)=1\,.$$

Recalling (4.16) and (4.17) obtains that

$$\liminf_{t\to\infty}\frac{1}{t}\ln\mathbf{P}\bigl(L_t^\pi<q'\bigr)\ge-\frac{1}{2}\int_{\mathbb{R}^l}|\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x)|^2m(x)\,dx\,, \quad (4.23)$$

so, (4.10a) follows from (4.11).

In order to prove (4.10b), we note that if $\pi_s^t = \hat{u}(X_s^t)$, then $\mathcal{E}_s^t = 0$ in (4.15), so

$$\int_0^1 M(\hat{u}(X_s^t), X_s^t)\, ds + \frac{1}{\sqrt{t}} \int_0^1 N(\hat{u}(X_s^t), X_s^t)^T\, dW_s^t = \int_0^1 M(\hat{u}(X_s^t), X_s^t)\, ds$$
$$+ \int_0^1 N(\hat{u}(X_s^t), X_s^t)^T (\hat{\lambda} N(\hat{u}(X_s^t), X_s^t) + \sigma(X_s^t)^T \nabla \hat{f}(X_s^t))\, ds$$
$$+ \frac{1}{\sqrt{t}} \int_0^1 N(\hat{u}(X_s^t), X_s^t)^T\, d\tilde{W}_s^t\,.$$

On recalling (4.1), similarly to (4.16),

$$\mathbf{P}\big(L_t^{\hat{\pi}} > q - 2\delta\big) = \hat{\mathbf{E}}^t \chi \Big\{ \int_0^1 \Big(M(\hat{u}(X_s^t), X_s^t) + N(\hat{u}(X_s^t), X_s^t)^T \big(\hat{\lambda} N(\hat{u}(X_s^t), X_s^t)$$
$$+ \sigma(X_s^t)^T \nabla \hat{f}(X_s^t) \big) \Big)\, ds + \frac{1}{\sqrt{t}} \int_0^1 N(\hat{u}(X_s^t), X_s^t)^T\, d\tilde{W}_s^t > q - 2\delta \Big\}$$
$$\exp\big(-\sqrt{t} \int_0^1 (\hat{\lambda} N(\hat{u}(X_s^t), X_s^t) + \sigma(X_s^t)^T \nabla \hat{f}(X_s^t))^T\, d\tilde{W}_s^t$$
$$- \frac{t}{2} \int_0^1 |\hat{\lambda} N(\hat{u}(X_s^t), X_s^t) + \sigma(X_s^t)^T \nabla \hat{f}(X_s^t)|^2\, ds\big)$$
$$\geq \chi \Big\{ \frac{1}{\sqrt{t}} \int_0^1 N(\hat{u}(X_s^t), X_s^t)^T\, d\tilde{W}_s^t > -\delta \Big\}$$
$$\chi \Big\{ \int_{\mathbb{R}^l} \Big(M(\hat{u}(x), x) + N(\hat{u}(x), x)^T \big(\hat{\lambda} N(\hat{u}(x), x) + \sigma(x)^T \nabla \hat{f}(x) \big) \Big)\, \nu^t(dx) \geq q - \delta \Big\}$$
$$\chi \Big\{ \frac{1}{\sqrt{t}} \int_0^1 (\hat{\lambda} N(\hat{u}(X_s^t), X_s^t) + \sigma(X_s^t)^T \nabla \hat{f}(X_s^t))^T\, d\tilde{W}_s^t \leq \delta \Big\}$$
$$\chi \Big\{ \int_{\mathbb{R}^l} |\hat{\lambda} N(\hat{u}(x), x) + \sigma(x)^T \nabla \hat{f}(x)|^2\, \nu^t(dx) - \int_{\mathbb{R}^l} |\hat{\lambda} N(\hat{u}(x), x) + \sigma(x)^T \nabla \hat{f}(x)|^2 \hat{m}(x)\, dx \leq 2\delta \Big\}$$
$$\exp\big(-2\delta t - \frac{t}{2} \int_{\mathbb{R}^l} |\hat{\lambda} N(\hat{u}(x), x) + \sigma(x)^T \nabla \hat{f}(x)|^2 \hat{m}(x)\, dx\big)\,. \quad (4.24)$$

One still has (4.20), (4.21), and (4.22). By (4.13) and (4.18),

$$\lim_{t\to\infty}\hat{\mathbf{P}}^t\Big(\int_{\mathbb{R}^l}\big(M(\hat{u}(x),x)+N(\hat{u}(x),x)^T(\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x))\big)\,\nu^t(dx)>q-\delta\Big)=1\,.$$

Recalling (4.24) yields

$$\liminf_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}\big(L_t^{\hat{\pi}}>q''\big)\geq-\frac{1}{2}\int_{\mathbb{R}^l}|\hat{\lambda}N(\hat{u}(x),x)+\sigma(x)^T\nabla\hat{f}(x)|^2\hat{m}(x)\,dx\,, \tag{4.25}$$

so, (4.10b) follows from (4.11).

Reversing the roles of q and q' in (4.10a) and reversing the roles of q and q'' in (4.10b) obtain that, if $q'<q$, then

$$\liminf_{t\to\infty}\frac{1}{t}\ln\mathbf{P}(L_t^{\pi}<q)\geq-J_{q'}^{\mathrm{s}}$$

and that, if $q''>q$, then

$$\liminf_{t\to\infty}\frac{1}{t}\ln\mathbf{P}(L_t^{\hat{\pi}}>q)\geq-J_{q''}^{\mathrm{o}}\,.$$

Letting $q'\to q$ and $q''\to q$ and using the continuity of J_q^{s} and J_q^{o}, respectively, which properties hold by Lemma 3.5, prove (2.28) and (2.38), respectively, provided $\hat{\lambda}<\overline{\lambda}$.

Suppose that $\hat{\lambda}=\overline{\lambda}<1$. Let $\hat{f}=f^{\hat{\lambda}}$ be as in Lemma 3.4. Then (4.23) and (4.25) hold by a similar argument to the one above. Since $\overline{\lambda}$ maximises $\lambda q-\breve{G}(\lambda,\hat{f},\hat{m})$ over λ we have that $(d/d\lambda)\,\breve{G}(\lambda,\hat{f},\hat{m})|_{\overline{\lambda}-}\leq q$. By (4.12) still holding, we have that in (4.13) the $=$ sign has to be replaced with $\leq$. By $\overline{\lambda}$ being positive, the first $=$ sign in (4.14) needs to be replaced with $\geq$, so does the $=$ sign in (4.11). By (4.23) and (4.25), one obtains (2.28) and (2.38), respectively.

Suppose that $\hat{\lambda}=\overline{\lambda}=1$. Since $\hat{\lambda}>0$, $J_q^{\mathrm{s}}=0$ and $J_q^{\mathrm{o}}>0$, so, (2.28) is a consequence of (2.27). We now work toward (2.38). Since $\lambda=1$ maximises $\lambda q-\breve{F}(\lambda,\hat{m})$ over λ and the function $\breve{F}(\lambda,\hat{m})$ is a convex function of λ, $\breve{F}(1,\hat{m})<\infty$ and $d/d\lambda\,\breve{F}(\lambda,\hat{m})\big|_{1-}\leq q$. Let $\nabla\hat{f}$ be defined as in part 3 of Lemma 3.6, i.e., let $\inf_{\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)}\breve{G}(1,\nabla f,\hat{m})$ be attained at $\nabla\hat{f}$. By (3.16) of Lemma 3.5, $d/d\lambda\,\breve{G}(\lambda,\nabla\hat{f},\hat{m})\big|_{1-}\leq q$. By part 3 of Lemma 3.6, $\breve{G}(1,\nabla\hat{f},\hat{m})$ being finite implies that, $\hat{m}(x)\,dx$–a.e.,

$$b(x)\sigma(x)^T\nabla\hat{f}(x)=b(x)\beta(x)-a(x)+r(x)\mathbf{1}\,. \tag{4.26}$$

By (3.14) of Lemma 3.5, if $\lambda<1$, then

$$\frac{d\breve{G}(\lambda,\nabla\hat{f},\hat{m})}{d\lambda}=\int_{\mathbb{R}^l}\big(M(u^{\lambda,\nabla\hat{f}}(x),x)+\lambda|N(u^{\lambda,\nabla\hat{f}}(x),x)|^2+N(u^{\lambda,\nabla\hat{f}}(x)^T\sigma(x)^T\nabla\hat{f}(x),x)\big)\hat{m}(x)\,dx\,,$$

where $u^{\lambda,\nabla\hat f}(x)$ is defined by (2.16) with $\nabla\hat f(x)$ as p. On noting that by (4.26) the limit, as $\lambda\uparrow 1$, in (2.16) with $\nabla\hat f(x)$ as p equals $c(x)^{-1}b(x)\beta$, we have, see Theorem 24.1 on p.227 in Rockafellar [38] for the first equality below, that

$$\frac{d}{d\lambda}\breve G(\lambda,\nabla\hat f,\hat m)\big|_{1-}=\lim_{\lambda\uparrow 1}\frac{d}{d\lambda}\breve G(\lambda,\nabla\hat f,\hat m)=\int\limits_{\mathbb{R}^l}\bigl(M(c(x)^{-1}b(x)\beta(x),x)$$
$$+|N(c(x)^{-1}b(x)\beta(x),x)|^2+N(c(x)^{-1}b(x)\beta(x),x)^T\sigma(x)^T\nabla\hat f(x)\bigr)\,\hat m(x)\,dx\,.$$

We recall that $\hat v(x)$ is defined to be a bounded continuous function with values in the range of $b(x)^T$ such that $|\hat v(x)|^2/2=q-d/d\lambda\,\breve F(\lambda,\hat m)\Big|_{1-}$ and $\hat u(x)=c(x)^{-1}b(x)(\beta(x)+\hat v(x))$. By Lemma 3.5, $d/d\lambda\,\breve F(\lambda,\hat m)\big|_{1-}=d/d\lambda\,\breve G(\lambda,\nabla\hat f,\hat m)\big|_{1-}$. Since the vectors $b(x)^Tc(x)^{-1}b(x)\beta(x)-\beta(x)$ and $b(x)^Tc(x)^{-1}b(x)\hat v(x)$ are orthogonal, with the former being in the null space of $b(x)$ and the latter being in the range of $b(x)^T$, substitution in (2.4a) and (2.4b) with the account of (2.21) yields

$$\int\limits_{\mathbb{R}^l}\bigl(M(\hat u(x),x)+|N(\hat u(x),x)|^2+N(\hat u(x),x)^T\sigma(x)^T\nabla\hat f(x)\bigr)\,\hat m(x)\,dx$$
$$=\frac{d}{d\lambda}\breve G(\lambda,\nabla\hat f,\hat m)\big|_{1-}+\int\limits_{\mathbb{R}^l}\frac{|\hat v(x)|^2}{2}\,\hat m(x)\,dx=q\,. \quad (4.27)$$

(As a consequence, (4.13) holds in this case too.)

We now invoke results in Puhalskii [36]. Let the process $\hat\Psi_t=(\hat\Psi^t_s\,,s\in[0,1])$ be defined by (2.6) with $\hat u(x)$ as $u(x)$. Since $\hat u(x)$ is a bounded continuous function, the random variables $N(\hat u(X^t_s),X^t_s)$ are uniformly bounded. Condition 2.2 in Puhalskii [36] is fulfilled because part 2 of condition (N) implies that the length of the projection of $N(\hat u(x),x)$ onto the nullspace of $\sigma(x)$ is bounded away from zero and, consequently, the quantity $|N(\hat u(x),x)|^2-N(\hat u(x),x)^T\sigma(x)(\sigma(x)\sigma(x)^T)^{-1}\sigma(x)^TN(\hat u(x),x)$ is bounded away from zero. Thus, Theorem 2.1 in Puhalskii [36] applies, so the pair $(\hat\Psi^t,\mu^t)$ satisfies the LDP in $\mathbb{C}([0,1])\times\mathbb{C}_\uparrow([0,1],\mathbb{M}(\mathbb{R}^l))$ for rate t, as $t\to\infty$, with the deviation function in (2.8), provided the function $\Psi=(\Psi_s,\,s\in[0,1])$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}_+$ and the function $\mu=(\mu_s(\Gamma))$, when considered as a measure on $[0,1]\times\mathbb{R}^l$, is absolutely continuous w.r.t. Lebesgue measure, i.e., $\mu(ds,dx)=m_s(x)\,dx\,ds$, where $m_s(x)$, as a function of x, belongs to $\hat{\mathbb{P}}$ for almost all s. If those conditions do not hold then $\mathbf{J}(\Psi,\mu)=\infty$. Since $L^{\hat\pi}_t=\hat\Psi^t_1$ and $\nu^t(\Gamma)=\mu^t([0,1],\Gamma)$, by projection, the pair $(L^{\hat\pi}_t,\nu^t)$ obeys the LDP in $\mathbb{R}\times\mathbb{M}(\mathbb{R}^l)$ for rate t with deviation function $\mathbf{I}^{\hat u}$, such that $\mathbf{I}^{\hat u}(L,\nu)=\inf\{\mathbf{J}(\Psi,\mu):\ \Psi_1=L\,,\mu([0,1],\Gamma)=\nu(\Gamma)\}$. Therefore,

$$\liminf_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}\bigl(L^{\hat\pi}_t>q\bigr)\ge-\inf_{(L,\nu):\,L>q}\mathbf{I}^{\hat u}(L,\nu)\,. \quad (4.28)$$

Calculations show that

$$\mathbf{I}^{\hat u}(L,\nu)=\sup_{\lambda\in\mathbb{R}}(\lambda L-\inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat u)\,\nu(dx))\,,$$

if $\nu(dx) = m(x)\,dx$, where $m \in \hat{\mathbb{P}}$, and $\mathbf{I}^{\hat{u}}(L,\nu) = \infty$, otherwise, where $\overline{H}(x;\lambda,f,v)$ is defined in (4.7). We have that the function $\lambda L - \inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx$ is concave in λ and is convex and lower semicontinuous in L. It is sup–compact in λ because $\mathbf{I}^{\hat{u}}(L,\nu)$ is a deviation function, i.e., it is inf–compact. (We provide a direct proof of the latter property in the appendix.) Therefore, by Theorem 7 on p.319 in Aubin and Ekeland [4],

$$\inf_{(L,\nu):\,L>q}\mathbf{I}^{\hat{u}}(L,\nu) \le \inf_{L>q}\sup_{\lambda\in\mathbb{R}}(\lambda L - \inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx)$$
$$= \sup_{\lambda\in\mathbb{R}}\inf_{L>q}(\lambda L - \inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx) = \sup_{\lambda\ge 0}(\lambda q - \inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx)\,. \tag{4.29}$$

By integration by parts, if $f \in \mathbb{C}_0^2$, then

$$\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,v)\hat{m}(x)\,dx = \int\limits_{\mathbb{R}^l}\big(\lambda M(v(x),x) + \frac{1}{2}\,|\lambda N(v(x),x) + \sigma(x)^T\nabla f(x)|^2 + \nabla f(x)^T\,\theta(x)$$
$$- \frac{1}{2}\,\nabla f(x)^T\frac{\operatorname{div}\big(\sigma(x)\sigma(x)^T\hat{m}(x)\big)}{\hat{m}(x)}\big)\hat{m}(x)\,dx\,. \tag{4.30}$$

As the righthand side depends on $f(x)$ through $\nabla f(x)$ only, similarly to developments above, we use the righthand side of (4.30) in order to define the lefthand side when $\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$. By the set of the gradients of $\mathbb{C}_0^2$–functions being dense in $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$,

$$\inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx = \inf_{\nabla f\in\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\,\hat{m}(x)\,dx\,.$$

Since $\overline{H}(x;1,f,\hat{u}) = H(x;1,f)$ (see (2.20) and (4.26)), $\int_{\mathbb{R}^l}\overline{H}(x;1,f,\hat{u})\hat{m}(x)\,dx = \breve{G}(1,\nabla f,\hat{m})$. By $\nabla\hat{f}$ minimising $\breve{G}(1,\nabla f,\hat{m})$ over $\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$, the function $q - \int_{\mathbb{R}^l}\overline{H}(x;1,f,\hat{u})\hat{m}(x)\,dx$ attains maximum over ∇f in $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$ at $\nabla\hat{f}$. Therefore, the partial derivative with respect to ∇f of $\lambda q - \int_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\hat{m}(x)\,dx$ equals zero at $(1,\nabla\hat{f})$. By (4.30), we can write (4.27) as $d/d\lambda\,\int_{\mathbb{R}^l}\overline{H}(x;\lambda,\hat{f},\hat{u})\hat{m}(x)\,dx\Big|_1 = q$, so, the partial derivative with respect to λ of $\lambda q - \int_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\hat{m}(x)\,dx$ at $(1,\nabla\hat{f})$ equals zero too. The function $\lambda q - \int_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\hat{m}(x)\,dx$ being concave in $(\lambda,\nabla f)$, it therefore attains a global maximum in $\mathbb{R}\times\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,\hat{m}(x)\,dx)$ at $(1,\nabla\hat{f})$, cf. Proposition 1.2 on p.36 in Ekeland and Temam [14]. Hence,

$$\sup_{\lambda\ge 0}\big(\lambda q - \inf_{f\in\mathbb{C}_0^2}\int\limits_{\mathbb{R}^l}\overline{H}(x;\lambda,f,\hat{u})\hat{m}(x)\,dx\big) = q - \breve{G}(1,\nabla\hat{f},\hat{m})\,.$$

The latter expression being equal to J_q^{o}, (4.28), and (4.29) imply the required lower bound (2.38).

Remark 4.1. The change of the measure in (4.3) is implicit in Puhalskii [36]. The idea of using a stochastic exponential in order to "absorb" control in (4.15) is borrowed from Hata, Nagai, and Sheu [20].

5 Proof of Theorem 2.3

For the first assertion of part 1, let us assume that $\lambda < \overline{\lambda}$. Let $\epsilon > 0$ be such that $\lambda(1+\epsilon) < \overline{\lambda}$ and the function $f^{\lambda(1+\epsilon)}$ is bounded below by an affine function of x. By (2.25), denoting $f_\epsilon = f^{\lambda(1+\epsilon)}$,

$$\limsup_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} \exp((1+\epsilon)\lambda t L_t^\pi + f_\epsilon(X_t) - f_\epsilon(X_0)) \le F((1+\epsilon)\lambda)\,.$$

By the reverse Hölder inequality,

$$\mathbf{E}\exp((1+\epsilon)\lambda t L_t^\pi + f_\epsilon(X_t) - f_\epsilon(X_0)) \ge \bigl(\mathbf{E}\exp(\lambda t L_t^\pi)\bigr)^{1+\epsilon}\bigl(\mathbf{E}\exp(-(1/\epsilon)(f_\epsilon(X_t) - f_\epsilon(X_0)))\bigr)^{-\epsilon},$$

so, since f_ϵ is bounded below by an affine function, $|X_0|$ is bounded, and ϵ can be chosen arbitrarily small, in analogy with the proof of (2.27),

$$\limsup_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} \exp(\lambda t L_t^\pi) \le F(\lambda)\,.$$

The latter inequality is trivially true if $\lambda > \overline{\lambda}$.

We address the lower bound. Let $0 < \lambda < \overline{\lambda}$. Then F is subdifferentiable at λ. Let q represent a subgradient of F at λ. Since $\lambda q - F(\lambda) = J_q^{\mathrm{o}}$, by (2.38),

$$\begin{aligned}\liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^\lambda}} \ge \liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^\lambda}} \chi_{\{L_t^{\hat{\pi}^\lambda} \ge q\}} \ge \lambda q + \liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{P}(L_t^{\hat{\pi}^\lambda} \ge q) \\ \ge \lambda q - J_q^{\mathrm{o}} = F(\lambda)\,. \quad (5.1)\end{aligned}$$

If $\lambda = \overline{\lambda}$ and F is subdifferentiable at $\overline{\lambda}$, a similar proof applies. Suppose that $\lambda = \overline{\lambda}$ and F is not subdifferentiable at $\overline{\lambda}$. By what has been just proved,

$$\liminf_{\check{\lambda}\uparrow\overline{\lambda}} \liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} e^{\check{\lambda} t L_t^{\hat{\pi}^{\check{\lambda}}}} \ge \liminf_{\check{\lambda}\uparrow\overline{\lambda}} F(\check{\lambda}) = F(\overline{\lambda})$$

and Hölder's inequality yields

$$\liminf_{\check{\lambda}\uparrow\overline{\lambda}} \liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} e^{\overline{\lambda} t L_t^{\hat{\pi}^{\check{\lambda}}}} \ge F(\overline{\lambda})\,.$$

By requiring $\pi_t^{\overline{\lambda}}$ to match $\pi_t^{\lambda_i}$ on certain intervals $[t_i, t_{i+1})$ where $\lambda_i \uparrow \overline{\lambda}$ and $t_i \to \infty$ appropriately, we can ensure that $\liminf_{t\to\infty}(1/t)\, \ln \mathbf{E} e^{\overline{\lambda} t L_t^{\hat{\pi}^{\overline{\lambda}}}} \ge F(\overline{\lambda})$.

Suppose that $\lambda > \overline{\lambda}$. If F is subdifferentiable at $\overline{\lambda}$, then, similarly to (5.1), on choosing q as a subgradient of F at $\overline{\lambda}$,

$$\liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{E} e^{\lambda t L_t^{\hat{\pi}^{\overline{\lambda}}}} \ge \lambda q + \liminf_{t\to\infty} \frac{1}{t}\, \ln \mathbf{P}(L_t^{\hat{\pi}^{\overline{\lambda}}} \ge q) \ge \lambda q - J_q^{\mathrm{o}} = (\lambda - \overline{\lambda})q + F(\overline{\lambda})\,. \quad (5.2)$$

Since q can be chosen arbitrarily great, $\lim_{t\to\infty}(1/t)\,\ln\mathbf{E}e^{\lambda t L_t^{\hat{\pi}^{\overline{\lambda}}}}=\infty$. If F is not subdifferentiable at $\overline{\lambda}$, then we pick λ_i and q_i such that $\lambda_i\uparrow\overline{\lambda}$, q_i is a subgradient of F at λ_i and $q_i\uparrow\infty$. Arguing along the lines of (5.2) yields

$$\liminf_{t\to\infty}\frac{1}{t}\,\ln\mathbf{E}e^{\lambda t L_t^{\hat{\pi}^{\lambda_i}}}\ge\lambda q_i+\liminf_{t\to\infty}\frac{1}{t}\,\ln\mathbf{P}(L_t^{\hat{\pi}^{\lambda_i}}\ge q_i)\ge(\lambda-\overline{\lambda})q_i+F(\lambda_i)\,,$$

so there exists π^λ such that $\lim_{t\to\infty}(1/t)\ln\mathbf{E}e^{\lambda t L_t^{\pi^\lambda}}=\infty$.

We prove part 2. Since $\mathbf{E}e^{\lambda t L_t^\pi}\ge e^{\lambda q t}\mathbf{P}(L_t^\pi\le q)$ provided $\lambda<0$, the inequality in (2.28) of Theorem 2.1 implies that

$$\liminf_{t\to\infty}\frac{1}{t}\,\ln\mathbf{E}e^{\lambda t L_t^\pi}\ge\sup_{q\in\mathbb{R}}(\lambda q-J_q^{\mathrm{s}})=F(\lambda)\,,$$

with the latter equality holding because by (2.26b) J_q^{s} is the Legendre–Fenchel transform of the function that equals $F(\lambda)$ when $\lambda\le 0$ and equals ∞, otherwise. Since $\lambda<0$, F is differentiable at λ by Lemma 3.7, so $q=F'(\lambda)$ and $\hat{\pi}^\lambda$ is well defined. The needed upper bound is nothing but (4.9).

Acknowledgement. I am grateful to Professor Stutzer for introducing me to the subject area of this research.

A The scalar case

In this subsection the one–dimensional affine setup is analysed in more detail. We will assume that $l=n=1$, so, in (2.41a)–(2.41e), Θ_1, θ_2, A_1, a_2, r_1, r_2, α_1, and α_2 are scalars, $\Theta_1<0$, σ is a $1\times k$–matrix, b is a $1\times k$–matrix, and β is a k–vector. Accordingly, c, $\sigma\sigma^T$, σb^T, $P_1(\lambda)$, $p_2(\lambda)$, $A(\lambda)$, $B(\lambda)$, and C are scalars. By (2.43), the equation for $P_1(\lambda)$ is

$$B(\lambda)P_1(\lambda)^2+2A(\lambda)P_1(\lambda)+\frac{\lambda}{1-\lambda}\,C=0\,. \tag{A.1}$$

Let

$$\tilde{\beta}=1+\frac{1}{\Theta_1^2}\,\frac{A_1-r_1}{c}\,\bigl(\sigma\sigma^T(A_1-r_1)-2\Theta_1\sigma b^T\bigr)\,. \tag{A.2}$$

(The latter piece of notation is modelled on that of Pham [34].) One can see, by (2.42a), (2.42b), and (2.42c), that $\tilde{\beta}\ge 0$ and

$$A(\lambda)^2-B(\lambda)\,\frac{\lambda}{1-\lambda}\,C=\Theta_1^2\,\frac{1-\lambda\tilde{\beta}}{1-\lambda}\,.$$

Hence, $P_1(\lambda)$ exists if and only if $\lambda\le 1/\tilde{\beta}$ and $\lambda<1$, so, $\tilde{\lambda}=(1/\tilde{\beta})\wedge 1$. (Not unexpectedly, if $\lambda<0$ then (A.1) has both a positive and a negative solution, whereas both solutions are positive if $0<\lambda\le\tilde{\lambda}$.) If $\lambda<\tilde{\lambda}$, then

$$P_1(\lambda)=\frac{1}{B(\lambda)}\,\bigl(-A(\lambda)-|\Theta_1|\sqrt{\frac{1-\lambda\tilde{\beta}}{1-\lambda}}\bigr) \tag{A.3}$$

and $F(\lambda)$ is determined by (2.45) and (2.48). The negative signature of the square root is chosen because $D(\lambda)$ in (2.14) has to be negative which is needed in order for the analogue of (2.47) to have a stationary distribution. Therefore,

$$D(\lambda) = \Theta_1 \sqrt{\frac{1-\lambda\tilde{\beta}}{1-\lambda}}\,. \tag{A.4}$$

The functions $D(\lambda)$ and $P_1(\lambda)$ are differentiable for $\lambda < 1 \wedge (1/\tilde{\beta})$. As in Pham [34], we distinguish between three cases: $\tilde{\beta} > 1$, $\tilde{\beta} < 1$, and $\tilde{\beta} = 1$.

Suppose that $\tilde{\beta} > 1$, so, $\tilde{\lambda} = 1/\tilde{\beta}$. Then $P_1(\lambda)$ and $D(\lambda)$ are continuous on $[0, 1/\tilde{\beta}]$ and differentiable on $(0, 1/\tilde{\beta})$. We have that $P_1(1/\tilde{\beta}) = -A(1/\tilde{\beta})/B(1/\tilde{\beta})$ and $D(1/\tilde{\beta}) = 0$. Also, $D(\lambda)/\sqrt{1/\tilde{\beta}-\lambda} \to -|\Theta_1|\sqrt{\tilde{\beta}}/\sqrt{1-1/\tilde{\beta}}$ and $(P_1(1/\tilde{\beta}) - P_1(\lambda))/\sqrt{1/\tilde{\beta}-\lambda} \to |\Theta_1|\sqrt{\tilde{\beta}}/(B(1/\tilde{\beta})\sqrt{1-1/\tilde{\beta}})$, as $\lambda \uparrow 1/\tilde{\beta}$. In addition, by (2.45) and (2.48), if $E(1/\tilde{\beta}) \neq 0$, then $|p_2(\lambda)| = |E(\lambda)/D(\lambda)| \to \infty$ and $F(\lambda) \to \infty$, so, $F(\lambda) = \infty$ when $\lambda \geq 1/\tilde{\beta}$, $\overline{\lambda} = 1/\tilde{\beta}$, and $\hat{\lambda} < \overline{\lambda}$. Suppose that $E(1/\tilde{\beta}) = 0$. By (2.45) and (2.46), $E(\lambda) = D(\lambda)Z(\lambda) + U(\lambda)$, where

$$Z(\lambda) = \frac{\lambda}{1-\lambda}\, b\sigma^T c^{-1}(a_2 - r_2 - \lambda b\beta) - \lambda\sigma\beta + \theta_2$$

and

$$U(\lambda) = \frac{\lambda}{1-\lambda}\,(A_1 - r_1)c^{-1}(a_2 - r_2 - \lambda b\beta) + \lambda(r_1 - \alpha_1) - \frac{A(\lambda)}{B(\lambda)}\, Z(\lambda)\,.$$

Therefore, for $\lambda < 1/\tilde{\beta}$,

$$p_2(\lambda) = -\frac{Z(\lambda)}{B(\lambda)} - \frac{U(\lambda)}{D(\lambda)}\,.$$

Since $E(1/\tilde{\beta}) = D(1/\tilde{\beta}) = 0$, $U(1/\tilde{\beta}) = 0$. By $U(\lambda)$ being linear in a neighbourhood of $1/\tilde{\beta}$, $p_2(\lambda)$ has a finite limit at $1/\tilde{\beta}$, so, we let, by continuity, $p_2(1/\tilde{\beta}) = -Z(1/\tilde{\beta})/B(1/\tilde{\beta})$, and $F(1/\tilde{\beta})$ is finite. Let us look at the derivative at $1/\tilde{\beta}$. We have that $(p_2(1/\tilde{\beta}) - p_2(\lambda))/\sqrt{1/\tilde{\beta}-\lambda} \to U'(1/\tilde{\beta})\sqrt{1-1/\tilde{\beta}}/(|\Theta_1|\sqrt{\tilde{\beta}})$, as $\lambda \uparrow 1/\tilde{\beta}$. By (2.48), $(F(1/\tilde{\beta}) - F(\lambda))/\sqrt{1/\tilde{\beta}-\lambda} \to (1/2)\,\sigma\sigma^T|\Theta_1|\sqrt{\tilde{\beta}}/(B(1/\tilde{\beta})\,\sqrt{1-1/\tilde{\beta}})$. Therefore, $F'(1/\tilde{\beta}-) = \infty$, so, F is essentially smooth, $\overline{\lambda} = 1/\tilde{\beta}$ and $\hat{\lambda} < \overline{\lambda}$.

Suppose that $\tilde{\beta} < 1$. By (A.2), $b\sigma^T \neq 0$. Also, $\tilde{\lambda} = \overline{\lambda} = 1$. By (A.3), (2.42a), and (2.42b), $P_1(\lambda)$ has limit $P_1(1) = -(A_1 - r_1)/(b\sigma^T)$ when $\lambda \uparrow 1$ and $(P_1(\lambda) - P_1(1))/\sqrt{1-\lambda} \to \Theta_1\sqrt{1-\tilde{\beta}}/((b\sigma^T)^2 c^{-1})$ as $\lambda \uparrow 1$. By (2.45), (2.46), and (A.4), $p_2(\lambda) \to -(a_2 - r_2 - b\beta)/b\sigma^T$, as $\lambda \uparrow 1$, which quantity we denote by $p_2(1)$. By (2.45), (2.46), (A.3) and (A.4), one can expand as follows (either by hand or by the use of Mathematica): as $\lambda \uparrow 1$,

$$p_2(\lambda) = p_2(1) - K_1\sqrt{1-\lambda} - K_2(1-\lambda) + o(1-\lambda)\,,$$

where

$$K_1 = \frac{1}{\Theta_1\sqrt{1-\tilde{\beta}}}\Big(\big(\Theta_1 - \frac{\sigma\sigma^T(A_1 - r_1)}{b\sigma^T}\big)\, p_2(1) + r_1 - \alpha_1 + P_1(1)(\theta_2 - \sigma\beta)\Big)$$

and

$$K_2 = \frac{\sigma\sigma^T}{(b\sigma^T)^2 c^{-1}}\, p_2(1) + \frac{b\beta}{b\sigma^T} \; + \frac{\theta_2 - \sigma\beta}{(b\sigma^T)^2 c^{-1}}\,.$$

By (2.48), $F(\lambda)$ has a finite limit as $\lambda \uparrow 1$, which we denote by $F(1)$. In addition,

$$\lim_{\lambda\uparrow 1} \frac{F(1) - F(\lambda)}{\sqrt{1-\lambda}} = \frac{\sigma\sigma^T |\Theta_1| \sqrt{1-\tilde{\beta}}}{2(b\sigma^T)^2 c^{-1}}\,,$$

implying that $F'(1-) = \infty$, so, F is essentially smooth and $\hat{\lambda} < \overline{\lambda}$.

Let us consider the case that $\tilde{\beta} = 1$, so, $(A_1 - r_1)\big(\sigma\sigma^T(A_1 - r_1) - 2\Theta_1 \sigma b^T\big) = 0$. One has that $\tilde{\lambda} = \overline{\lambda} = 1$, $D(\lambda) = \Theta_1$, $P_1(\lambda) = (-\sigma b^T c^{-1}(A_1 - r_1))/\big((1-\lambda)/\lambda\, \sigma\sigma^T + \sigma b^T c^{-1} b\sigma^T\big)$ and $p_2(\lambda) = -E(\lambda)/\Theta_1$. Thus, if $b\sigma^T = 0$, then $A_1 - r_1 = 0$ and $P_1(\lambda) = 0$. If $b\sigma^T \neq 0$, then $P_1(1) = -(A_1 - r_1)/(b\sigma^T)$, $P_1'(1) = -\sigma\sigma^T(A_1 - r_1)/((b\sigma^T)^3 c^{-1})$, and $P_1''(1) = 2\sigma\sigma^T(A_1 - r_1)/\big((b\sigma^T)^3 c^{-1}\big)\big(1 - \sigma\sigma^T/\big((b\sigma^T)^2 c^{-1}\big)\big)$. Since

$$A_1 - r_1 + b\sigma^T P_1(1) = 0\,, \tag{A.5}$$

$E(\lambda)$ is continuous and differentiable on $[0,1]$, see (2.46), so is $p_2(\lambda)$. By (2.48), if $a_2 - r_2 - b\beta + b\sigma^T p_2(1) \neq 0$, then $F(\lambda) \to \infty$, as $\lambda \uparrow 1$, so $\hat{\lambda} < \overline{\lambda}$. If

$$a_2 - r_2 - b\beta + b\sigma^T p_2(1) = 0\,, \tag{A.6}$$

then

$$F(1) = \frac{1}{2}\,\sigma\sigma^T p_2(1)^2 + (-\sigma\beta + \theta_2) p_2(1) + r_2 - \alpha_2 + |\beta|^2 + \frac{1}{2}\,\sigma\sigma^T P_1(1)$$

and

$$F'(1-) = \sigma\sigma^T p_2'(1-) p_2(1) + \frac{1}{2c}\,(b\sigma^T p_2'(1-) - b\beta)^2 - \beta^T\sigma^T p_2(1) + (-\sigma\beta + \theta_2) p_2'(1-) \\ + r_2 - \alpha_2 + \frac{3}{2}\,|\beta|^2 + \frac{1}{2}\,\sigma\sigma^T P_1'(1-)\,.$$

As one can see, $F(\lambda)$ is not essentially smooth. We obtain that $\hat{\lambda} < \overline{\lambda}$ if and only if $F'(1-) > q$, otherwise $\hat{\lambda} = 1$. It is noteworthy that (A.5) and (A.6) represent conditions (2.50a) and (2.50b), respectively.

The cases where $\tilde{\beta} \geq 1$ and $F(\lambda) \to \infty$ as $\lambda \uparrow 1/\tilde{\beta}$ and where $\tilde{\beta} < 1$ have been analysed in Pham [34].

B Proof of Lemma 2.1

Suppose that the matrix $\sigma(x) Q_1(x) \sigma(x)^T$ is uniformly positive definite. Then $|Q_1(x)\sigma(x)^T y| \geq k_1 |y|$, for some $k_1 > 0$, all $x \in \mathbb{R}^l$ and all $y \in \mathbb{R}^k$. Since $|\sigma(x)^T y|^2 = y^T \sigma(x)\sigma(x)^T y \leq k^2|_2 |y|^2$, for some $k_2 \geq k_1$, we have that

$$\frac{|(I_k - Q_1(x))\sigma(x)^T y|}{|\sigma(x)^T y|} \leq \frac{\sqrt{|\sigma(x)^T y|^2 - k_1^2 |y|^2}}{|\sigma(x)^T y|} \leq \sqrt{1 - \frac{k_1^2}{k_2^2}}\,.$$

Therefore, since $I_k - Q_1(x)$ represents the operator of the orthogonal projection on the range of $b(x)^T$, given $z \in \mathbb{R}^n$,

$$(\sigma(x)^T y)^T b(x)^T z \leq \sqrt{1 - \frac{k_1^2}{k_2^2}}\, |\sigma(x)^T y||b(x)^T z| \,,$$

so nonzero vectors from the ranges of $\sigma(x)^T$ and of $b(x)^T$ are at angles uniformly bounded away from zero. Conversely, if $(\sigma(x)^T y)^T b(x)^T z \leq \rho_1 |\sigma(x)^T y||b(x)^T z|$, for some $\rho_1 \in (0,1)$, then $|(I_k - Q_1(x))\sigma(x)^T y| \leq \rho_1 |\sigma(x)^T y|$ so that $|Q_1(x)\sigma(x)^T y| = \sqrt{|\sigma(x)^T y|^2 - |(I_k - Q_1(x))\sigma(x)^T y|^2} \geq (1-\rho_1)|\sigma(x)^T y| \geq (1-\rho_1)\rho_2 |y|$, the latter inequality holding by $\sigma(x)\sigma(x)^T$ being uniformly positive definite, where $\rho_2 > 0$. Thus, the matrix $\sigma(x)Q_1(x)\sigma(x)^T$ is uniformly positive definite if and only if "the angle condition" holds. Since the angle condition is symmetric in $\sigma(x)$ and $b(x)$, it is also equivalent to the matrix $c(x) - b(x)\sigma(x)^T(\sigma(x)\sigma(x)^T)^{-1}\sigma(x)b(x)^T$ being uniformly positive definite.

In order to prove the second assertion of the lemma, let us observe that

$$\beta(x)^T Q_2(x)\beta(x) = \beta(x)^T Q_1(x)\big(I_k - Q_1(x)\sigma(x)^T(\sigma(x)Q_1(x)Q_1(x)\sigma(x)^T)^{-1}\sigma(x)Q_1(x)\big)Q_1(x)\beta(x) \,,$$

so, if $\beta(x)^T Q_2(x)\beta(x)$ is bounded away from zero, then, by $|Q_1(x)\beta(x)|$ being bounded, there exists $\rho_3 \in (0,1)$ such that, for all $x \in \mathbb{R}^l$,

$$(1-\rho_3)|Q_1(x)\beta(x)| > \big(Q_1(x)\sigma(x)^T(\sigma(x)Q_1(x)Q_1(x)\sigma(x)^T)^{-1}\sigma(x)Q_1(x)\big)Q_1(x)\beta(x) \,.$$

The righthand side representing the orthogonal projection of $Q_1(x)\beta(x)$ onto the range of $(\sigma(x)Q_1(x))^T$ implies that, given $y \in \mathbb{R}^l$,

$$|(Q_1(x)\beta(x))^T Q_1(x)\sigma(x)^T y| \leq (1-\rho_3)|Q_1(x)\beta(x)||Q_1(x)\sigma(x)^T y| \,,$$

which means that $Q_1(x)\beta(x)$ is at angles to $Q_1(x)\sigma(x)^T y$ which are bounded below uniformly over y. The converse is proved similarly.

C

Lemma C.1. *Given $L \in \mathbb{R}$, $m \in \hat{\mathbb{P}}$, and $v \in \mathbb{L}^2(\mathbb{R}^l, \mathbb{R}^n, m(x)\,dx)$, the sets*

$$\{\lambda \in \mathbb{R} :\, \lambda L - \inf_{f \in \mathbb{C}_0^2} \int_{\mathbb{R}^l} \overline{H}(x;\lambda,f,v)\, m(x)\, dx \geq \alpha\}$$

are compact for all $\alpha \in \mathbb{R}$.

Proof. By (4.7),

$$\inf_{f \in \mathbb{C}_0^2} \int_{\mathbb{R}^l} \overline{H}(x;\lambda,f,v)\, m(x)\, dx = \inf_{\nabla f \in \mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)} \int_{\mathbb{R}^l} \big(\lambda M(v(x),x)$$
$$+ \frac{1}{2}\, |\lambda N(v(x),x) + \sigma(x)^T \nabla f(x)|^2 + \nabla f(x)^T \theta(x) - \frac{1}{2}\, \nabla f(x)^T \frac{\operatorname{div}(\sigma(x)\sigma(x)^T m(x))}{m(x)}\big)\, m(x)\, dx \,.$$

The infimum is attained at $\nabla f(x) = \lambda g_1(x) + g_2(x)$, where

$$g_1 = -\Pi\big((\sigma(\cdot)\sigma(\cdot)^T)^{-1}\sigma(\cdot)^T N(v(\cdot),\cdot)\big),$$
$$g_2 = \Pi\big((\sigma(\cdot)\sigma(\cdot)^T)^{-1}\big(-\theta(\cdot) + \frac{\operatorname{div}(\sigma(\cdot)\sigma(\cdot)^T m(\cdot))}{2m(\cdot)}\big)\big)\,,$$

with Π representing the operator of the orthogonal projection on $\mathbb{L}_0^{1,2}(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ in $\mathbb{L}^2(\mathbb{R}^l,\mathbb{R}^l,m(x)\,dx)$ with respect to the inner product $\langle h_1,h_2\rangle = \int_{\mathbb{R}^l} h_1(x)^T\sigma(x)\sigma(x)^T h_2(x)\,m(x)\,dx$. Therefore,

$$\begin{aligned}\lambda L - \inf_{f\in\mathbb{C}_0^2}\int_{\mathbb{R}^l}&\overline{H}(x;\lambda,f,v)\,m(x)\,dx\\ &= \lambda\big(L - \int_{\mathbb{R}^l} M(v(x),x)m(x)\,dx - \int_{\mathbb{R}^l} g_1(x)^T\sigma(x)\sigma(x)^T g_2(x)m(x)\,dx\big)\\ +\frac{1}{2}\int_{\mathbb{R}^l} g_2(x)^T\sigma(x)\sigma(x)^T g_2(x)m(x)\,dx &- \frac{\lambda^2}{2}\int_{\mathbb{R}^l}\big(|N(v(x),x)|^2 - g_1(x)^T\sigma(x)\sigma(x)^T g_1(x)\big)m(x)\,dx\,.\end{aligned} \tag{C.1}$$

Since projection is a contraction operator,

$$\int_{\mathbb{R}^l} g_1(x)^T\sigma(x)\sigma(x)^T g_1(x)m(x)\,dx \le \int_{\mathbb{R}^l} N(v(x),x)^T\sigma(x)^T(\sigma(x)\sigma(x)^T)^{-1}\sigma(x)N(v(x),x)m(x)\,dx\,.$$

As mentioned, by condition (N), $\beta(x)$ does not belong to the sum of the ranges of $b(x)^T$ and of $\sigma(x)^T$. By (2.4b), $N(u,x)$ does not belong to the range of $\sigma(x)^T$, for any u and x . Therefore, the projection of $N(v(x),x)$ onto the null space of $\sigma(x)$ is nonzero which implies that $|N(v(x),x)|^2 - N(v(x),x)^T\sigma(x)^T(\sigma(x)\sigma(x)^T)^{-1}\sigma(x)N(v(x),x)$ is positive for any x , so, the coefficient of λ^2 on the righthand side of (C.1) is positive, yielding the needed property. ☐

The next result seems to be "well known". I haven't been able to find a reference, though.

Lemma C.2. *For arbitrary $\kappa > 0$,*

$$\limsup_{t\to\infty}\mathbf{E}e^{\kappa|X_t|} < \infty\,.$$

Proof. We prove that, if $\gamma > 0$ and is small enough, then

$$\limsup_{t\to\infty}\mathbf{E}e^{\gamma|X_t|^2} < \infty\,.$$

By (2.2), there exist $K_1 > 0$ and $K_2 > 0$ such that, for all $x \in \mathbb{R}^l$, $\theta(x)^T x \le -K_1|x|^2 + K_2$. On applying Itô's lemma to (2.1) and recalling that $\sigma(x)\sigma(x)^T$ is bounded, we have that, for some $K_3 > 0$ and all $i \in \mathbb{N}$, $t \ge 0$,

$$d\mathbf{E}|X_t|^{2i} \le -2iK_1\mathbf{E}|X_t|^{2i}\,dt + 2i^2K_3\mathbf{E}|X_t|^{2i-2}\,dt\,.$$

Hence,

$$\mathbf{E}|X_t|^{2i} \leq \mathbf{E}|X_0|^{2i}e^{-2iK_1t} + 2i^2K_3e^{-2iK_1t}\int\limits_0^t e^{2iK_1s}\mathbf{E}|X_s|^{2i-2}\,ds\,. \tag{C.2}$$

Let

$$M_i(t) = \frac{1}{i!}\sup_{s\leq t}\mathbf{E}|X_s|^{2i}\,.$$

By (C.2),

$$M_i(t) \leq \frac{\mathbf{E}|X_0|^{2i}}{i!} + \frac{K_3}{K_1}\,M_{i-1}(t)\,.$$

Hence, if $\gamma K_3/K_1 < 1$, then

$$\mathbf{E}e^{\gamma|X_t|^2} \leq \sum_{i=0}^{\infty}\gamma^i M_i(t) \leq \frac{1}{1-\gamma K_3/K_1}\sum_{i=0}^{\infty}\frac{\gamma^i\mathbf{E}|X_0|^{2i}}{i!} = \frac{1}{1-\gamma K_3/K_1}\,\mathbf{E}e^{\gamma|X_0|^2}\,.$$

□

References

[1] R.A. Adams and J.J.F. Fournier. *Sobolev spaces*. Academic Press, 2nd edition, 2003. Pure and Applied Mathematics, Vol. 140.

[2] Sh. Agmon. The L_p approach to the Dirichlet problem. I. Regularity theorems. *Ann. Scuola Norm. Sup. Pisa (3)*, 13:405–448, 1959.

[3] J.-P. Aubin. *Optima and equilibria*, volume 140 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993. An introduction to nonlinear analysis, Translated from the French by Stephen Wilson.

[4] J.-P. Aubin and I. Ekeland. *Applied nonlinear analysis*. Wiley, 1984.

[5] T.R. Bielecki and S.R. Pliska. Risk-sensitive dynamic asset management. *Appl. Math. Optim.*, 39(3):337–360, 1999.

[6] T.R. Bielecki and S.R. Pliska. Risk-sensitive ICAPM with application to fixed-income management. *IEEE Trans. Automat. Control*, 49(3):420–432, 2004.

[7] T.R. Bielecki, S.R. Pliska, and S.J. Sheu. Risk sensitive portfolio management with Cox-Ingersoll-Ross interest rates: the HJB equation. *SIAM J. Control Optim.*, 44(5):1811–1843, 2005.

[8] V. I. Bogachev, N. V. Krylov, and M. Röckner. On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions. *Comm. Partial Differential Equations*, 26(11-12):2037–2080, 2001.

[9] V.I. Bogachev, N.V. Krylov, and M. Röckner. Elliptic and parabolic equations for measures. *Uspekhi Mat. Nauk*, 64(6(390)):5–116, 2009.

[10] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems.* Springer Series in Operations Research. Springer-Verlag, New York, 2000.

[11] M. Davis and S. Lleo. Risk-sensitive benchmarked asset management. *Quant. Finance*, 8(4):415–426, 2008.

[12] M. Davis and S. Lleo. Jump-diffusion risk-sensitive asset management I: diffusion factor model. *SIAM J. Financial Math.*, 2(1):22–54, 2011.

[13] M. Davis and S. Lleo. Jump-diffusion risk-sensitive asset management II: Jump-diffusion factor model. *SIAM J. Control Optim.*, 51(2):1441–1480, 2013.

[14] I. Ekeland and R. Temam. *Convex analysis and variational problems.* North Holland, 1976.

[15] S.N. Ethier and T.G. Kurtz. *Markov processes. Characterization and convergence.* Wiley, 1986.

[16] W.H. Fleming and S.J. Sheu. Optimal long term growth rate of expected utility of wealth. *Ann. Appl. Probab.*, 9(3):871–903, 1999.

[17] W.H. Fleming and S.J. Sheu. Risk-sensitive control and an optimal investment model. *Math. Finance*, 10(2):197–213, 2000. INFORMS Applied Probability Conference (Ulm, 1999).

[18] W.H. Fleming and S.J. Sheu. Risk-sensitive control and an optimal investment model. II. *Ann. Appl. Probab.*, 12(2):730–767, 2002.

[19] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 1983.

[20] H. Hata, H. Nagai, and S.J. Sheu. Asymptotics of the probability minimizing a "downside" risk. *Ann. Appl. Probab.*, 20(1):52–89, 2010.

[21] N. Ichihara. Recurrence and transience of optimal feedback processes associated with Bellman equations of ergodic type. *SIAM J. Control Optim.*, 49(5):1938–1960, 2011.

[22] A. D. Ioffe and V. M. Tihomirov. *Theory of extremal problems*, volume 6 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam-New York, 1979. Translated from the Russian by Karol Makowski.

[23] H. Kaise and S.J. Sheu. On the structure of solutions of ergodic type Bellman equation related to risk-sensitive control. *Ann. Probab.*, 34(1):284–320, 2006.

[24] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1988.

[25] K. Koncz. On the parameter estimation of diffusional type processes with constant coefficients (elementary Gaussian processes). *Anal. Math.*, 13(1):75–91, 1987.

[26] K. Kuroda and H. Nagai. Risk-sensitive portfolio optimization on infinite time horizon. *Stoch. Stoch. Rep.*, 73(3-4):309–331, 2002.

[27] O.A. Ladyzhenskaya and N.N. Ural′tseva. *Linear and quasilinear elliptic equations.* Translated from the Russian by Scripta Technica, Inc. Translation editor: Leon Ehrenpreis. Academic Press, New York-London, 1968.

[28] R. Liptser. Large deviations for two scaled diffusions. *Probab. Theory Related Fields*, 106(1):71–104, 1996.

[29] R.S. Liptser and A.N. Shiryayev. *Statistics of random processes. I.* Springer-Verlag, New York-Heidelberg, 1977. General theory, Translated by A. B. Aries, Applications of Mathematics, Vol. 5.

[30] R.Sh. Liptser and A.N. Shiryayev. *Theory of martingales.* Kluwer, 1989.

[31] G. Metafune, D. Pallara, and A. Rhandi. Global properties of invariant measures. *J. Funct. Anal.*, 223(2):396–424, 2005.

[32] H. Nagai. Optimal strategies for risk-sensitive portfolio optimization problems for general factor models. *SIAM J. Control Optim.*, 41(6):1779–1800, 2003.

[33] H. Nagai. Downside risk minimization via a large deviations approach. *Ann. Appl. Probab.*, 22(2):608–669, 2012.

[34] H. Pham. A large deviations approach to optimal long term investment. *Finance Stoch.*, 7(2):169–195, 2003.

[35] A.A. Puhalskii. On portfolio choice by maximizing the outperformance probability. *Math. Finance*, 21(1):145–167, 2011.

[36] A.A. Puhalskii. On large deviations of coupled diffusions with time scale separation. *Ann. Probab.*, 44(4):3111–3186, 2016.

[37] A.A. Puhalskii and M.J. Stutzer. On a portfolio's shortfall probability. arXiv:1602.02192, 2016.

[38] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[39] A.V. Skorokhod. *Asymptotic methods in the theory of stochastic differential equations*, volume 78 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1989. Translated from the Russian by H. H. McFaden.

[40] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes.* Springer, 1979.

[41] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009. Old and new.

[42] J.C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Trans. Automatic Control*, AC-16:621–634, 1971.

[43] W.M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM J. Control*, 6:681–697, 1968.