Simulation of Implied Volatility Surfaces via Tangent Lévy Models[*]

Sergey Nadtochiy [†] and Thaleia Zariphopoulou[‡]

Abstract. In this paper, we construct a solution to the optimal contract problem for delegated portfolio management of the first-best (risk-sharing) type. The novelty of our result is (i) in the robustness of the optimal contract with respect to perturbations of the wealth process (interpreted as capital injections), and (ii) in the more general form of principal's objective function, which is allowed to depend directly on the agent's strategy, as opposed to being a function of the generated wealth only. In particular, the latter feature allows us to incorporate endogenous trading constraints in the contract. We reduce the optimal contract problem to the following inverse problem: for a given portfolio (defined in a feedback form, as a random field), construct a stochastic utility for which the given portfolio is optimal. We characterize the solution to this problem through a Stochastic Partial Differential Equation (SPDE), prove its well-posedness, and compute the solution explicitly in the Black-Scholes model.

1. Introduction.

In this paper, we study a problem of delegated portfolio management. The basic formulation of the problem (on which we build our setup) is as follows. An investor hires a fund manager (referred to as the agent) for a specified period of time, to invest her capital dynamically in the available assets. At the end of the time period, the investor receives the wealth generated by the manager and, in return, pays the fees prescribed by the contract. The fees are allowed to depend on the wealth level and on all publicly observed factors that generate the market filtration. As we assume a non-degenerate Itô's market model, the agent's strategy is uniquely determined by the (terminal value of the) generated wealth process,[1] and since the agent does not have any superior information about the market relative to the investor,[2] the investor can choose any target strategy and restrict the payoff of the contract to the set on which the generated wealth coincides with the one associated with the target strategy, thus, forcing the agent to follow this strategy (see Subsection 4.2 for more). As a result, in such a setting, the optimal strategy of the agent and the associated optimal contract can be chosen so that they jointly maximize the investor's objective. Such problems are known as "first-best", or the optimal risk-sharing problems.

The existing literature on the optimal contract design for the delegated portfolio management problem, of the first-best type, includes [16], [17], [12], [2], and the references therein.[3]

[*] Submitted to the editors on February 22, 2018.

[†] Department of Applied Mathematics, Illinois Institute of Technology (snadtochiy@iit.edu). This author is partially supported by the NSF CAREER grant DMS-1855309.

[‡] Department of Mathematics, University of Texas Austin (zariphop@math.utexas.edu).

[1] This follows directly from the uniqueness of the martingale representation. It is important to notice that the agent's strategy is still not directly observable, in the sense that it cannot be deduced from the wealth value on a single random outcome. However, if the wealth is known as a random variable, then, the strategy can be deduced uniquely from it.

[2] Relaxing this assumption would, e.g., correspond to assuming that the investor does not observe the prices of some of the assets, and it would introduce the information asymmetry, or "moral hazard", in the model. We do not consider such a version of the problem in the present paper.

[3] In this paper, we limit our literature review to the papers that are dealing with the delegated portfolio management problem specifically, leaving aside the discussion of general optimal contract theory, such as the

Single period models are analyzed in [16] and [17], while [12] considers the Black-Scholes model, with the investor and the fund manager having either exponential or power utilities. A general market model and general utilities are considered in [2], which, in particular, constructs an optimal contract explicitly when the market is complete.

The present work differs from the existing results in that, herein, (i) we require that the contract is robust with respect to the perturbations of wealth process, and (ii) we consider a more general optimality criterion for a contract than the classical expected utility of terminal wealth. Our main motivation to consider the perturbations of wealth process is to include (unanticipated) *capital injections* made by the investor after the contract is initiated. Namely, we assume that the contract allows the investor (as, e.g., most fee structures of mutual funds do) to add an arbitrary amount of additional capital to her account with the manager, at any time when she wishes to do so, and with the fee structure for the manager remaining the same (i.e. the contract remains the same). Note that these times and amounts, and even their probabilistic structure, may not be initially known to either one of the two parties. However, the inflow of capital in the fund may change the investment strategy of the fund manager drastically (see, e.g. [1], and the references therein, for more on the effects of capital inflows and outflows on the behavior of a fund manager). Thus, when designing an optimal contract, one needs to take into account the agent's optimal strategy, induced by this contract, for any intermediate time and wealth level. Mathematically, this means that the agent's strategy should be viewed as a random field, defined for all possible initial investment times and wealth levels. Another implication of this feature is that it is no longer obvious how the principal can force the agent to follow a chosen strategy. Indeed, since the agent's strategy is not directly observable by the principal, it can only be enforced indirectly, through the wealth process. Even though it turns out that such enforcement can be implemented in our setting, the implementation itself (i.e., the form of a contract) is far from obvious (unlike the classical case, with no capital injections). Subsection 4.2 contains a more detailed discussion of this issue.

Another special feature of our setting is a more general optimality criterion for the contract. Namely, we assume that the entity designing the contract (referred to as the principal) may be concerned directly with the strategy used by the agent, in addition to the wealth generated by this strategy.[4] Our main motivating example of such preference structure is the case of *constrained maximization* of expected utility of terminal wealth, with the constraint that no investment is made in certain assets. In such a case, the principal's objective contains an infinite penalty for investing in the prohibited assets, and the contract must be designed so that the agent follows this rule. For example, a regulator or the board of directors of a mutual fund may want to enforce a ban on investments in certain "socially irresponsible" assets, or in the assets of companies subject to sanctions (we refer the reader to [14], and the reference therein, for more on the so called "socially responsible" funds). However, the principal cannot put such a rule into a contract directly, as she does not observe the agent's actions. Hence, these constraints need to be enforced implicitly, through the design of the contract, which can only depend on the generated wealth and on the publicly observed factors – this is what we

seminal work [5].

[4] As explained in the next paragraph, the investor may not coincide with the principal, in our setting.

refer to as the *endogenous constraints*.

Let us describe a specific setting in which the robustness of the contract with respect to capital injections and the endogenous constraints are important (a more detailed formulation is given in Section 4). First, we assume that the principal, who designs the contract, may not coincide with the investor (at least, they may not coincide for the entire duration of the contract). For example, the fee structure of a mutual fund is very often prescribed a priori, and an individual investor can either take it or leave it. In this case, the principal may be a regulator or the board of directors of the mutual fund.[5] Even though the principal may not coincide with the investor, we assume that she aims to design the contract so that the investor is satisfied: e.g. the board of directors of a mutual fund wants to keep their investors happy, in order not to lose them to the competitors. At the same time, the principal also wants to ensure that the agent does not invest in the prohibited assets.[6] Thus, the principal finds a strategy that maximizes the investor's expected utility of terminal wealth, subject to the constraint that no investment is made in the prohibited assets, and aims to design a contract (which is only allowed to depend on the generated wealth and on the publicly observable market factors) which would make this strategy optimal for the agent. This task is complicated by the fact that the investor may perform capital injections, whose times and sizes are unknown (i.e. not modeled) initially. Namely, the investor, unlike the principal, may not be concerned about investing in prohibited assets, hence, she may perform a capital injection even if it encourages the agent to violate this constraint. Thus, the contract has to be chosen by the principal so that the agent has no incentive to violate the constraint even in the presence of capital injections – this is what we refer to as the *robustness with respect to capital injections*.

On the mathematical side, this paper solves the following inverse problem: given a regular enough random field, find a stochastic utility whose optimal investment strategy, in the feedback form, coincides with this random field. We characterize the solution through a linear stochastic partial differential equation (SPDE), prove its well-posedness, and compute the solution explicitly in the Black-Scholes model.

Let us describe how the proposed method for constructing an optimal contract compares to the existing literature. The typical construction of an optimal contract starts by finding a convenient description of a sufficiently large family of potential contracts (so that the principal's maximization problem can be reduced to this family without loss of optimality). The "convenience" is determined by how tractable is the description of the contract and the associated optimal strategy of the agent. For example, [12] assumes a convenient differential representation of a contract (which, potentially, may lead to a loss of optimality), for which the optimal effort of the agent can be computed explicitly in a feedback form. In [2], the market completeness allows for the use of convex duality in oder to describe the optimal strategy of the agent. More generally, it was proposed in the seminal work [15] that a convenient

[5]Alternatively, the principal may be an initial investor, who enters into a long-term contract with the fund manager and passes on her wealth to the successors. The successors cannot withdraw funds before the deadline, but they may be allowed to add capital, keeping the fee structure as prescribed by the principal.

[6]For example, a university may want to ensure that its endowment is not invested in the stocks of the companies engaged in the production and distribution of fossil fuel, or tobacco and alcohol. We thank the anonymous referee for this example.

parameterization of the space of potential contracts is through the associated (continuation) value processes. For example, in a finite-horizon setting, the contract is given by the value process at the terminal time. In many relevant problems, the Pontryagin principle provides a representation of the value process of the agent as the solution to a backward stochastic differential equation (BSDE), and the associated optimal strategy is expressed through the value process and its diffusion coefficient. Then, in order to describe all contracts and the associated optimal strategies, it suffices to describe all solutions to the aforementioned BSDE. It turns out that, since the terminal condition of the BSDE is not fixed, the space of its solutions can be described by "reversing the time" and viewing the BSDE as a family of forward stochastic differential equations (SDEs). Any solution to such SDE also solves the desired BSDE, which makes it a value process and yields a tractable representation for the associated contract and the optimal strategy of the agent. In turn, the aforementioned family of SDEs can be parameterized explicitly by varying the diffusion coefficient of this equation (viewed as a stochastic process), and its initial condition, over a tractable set.[7] For any choice of the diffusion coefficient and the initial condition, one obtains the associated contract and the agent's optimal strategy explicitly, and the latter determine the objective of the principal. Thus, the principal's problem is reduced to a standard optimal control problem, where the coefficients of the state process depend explicitly on the control (the latter represents the diffusion coefficient of the value process). This approach appears to be prevailing in the modern literature: see, e.g., [4], [3]. Our approach follows the same general methodology. However, the main difference is that, herein, we need to consider value processes for all possible initial wealth levels, in order to ensure that the optimal strategy of the agent is stable with respect to capital injections. Such formulation does not allow for the use of Pontryagin principle, as it only describes the optimal wealth dynamics starting from a given initial point, but does not characterize the optimal wealth processes starting form other intermediate levels. Hence, we replace the Pontryagin BSDE by the stochastic Hamilton-Jacobi-Belman (HJB) equation (3.3), also known as the forward performance SPDE. The latter describes the optimal wealth processes and the associated optimal strategies for all initial wealth levels and initial times. The main mathematical challenge associated with this modification is that, unlike the case of BSDE, one cannot simply "reverse the time" in a parabolic SPDE: the resulting time-reversed equation is, typically, not well posed for most initial conditions (see the discussion of this issue for parabolic PDEs in [11]). Thus, another interpretation of the solution to the inverse problem mentioned in the preceding paragraph is as follows: we find an explicit representation of a sufficiently large family of solutions to the stochastic HJB equation via the associated optimal strategies. Having established such a representation, we follow the standard approach and parameterize the space of contracts explicitly via the agent's optimal strategies: each contract is computed from the terminal value of the solution to an SPDE, whose coefficients depend on the agent's strategy. As a result, the principal's problem can be reformulated as an optimal control problem over the agent's strategies (viewed as random fields). The latter, in general, becomes an infinite-dimensional control problem (i.e., control of SPDE), which, however, is made trivial in the present setting by assuming the additive structure of the principal's objec-

[7] For the sake of simplicity, this discussion is restricted to diffusion-based models, and it excludes the adverse-selection, or "third-best", problems, where 2BSDEs are used (cf. [3]).

tive with respect to the contract fees. It is also worth mentioning that our description of the solutions to stochastic HJB equation works only in the case of symmetric information. The additional challenges arising in the asymmetric case (i.e., in the case of "moral hazard") are discussed in Remark 10.

The rest of the paper is organized as follows. In Section 2, we formulate the optimal contract problem precisely, in mathematical terms. Subsection 2.1 is concerned with the market model, and Subsection 2.2 introduces the notions of admissible and optimal contracts. Section 3 presents a general solution to the problem, which reduces to the inverse problem of constructing an optimization criterion that generates a given optimal strategy (viewed as a random field), for all initial wealth levels. Proposition 3.1 connects this problem to a nonlinear SPDE, and Proposition 3.2 shows how to linearize this SPDE and proves the well-posedness of the resulting equation. Finally, Theorem 3.3 connects these results to the optimal contract problem. In Section 4, we consider a specific setting in which the proposed notion of optimal contract is natural, and use the general results of preceding sections to construct an optimal contract in closed form, in the Black-Scholes model. Remarkably, the optimal contract constructed in Section 4 depends only on the values of the wealth process and of the tradable assets at the terminal time.

2. Problem formulation.

2.1. Market model and investment strategies. We fix a stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$ and assume that the publicly observed filtration $\mathbb{F}$ (also referred to as the market filtration) is an augmentation of the filtration generated by W, a standard Brownian motion in $\mathbb{R}^d$. In addition, we assume that the price process of traded assets $S = (S^1, \dots, S^k)^T$ is an Itô process in $\mathbb{R}^k$ with positive entries, given by

$$d \log S_t = \tilde{\mu}_t dt + \sigma_t^T dW_t, \tag{2.1}$$

where the logarithm is taken entry-wise, $\tilde{\mu}$ is a locally integrable stochastic process with values in $\mathbb{R}^k$, and σ is a $d \times k$ matrix of locally square integrable processes, with $d \geq k$, and with linearly independent columns. The latter assumptions is interpreted as the absence of redundant assets. We use the notation "A^T" to denote the transpose of a matrix (vector) A. For simplicity, we set the riskless interest rate to zero (equivalently, we work with discounted units). We introduce the d-dimensional stochastic process λ, frequently called the market price of risk, via

$$\lambda_t := \left(\sigma_t^T\right)^+ \mu_t, \tag{2.2}$$

where $(\sigma_t^T)^+$ is the Moore-Penrose pseudo-inverse of the matrix σ_t^T, and μ is the drift of S: $\mu_t^i = \tilde{\mu}_t^i + \|\sigma_t^i\|^2/2$, for $i = 1, \dots, k$, with σ_t^i being the i-th column of σ_t. In particular, we have $\sigma_t^T \lambda_t = \mu_t$. The existence of such a process λ follows from the assumption of absence of arbitrage in the model. Denote by $\mathcal{X}$ a set of pairs (ξ, τ), with $\tau \in \mathcal{T}$ and $\xi \in L_+^0(\mathcal{F}_\tau)$, where $\mathcal{T}$ is the set of all $\mathbb{F}$-stopping times, and $L_+^0(\mathcal{F}_\tau)$ is the set of all positive $\mathcal{F}_\tau$-measurable random variables. Starting from any $(\xi, \tau) \in \mathcal{X}$, the cumulative wealth process $X^{\pi,\xi,\tau}$ is given by

$$dX_s^{\pi,\xi,\tau} = \pi_s^T \sigma_s^T \lambda_s ds + \pi_s^T \sigma_s^T dW_s, \quad s \in (\tau, T], \quad X_\tau^{\pi,\xi,\tau} = \xi, \tag{2.3}$$

for any progressively measurable process π, representing the self-financing trading strategy, for which the integrals associated with the right hand side of (2.3) are well defined. We assume that π is such that $X^{\pi,\xi,\tau}$ is almost surely strictly positive at all times. For each pair (ξ, τ), we fix a subset of such strategies $\mathcal{A}(\xi, \tau)$, and call any $\pi \in \mathcal{A}(\xi, \tau)$ (ξ, τ)-admissible (or, just admissible, if the rest is clear from the context).

Remark 1. *It is possible to drop the restriction to strictly positive wealth processes. However, in this case, the construction (and the form) of the optimal contract would change accordingly (cf. the proof of Proposition 3.2 and Remark 8).*

2.2. Optimal contract. Consider an investor who hires a risk-neutral agent in order to invest her initial capital $X_0 > 0$ in the market described above. The agent is offered a contract, which is represented by a measurable mapping $C : \Omega \times (0, \infty) \to \mathbb{R}$, which maps the terminal value of a wealth process (produced by the agent, via a chosen trading strategy π) into the payment (received by the agent at time T). The agent is risk-neutral, in that he aims to maximize his expected objective:

$$\max_{\pi} \mathbb{E}\, C(X_T^{\pi}), \tag{2.4}$$

where the maximization is performed over all admissible strategies π, with C being fixed. The agent will not enter into a contract if his expected payment does not reach a given level $u_0 > 0$. We define an admissible contract as a contract for which the agent's optimization problem is well posed, and such that the participation constraint is satisfied.

Definition 2.1. *We call C an admissible contract if the following holds.*

- *For any $(\xi, \tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi, \tau)$, $C(X_T^{\pi,\xi,\tau})$ is absolutely integrable.*
- *There exists a progressively measurable random field $\pi^* : [0, T] \times \Omega \times (0, \infty) \to \mathbb{R}$, such that:*
  - *for any $(\xi, \tau) \in \mathcal{X}$, there exists a unique $X^{*,\xi,\tau}$ satisfying (2.3), with $\pi = (\pi_t^*(X_t^{*,\xi,\tau}))$,*
  - *for any $(\xi, \tau) \in \mathcal{X}$, $(\pi_t^*(X_t^{*,\xi,\tau})) \in \mathcal{A}(\xi, \tau)$,*
  - $\mathbb{E}\, C\left(X_T^{*,X_0,0}\right) \geq u_0$,
  - *for any $(\xi, \tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi, \tau)$,*
  $$\mathbb{E}\left(C(X_T^{\pi,\xi,\tau}) \mid \mathcal{F}_\tau\right) \leq \mathbb{E}\left(C(X_T^{*,\xi,\tau}) \mid \mathcal{F}_\tau\right), \quad a.s.,$$
  and the equality is only possible if $\pi = (\pi_t^*(X_t^{*,\xi,\tau}))$ for a.e. (t, ω) in the stochastic interval $[\tau, T]$.

Any such strategy π^* is called C-optimal.

Throughout the rest of the paper, a "strategy" may refer to either a stochastic process (i.e. an element of $\mathcal{A}(\xi, \tau)$) or a random field (whose values along the generated wealth form an element of $\mathcal{A}(\xi, \tau)$), whenever the meaning is clear from the context. Similarly, in $X^{\pi,\xi,\tau}$, the term π may represent either a stochastic process or a random field – in either case, it is clear how (π, ξ, τ) generates $X^{\pi,\xi,\tau}$.

The special feature of the above definition, which differentiates it from the classical setup, is that the agent is allowed to re-evaluate his strategy at intermediate times, and starting from various wealth levels, which, in particular, may not coincide with the wealth generated by his strategy thus far. In addition, at each re-evaluation, the agent has to follow the exact strategy prescribed by the optimal random field: i.e. the optimal strategy is time-consistent and unique. A motivation for such strong definition of an optimal contract is given in the discussion following Definition 2.2, and a specific problem is described in Section 4.

The contract is designed by a principal who aims to maximize the expectation of her individual objective J, which maps any progressively measurable random field $\pi : [0,T] \times \Omega \times (0,\infty) \to \mathbb{R}$ into an $\mathcal{F}_T$-measurable random variable $J(\pi)$, applied to the strategy used by the agent, less the payment to the agent:

$$\max_{C} \mathbb{E}\left[J(\pi) - C\left(X_T^{\pi}\right)\right]. \tag{2.5}$$

The above maximization is performed over all admissible contracts C, with the strategy π being C-optimal.

Definition 2.2. *An admissible contract C^* is a solution to the optimal contract problem (2.4)–(2.5), also referred to as an optimal contract, if, for any C^*-optimal strategy π^*, any admissible contract C, and any C-optimal π, we have*

$$\mathbb{E}\left(J(\pi) - C\left(X_T^{\pi,X_0,0}\right)\right) \leq \mathbb{E}\left(J(\pi^*) - C^*\left(X_T^{*,X_0,0}\right)\right),$$

where $X^{\pi,X_0,0}$ and $X^{*,X_0,0}$ are the wealth processes associated with π and π^*, respectively, and with the initial condition X_0 at time zero.

The main difference between the above formulation of the optimal contract problem and the classical one is that, in the present case, the principal needs to predict the agent's strategy for various initial wealth levels, which may not correspond to the levels generated by the strategy itself. The reason for such a formulation is explained in the introduction: on the one hand, we want to allow for (positive) capital injections after the contract is initiated, on the other hand, we do not want to impose any probabilistic structure on the times or the sizes of these injections. In such a robust formulation, the capital under management may change (increase) in an "unpredictable way" at any given time, which, naturally, forces the agent to change his strategy. However, Definition 2.1 ensures that, even if an injection is made, the agent's optimal strategy is still given by the same random field (only started from a different wealth level). Thus, in the presence of unknown capital injections, the contract can only determine the agent's optimal strategy as a random field. This makes it natural to define the principal's objective as a function of such random field. A specific example that leads to an optimal contract problem of the present type is described in Section 4.

It is worth mentioning that, in the classical formulation of the problem, if we assume no capital injections and view strategies as stochastic processes, with a fixed initial wealth, the optimal contract problem typically reduces to the "first best" type, which has a trivial solution. This is due to the fact that, in a non-degenerate market, one can infer the trading strategy from a terminal value of the wealth process (viewed as a random variable). An example of such trivial construction is given in Subsection 4.2. However, the mapping from wealth to strategy (viewed as a stochastic process) depends on the initial capital, hence, the resulting, trivial, solution is not robust with respect to capital injections. The optimal contract defined above (with an example constructed in Subsection 4.3) is robust with respect to such injections, and it is also optimal in the classical formulation. Thus, in particular, it provides another, non-trivial, solution to the classical problem.

It is also important that $J(\pi)$ may depend on π in a more general way – not only through X^π. Otherwise, the problem becomes trivial in many cases of interest, as illustrated in Subsection 4.2. As discussed in the introduction, our main motivation for considering general dependence on π is the presence of endogenous constraints. Namely, we assume that the principal does not want the agent to invest in certain stocks but cannot simply include it in the contract, as the agent's strategy is not directly observable.

Remark 2. *Note that we allow the principal's individual objective, J, and the contract, C, to be quite general. However, the principal's total objective combines them in the additive way: $J - C$. From an economic point of view, it may be more natural to include the agent's fees inside J, but it is not allowed by the current setting. As mentioned in the introduction, if the agent's fees are included inside J, the problem turns into an optimal control of SPDE, which is significantly more challenging. Nevertheless, the subsequent sections show that, in the present setting, the optimal contract is constructed as $C(x) = \overline{C}_T(x)$, where $\overline{C}$ is a sufficiently smooth random filed, so that we can define*

$$C(X_T^\pi) = \overline{C}_0(X_0) + \int_0^T d\,\overline{C}_t(X_t^\pi).$$

As we assume no discounting (equivalently, we work with discounted units), the above representation can be interpreted as a flow of payments from the principal to the agent. As these payments are spread over the entire time interval $[0,T]$, it is possible to justify their appearance in the additive form in the principal's objective. Indeed, it is natural for the principal to treat differently (i.e. have different types of utility for) the terminal and the intermediate payments.

Remark 3. *The assumption of risk-neutrality of the agent can be relaxed by assuming that he maximizes the expected utility of his fees, $U(C)$. However, in such a case, we would either have to replace C in the principal's objective by $U(C)$, or the agent's participation constraint would have to be formulated in terms of expected fees (as opposed to expected utility of his fees), none of which is very natural. In addition, we do not allow for a cost of effort in the agent's objective. These are the limitations we have to accept in order to be able to use our solution approach. We leave the case of more general preferences and cost structures for future research.*

Remark 4. *The optimal contract constructed herein is also robust with respect to maturity. Namely, our method allows one to construct an entire family of optimal contracts, $\{C_T\}$, for all maturities $T > 0$. Thus, we also solve a slightly more general optimal contract problem (of the so-called "third best" type), in which the agent is allowed to choose the time horizon (when the contract is initiated), and the principal does not know which horizon the agent prefers, hence, she offers him a menu of contracts, for all possible horizons.*

Remark 5. *A very desirable feature of a contract is its limited liability: i.e. the condition $C \geq 0$. Note that we do not require limited liability in the definition of admissible contract, and our general results do not guarantee that this property is satisfied by the optimal con-*

tract. However, the optimal contract constructed in Section 4 does satisfy the limited liability condition.

3. Solution. Let us outline, heuristically, the solution approach. First, we notice that, if C is an admissible contract and π^* is C-optimal, with the associated optimal wealth X^*, the contract

$$\tilde{C} := C \frac{u_0}{\mathbb{E}\, C(X^*_T)} \tag{3.1}$$

is also admissible, and the set of $\tilde{C}$-optimal strategies is the same as the set of C-optimal strategies. In addition,

$$\mathbb{E}\, \tilde{C}(X^*_T) = u_0.$$

Thus, there is no loss of optimality in restricting the candidate contracts C to those admissible contracts for which $\mathbb{E}C(X^\pi) = u_0$, for every C-optimal π. This implies that we can drop the expected payment to the agent in the principal's objective and solve the relaxed problem: find a random field π^* and the associated optimal wealth X^* (with initial condition $(X_0, 0)$), such that

$$\pi^* \in \operatorname{argmax} \mathbb{E}\, J(\pi),$$

where the maximization is performed over all random fields π, such that, for any $(\xi, \tau) \in \mathcal{X}$, $(\pi_t(X^{\pi,\xi,\tau}_t)) \in \mathcal{A}(\xi, \tau)$. Then, for the π^* obtained as above, we need to construct an admissible contract C, such that π^* is the only C-optimal strategy. Normalizing C as in (3.1), we obtain the desired optimal contract.

Thus, the construction of an optimal contract reduces to solving the following *inverse problem*: given a strategy π^* (viewed as a random field), find an admissible contract C, such that, for any $(\xi, \tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi, \tau)$,

$$\mathbb{E}\left(C(X^{\pi,\xi,\tau}_T) \mid \mathcal{F}_\tau\right) \leq \mathbb{E}\left(C(X^{*,\xi,\tau}_T) \mid \mathcal{F}_\tau\right), \quad a.s.,$$

and the equality is only possible if $\pi = (\pi^*_t(X^{*,\xi,\tau}_t))$ for a.e. (t, ω) in the stochastic interval $[\tau, T]$. Fortunately, a solution to such problem is offered by the so-called forward performance SPDE. In the remainder of this section, we describe this solution, given by a random field $(U_t(x))_{t\geq 0,\, x>0}$, and show that

$$C(x) = u_0 \frac{U_T(x)}{U_0(X_0)},$$

is the desired optimal contract.

3.1. Forward performance SPDE. We start this subsection by recalling (heuristically) the derivation of the stochastic Hamilton-Jacobi-Bellman (HJB) equation in the classical utility maximization problem (also known as Merton problem). These derivations, leading up to Proposition 3.1, can be found, e.g., in [10], [18], as well as in many subsequent publications, but we provide them here for the sake of completeness. This derivation is based on the following martingale principle. Consider a progressively measurable random field $U = (U_t(x))$, representing the stochastic value function – i.e., the maximal expected utility of terminal wealth that can be generated from the initial wealth level x and initial time t. Then (under additional technical assumptions), this random field satisfies the following two properties:

- for any attainable wealth process X, the process $(U_t(X_t))$ is a supermartingale;
- there exists a wealth process X^*, such that the process $(U_t(X_t^*))$ is a martingale.

The converse is also true: if one can find a progressively measurable random field U, satisfying the above properties on a time interval $[0,T]$, then, this random field is the value function of the Merton problem with stochastic utility $U_T(\cdot)$, and X^* is the associated optimal wealth process. Relaxing the supermartingale and martingale conditions to their local versions, and assuming enough regularity of the random field U (in the sense of [9]), we translate the two defining properties of a stochastic value function into the following: the drift of $(U_t(X_t))$ is non-positive for any admissible X, and it is zero at some X. The latter, in turn, is equivalent to: the maximum drift of $(U_t(X_t))$, over all admissible portfolios, is zero. Assuming an Itô representation for U,

$$dU_t(x) = b_t(x)dt + a_t^T(x)dW_t, \tag{3.2}$$

with sufficiently regular random fields a and b (we refer to a as the volatility of U), we apply the Itô-Ventzel formula to $(U_t(X_t))$, to obtain:

$$dU_t\left(X_t\right) = \left(b_t + \partial_x U_t\, \pi_t^T \sigma_t^T \lambda_t + \frac{1}{2}\partial_{xx}^2 U_t\, \pi_t^T \sigma_t^T \sigma_t \pi_t + \partial_x a_t^T\, \sigma_t \pi_t\right) dt + (\cdots)dW_t,$$

where π is the strategy generating X, via (2.3). Assuming $\partial_{xx}^2 U_t < 0$, the maximum over π of the drift in the above expression can be computed explicitly. Equating it to zero, we obtain a formula for b. Substituting this formula into (3.2), we obtain

$$dU_t(x) = \frac{1}{2}\frac{\|\partial_x U_t(x)\lambda_t + (\sigma_t^T)^+ \sigma_t^T \partial_x a_t(x)\|^2}{\partial_{xx}^2 U_t(x)} dt + a_t^T(x)dW_t, \quad t \in [0,T]. \tag{3.3}$$

In a Markovian model, $U_t(x)$ becomes a deterministic function of (t,x) and the levels of relevant stochastic factors. Using Itô's formula, we represent a and the left hand side of (3.3) through the derivatives of U and obtain the classical HJB equation. Hence, we refer to (3.3) as the stochastic HJB equation, although it was initially named "forward performance SPDE", by the authors of [10], [18].[8] The following proposition formalizes the above discussion.

Proposition 3.1. *Assume that* $a = (a_t(x))_{t\in[0,T],\, x>0}$ *and* $U = (U_t(x))_{t\in[0,T],\, x>0}$, *respectively, are once and twice continuously differentiable random fields (in the sense of [9]), satisfying (3.3), and such that U is strictly concave in x (almost surely, for all times). Then, the following holds.*

1. *For any* $(\xi,\tau) \in \mathcal{X}$ *and any* $\pi \in \mathcal{A}(\xi,\tau)$, *the process* $\left(U_t\left(X_t^{\pi,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ *is a local supermartingale (in the sense that there exists a localizing sequence that makes it a supermartingale).*[9]
2. *Assume that there exists a progressively measurable random field* π^*, *satisfying almost surely, for all* $t \in [0,T]$,

$$\sigma_t \pi_t^*(x) = -\frac{\lambda_t \partial_x U_t(x) + (\sigma_t^T)^+ \sigma_t^T \partial_x a_t(x)}{\partial_{xx}^2 U_t(x)}, \qquad \forall x > 0, \tag{3.4}$$

[8]In the forward performance theory, it is crucial that the SPDE (3.3) holds for all $t \geq 0$. For the problem under consideration, it is not important, and we can consider $t \in [0,T]$.

[9]Throughout the paper, such process is always defined with respect to the filtration $(\mathcal{F}_{\tau\vee t})_{t\in[0,T]}$, and its value on $[0,\tau]$ is $U_\tau(\xi)$.

and such that, for any initial condition $(\xi,\tau) \in \mathcal{X}$, there exists a unique (strong) solution $X^{*,\xi,\tau}$ to

$$dX_t^{*,\xi,\tau} = \left(\sigma_t \pi_t^*(X_t^{*,\xi,\tau})\right)^T \lambda_t dt + \left(\sigma_t \pi_t^*(X_t^{*,\xi,\tau})\right)^T dW_t, \quad t \in [\tau, T], \quad X_\tau^{*,\xi,\tau} = \xi. \tag{3.5}$$

Then, $\left(U_t\left(X_t^{*,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a local martingale.

3. *Assume that the conditions of the previous two items are satisfied, and that, in addition, the aforementioned local martingale and local supermartingales are a true martingale and true supermartingales, respectively. Then, for any $(\xi,\tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi,\tau)$,*

$$\mathbb{E}\left(U_T(X_T^{*,\xi,\tau}) \mid \mathcal{F}_\tau\right) \geq \mathbb{E}\left(U_T(X_T^{\pi,\xi,\tau}) \mid \mathcal{F}_\tau\right) \quad a.s.,$$

and the equality is only possible if $\pi = (\pi_t^*(X_t^{*,\xi,\tau}))$ for a.e. (t,ω) in the stochastic interval $[\tau, T]$.

Proof:

The proof follows easily from the preceding discussion. In particular, the first two claims follow immediately. For the last claim, we only need to notice that the drift of $U_t(X_t^{\pi,\xi,\tau})$ is strictly negative unless $\pi_t = \pi_t^*(X_t^{\pi,\xi,\tau})$, with π^* given by (3.4). The inequality between conditional expectations follows directly from the martingale and supermartingale properties. ∎ None

The last item of the above theorem implies that $(\pi_t^*(X_t^*))$ maximizes the criterion $\mathbb{E}U_T(X_T^\pi)$ over all admissible strategies, provided it is, itself, admissible. Of course, to establish this, one needs to (i) solve the SPDE (3.3), (ii) ensure the existence of π^* and X^*, and (iii) drop "local" in the supermartingale and martingale properties. One way to ensure that the local supermartingale $(U_t(X_t^\pi))_{t\geq 0}$ is a true supermartingale, is to construct U so that $\inf_{t,x} U_t(x)$ is bounded from below by an absolutely integrable random variable, and to restrict the initial wealth to absolutely integrable random variables. Then, one can also show by a standard argument that the local martingale $(U_t(X_t^*))_{t\geq 0}$ is a true martingale if and only if its expectation at any time coincides with its initial value. Of course, there also exist other ways to address (iii).

To address (i) and (ii), one needs to solve (3.3). However, the latter equation presents numerous difficulties associated with its nonlinear nature and, even more importantly, with the fact that it has "time running in a wrong direction" (cf. [11], for a more detailed discussion of the latter issue). To date, there exist no existence or uniqueness results for the solutions to (3.3) in its general form. Nevertheless, in the next subsection, we choose a specific form of the volatility process a and show how to construct a unique solution to (3.3), for any given (sufficiently regular) strategy π^*, given as a random field. If, in addition, (iii) is resolved and $(\pi_t^*(X_t^*))$ is admissible, we obtain a solution to the optimal contract problem formulated in Subsection 2.2. Indeed, if π^* is the optimal strategy of the principal (i.e. the strategy she would like the agent to follow), the associated $U_T(x)$, normalized appropriately, produces the desired optimal contract.

3.2. Solving the forward performance SPDE. Assume that we are given a random field

$$\pi^* : (\mathbb{R}_+ \times \Omega \times (0,\infty), \mathcal{P} \otimes \mathcal{B}((0,\infty))) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})), \tag{3.6}$$

where $\mathcal{P}$ is the sigma-algebra of progressively measurable sets. As usual, we suppress the dependence upon $\omega \in \Omega$. We assume that π^* is a sufficiently smooth random field, with the precise assumptions stated below. In this subsection, we construct a solution to (3.3), such that (3.4) holds with the given π^*.

Our goal is to reduce (3.3) to a linear SPDE. To this end, we assume that U solves (3.3) and that

$$a_t(x) = a(t, x, U_t, \partial_{xx}^2 U_t) := a_t(\bar{x}) - \lambda_t (U_t(x) - U_t(\bar{x})) - \int_{\bar{x}}^x \sigma_t \pi_t^*(y) \partial_{yy}^2 U_t(y) dy, \tag{3.7}$$

where $\bar{x} > 0$ is a fixed constant, and $(a_t(\bar{x}))_{t \geq 0}$ is an arbitrary locally square integrable process in $\mathbb{R}^d$. With such a choice, we have:

$$\partial_x a_t(x) = -\sigma_t \pi_t^*(x) \partial_{xx}^2 U_t(x) - \partial_x U_t(x) \lambda_t. \tag{3.8}$$

Then, recalling that the columns of σ_t are linearly independent, we obtain

$$\partial_x U_t(x) \lambda_t + (\sigma_t^T)^+ \sigma_t^T \partial_x a_t(x) = -\sigma_t \pi_t^*(x) \partial_{xx}^2 U_t(x), \tag{3.9}$$

and (3.3) becomes

$$\begin{aligned} dU_t(x) &= \frac{1}{2} \|\sigma_t \pi_t^*(x)\|^2 \partial_{xx}^2 U_t(x) dt \\ &+ \left(a_t(\bar{x}) - \lambda_t (U_t(x) - U_t(\bar{x})) - \int_{\bar{x}}^x \sigma_t \pi_t^*(y) \partial_{yy}^2 U_t(y) dy \right)^T dW_t \end{aligned} \tag{3.10}$$

The following derivations (until Assumption 1) are heuristic and are meant to motivate the main result of this subsection, Proposition 3.2. Introducing $V_t(x) := \partial_x U_t(x)$, we differentiate the above equation, to obtain

$$dV_t(x) = \frac{1}{2} \partial_x \left(\|\sigma_t \pi_t^*(x)\|^2 \partial_x V_t(x) \right) dt - \left(\sigma_t \pi_t^*(x) \partial_x V_t(x) + \lambda_t V_t(x) \right)^T dW_t. \tag{3.11}$$

Next, we introduce $R_t(x) := -\partial_x V_t(x) = -\partial_{xx}^2 U_t(x)$, and differentiate the above equation, to obtain

$$\begin{aligned} dR_t(x) = \frac{1}{2} &\left[\partial_x \left(\|\sigma_t \pi_t^*(x)\|^2 \partial_x R_t(x) \right) + \partial_x \left(\|\sigma_t \pi_t^*(x)\|^2 \right) \partial_x R_t(x) \right. \\ &\left. + \partial_{xx}^2 \left(\|\sigma_t \pi_t^*(x)\|^2 \right) R_t(x) \right] dt - \left[\sigma_t \pi_t^*(x) \partial_x R_t(x) + \left(\lambda_t + \sigma_t \partial_x \pi_t^*(x) \right) R_t(x) \right]^T dW_t, \end{aligned} \tag{3.12}$$

with the deterministic initial condition $R_0(x) = -\partial_{xx}^2 U_0(x)$.

In order to solve (3.12), we will apply the results of [8] (see Appendix A for more details). To this end, we need to introduce two assumptions.

Assumption 1. *Let σ and π^*, respectively, be the volatility matrix (defined by (2.1)) and the candidate optimal portfolio (as in (3.6)). We assume that, almost surely, for each $t \geq 0$, the function $\pi_t^*(\cdot)$ is five times continuously differentiable and*

$$\sup_{z\in\mathbb{R}} \left| \sum_{j=1}^{k} \sigma_t^{ij} (\partial_z)^m \left(e^{-z} \pi_t^{*j}(e^z) \right) \right| \leq \xi_t, \quad \forall m = 0, \dots, 5,\ i = 1, \dots, d,$$

for some progressively measurable stochastic process ξ with locally bounded paths.

Remark 6. *The above assumption is purely technical. If π^* arises from a utility maximization problem (which is natural), Assumption 1 implies certain restrictions on the underlying model: i.e., on the growth rate and smoothness of the diffusion coefficients and of the utility function. Nevertheless, from a practical perspective, this assumption is not a serious limitation. Indeed, if the target optimal strategy can be approximated with the ones that satisfy Assumption 1, our algorithm can produce a contract that is ε-optimal for the principal, with arbitrary precision $\varepsilon > 0$.*

For any function $\phi : \mathbb{R} \to \mathbb{R}$, m-times weakly differentiable, we define the norm

$$\|\phi\|_m := \left(\sum_{j=0}^{m} \int_{\mathbb{R}} r^2(z) \left(\phi^{(j)}(z) \right)^2 dz \right)^{1/2},$$

with

$$r(z) := \exp\left(\eta\sqrt{1+z^2} \right), \tag{3.13}$$

with some constant $\eta > 1$. Following [8], we define the weighted Sobolev space $\mathbb{W}^m$ (consisting of m-times weakly differentiable functions from $\mathbb{R}$ to $\mathbb{R}$) as the closure of $C_0^\infty(\mathbb{R})$ in the $\|.\|_m$ norm. We also consider a measurable function $U_0 : (0, \infty) \to \mathbb{R}$.

Assumption 2. *Let λ and U_0, respectively, be the market price of risk (defined by (2.2)) and the candidate initial condition (as above). We assume that $\partial_{xx}^2 U_0(\exp(\cdot)) \in \mathbb{W}^3$, and that $|\lambda|$ has locally integrable paths.*

We now present one of the main results of this paper.

Proposition 3.2. *Let π^*, U_0, σ, and λ, satisfy Assumptions 1 and 2. Then, there exists a unique random field R which solves (3.12), with the initial condition $R_0 = -\partial_{xx}^2 U_0$, and is such that $R_t(\log \cdot)$ takes values in $\mathbb{W}^3$. The random field $R_\cdot(\cdot)$ is almost surely continuous.*

If, in addition, R is strictly positive, then, for any constant $\bar{x} > 0$ and any locally square integrable $\mathbb{R}^d$-valued process $(a_t(\bar{x}))_{t\geq 0}$, the random field $(U_t(x))_{t\geq 0,\, x>0}$, given by

$$U_t(x) = \zeta_t + \int_{\bar{x}}^{x} \int_{y}^{\infty} R_t(z) dz dy, \tag{3.14}$$

with

$$d\zeta_t = -\frac{1}{2} \|\sigma_t \pi_t^*(\bar{x})\|^2 R_t(\bar{x}) dt + a_t^T(\bar{x}) dW_t, \quad \zeta_0 = U_0(\bar{x}),$$

is strictly concave and strictly increasing in x, and satisfies (3.3), with the volatility a given by (3.7). Moreover, for the given π^*, (3.4) holds, and there exists a unique solution to (3.5), for any $(\xi, \tau) \in \mathcal{X}$.

Proof:

First, we transform (3.12) with the simple change of variables, $x = \exp(z)$, introducing $\tilde{R}_t(z) := R_t(e^z)$, and (3.12) becomes

$$
\begin{aligned}
d\tilde{R}_t(z) = \frac{1}{2} \Big[&(\partial_z + 1) \left(\|e^{-z}\sigma_t \pi_t^*(e^z)\|^2 \partial_z \tilde{R}_t(z) \right) + (\partial_z + 2) \left(\|e^{-z}\sigma_t \pi_t^*(e^z)\|^2 \right) \partial_z \tilde{R}_t(z) \\
&+ (\partial_{zz}^2 + 3\partial_z + 2) \left(\|e^{-z}\sigma_t \pi_t^*(e^z)\|^2 \right) \tilde{R}_t(z) \Big] dt \\
- \Big[e^{-z}\sigma_t \pi_t^*(e^z) \partial_z \tilde{R}_t(z) &+ \left(\lambda_t + (\partial_z + 1) \left(e^{-z}\sigma_t \pi_t^*(e^z) \right) \right) \tilde{R}_t(z) \Big]^T dW_t,
\end{aligned}
\tag{3.15}
$$

Notice that the SPDE (3.15) is linear and (degenerate) parabolic. In particular, it belongs to the class of equations analyzed in [8]. For convenience, we provide a summary of the relevant results from [8] in Appendix A. More specifically, we refer to Example 2.2 in [8], and the preceding discussion, to conclude that the conditions of Theorem 2.5 in [8] are satisfied, with $d = 1$, $m = 3$, and $\Gamma = 1$. The latter theorem states that there exists a unique generalized solution $\tilde{R}$ to (3.15), with $\tilde{R}_0(z) = -\partial_{xx}^2 U_0(e^z)$, which is a progressively measurable process with values in $\mathbb{W}^3$, having continuous paths in $\mathbb{W}^2$. Notice that $\tilde{R}_t \in \mathbb{W}^3$ implies that $\tilde{R}_t(.)$ is twice continuously differentiable. Hence, the random field $\tilde{R}_\cdot(\cdot)$ is almost surely continuous, and the spatial derivatives in (3.15) can be understood in the classical sense. Then, changing the variables back to $x = \exp(z)$, we conclude that $R_t(x) := \tilde{R}_t(\log x)$ solves (3.12). Reverting these arguments, we obtain uniqueness of the solution to (3.12).

Next, assume that R is strictly positive. We need to verify that the random field U, defined by (3.14), is well defined and has the desired properties. To this end, we define

$$
V_t(x) = \int_x^\infty R_t(y) dy.
$$

Note that the above integral is well defined due to the choice of r (cf. (3.13)) and the fact that $\tilde{R}_t = R_t(\exp(\cdot))$ takes values in $\mathbb{W}^3 \subset \mathbb{W}^0$:

$$
\int_x^\infty R_t(y) dy = \int_{\log x}^\infty e^z \tilde{R}_t(z) dz \le \left(\int_{\log x}^\infty r^2(z) \tilde{R}_t^2(z) dz \right)^{1/2} \left(\int_{\log x} e^{2z - 2\eta\sqrt{1+z^2}} dz \right)^{1/2} < \infty
$$

Similarly, it is easy to deduce that $\partial_x R_t(\cdot)$ and $\partial_{xx}^2 R_t(\cdot)$ are absolutely integrable over (ε, ∞), for any $\varepsilon > 0$. Applying the stochastic Fubini theorem (cf. Theorem 64 in [13]), we integrate (3.12) to deduce that V satisfies (3.11), with the initial condition $V_0(x) = \partial_x U_0(x)$.[10] Applying stochastic Fubini theorem again, we integrate (3.11), to show that U, defined by (3.14), satisfies the SPDE (3.10). It is clear that $U_t(\cdot)$ is strictly concave, as R is strictly positive. Then, choosing a_t via (3.7), we conclude that U satisfies (3.3). In turn, equation (3.9) yields

[10] Strictly speaking, in order to apply Theorem 64 in [13], we need to localize R and pass to the limit in the integrals over finite domain. We skip these routine arguments for the sake of brevity.

(3.4). Finally, Assumption 1 implies that $\sigma_t \pi_t^*(\cdot)$ is globally Lipschitz, uniformly over (t, ω), which yields the existence and uniqueness of the solution to (3.5), for any initial condition $(\xi, \tau) \in \mathcal{X}$. ∎ None

Remark 7. *Proposition 3.2 can be extended to hold with any positive weight function r, satisfying the condition $(\tilde{W})$ in [8], and such that*

$$\int_x^\infty \frac{e^{2z}}{r^2(z)} dz < \infty, \quad \forall\, x \in \mathbb{R}.$$

Remark 8. *It is straight-forward to formulate the version of Proposition 3.2 for the case where the wealth variable x takes values in $\mathbb{R}$ (as opposed to being restricted to $(0, \infty)$). This would correspond to the investment problems in which the wealth is not restricted to remain positive (cf. Remark 1). We did not find a unifying formulation that would allow us to treat both cases (i.e. $x \in \mathbb{R}$ and $x > 0$) simultaneously, and we chose to consider the case $x > 0$. This choice is motivated by the example in Section 4 which shows that, in the case $x > 0$, in the Black-Scholes model, one can construct explicitly an optimal contract which also satisfies the limited liability condition. Currently, we do not know how to ensure the limited liability condition for the case $x \in \mathbb{R}$, even in the context of this simple example.*

Remark 9. *An alternative description of the solutions to (3.3), using duality methods, is given in [6], [7]. However, the present construction is much shorter and more direct, and it allows us to obtain explicit solutions, as illustrated in Section 4. It is also worth mentioning that the Markovian solutions to (3.3) are analyzed in [11].*

Propositions 3.1 and 3.2 allow us to establish the following characterization of an optimal contract, which is the main result of this paper.

Theorem 3.3. *Consider any initial capital $X_0 > 0$, as well as any λ and U_0, satisfying Assumption 2 and such that $U_0(X_0) > 0$. Assume that a progressively measurable random field π^* and the process σ satisfy Assumption 1, that $(\pi_t^*(X_t^{*,X_0,0})) \in \mathcal{A}(X_0, 0)$, and that*

$$\mathbb{E}J(\pi) \leq \mathbb{E}J(\pi^*),$$

for any π that is C-optimal for some admissible contract C. Let R be the unique solution to (3.12), with the initial condition $R_0 = -\partial_{xx}^2 U_0$. Assume that R is strictly positive and consider the associated U, as in Proposition 3.2, with any constant $\bar{x} > 0$ and any locally square integrable $\mathbb{R}^d$-valued process $(a_t(\bar{x}))_{t\geq 0}$. Then, the following holds.

1. *For any $(\xi, \tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi, \tau)$, the process $\left(U_t\left(X_t^{\pi,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a local supermartingale.*
2. *For any $(\xi, \tau) \in \mathcal{X}$, there exists a unique solution $X^{*,\xi,\tau}$ to (3.5), and the process $\left(U_t\left(X_t^{*,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a local martingale.*
3. *If the aforementioned local martingale and local supermartingales are a true martingale and true supermartingales, respectively, then,*

$$C^*(x) := U_T(x) \frac{u_0}{U_0(X_0)}$$

is an optimal contract.

Proof:

Proposition 3.2 implies that U, a, and π^*, satisfy all the assumptions of Proposition 3.1. The first two statements of the theorem follow immediately. To show the last statement, we notice that the admissibility of $(\pi^*_t(X^*_t))$, the integrability of $C^*(X_T^{\pi,\xi,\tau})$, and the last part of Proposition 3.1, imply that C^* is an admissible contract and that π^* is C^*-optimal. To conclude, consider any admissible contract C and any C-optimal π. Then, we have

$$\mathbb{E}\left[J(\pi) - C\left(X_T^{\pi,X_0,0}\right)\right] \leq \mathbb{E}J(\pi) - u_0 \leq \mathbb{E}J\left(\pi^*\right) - u_0 = \mathbb{E}\left[J\left(\pi^*\right) - C^*\left(X^{*,X_0,0}\right)\right],$$

where the first inequality follows from the admissibility of C and the C-optimality of π, and the second inequality follows from the assumptions of the theorem. ∎ None

Remark 10. *Note that a solution to (3.3) yields an optimal contract only in the case of symmetric information: i.e., when the contract is allowed to be measurable with respect to the full filtration of the agent, $\mathcal{F}_T$. In the asymmetric ("moral hazard") case, the principal's observations may not include the entire $\mathbb{F}$ (e.g., she may not see the prices of some of the assets), hence, the contract needs to be measurable with respect to a sigma-algebra that is strictly smaller than $\mathcal{F}_T$. This leads to the problem of describing all solutions to (3.3) that satisfy an additional measurability constraint at time T. The contract constructed in Theorem 3.3 may not satisfy this measurability constraint, as the random field U_T, constructed in Proposition 3.2, does not satisfy such a constraint.*

In the asymmetric case, the aforementioned measurability constraint implies that not every random field π^*, satisfying Assumption 1, has an admissible associated random field U and, in turn, an admissible contract. Thus, as expected, the asymmetry of information reduces the set of strategies which admit an optimal contract (i.e., which can be optimal for the agent), relative to the symmetric (first-best) case. In particular, the best strategy, from the point of view of the principal, may no longer be attainable. The distance between this strategy and the space of attainable ones – i.e., the strategies associated with the solutions to (3.3) that satisfy the additional measurability constraint – quantifies the loss to the principal due to information asymmetry. However, the problem of establishing a convenient representation of all solutions to (3.3) that satisfy additional measurability constraints, in general, remains open. A related problem is solved in [11], but the latter paper imposes even stronger constraints on the space of admissible solutions to (3.3).

The next section illustrates the application of the above theorem. It describes a specific market model and a concrete contract design problem, for which the present definition of optimal contract is natural, and it shows how to construct an optimal contract explicitly. Moreover, the resulting optimal contract satisfies the limited liability condition: $C \geq 0$ (note, however, that this condition is not guaranteed by Theorem 3.3).

4. Explicit optimal contract in the Black-Scholes model. In this section, we assume that $d = k = 2$, and

$$d\log(S_t^1) = (\mu_1 - \sigma_1^2/2)dt + \sigma_1 dW_t^1,$$
$$d\log(S_t^2) = (\mu_2 - \sigma_2^2/2)dt + \sigma_2(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2),$$

with some $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, and $\rho \in (-1, 1)$. In other words,

$$\sigma = \begin{pmatrix} \sigma_1 & \sigma_2\rho \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \lambda = (\sigma^T)^{-1}\mu = \begin{pmatrix} \mu_1/\sigma_1 \\ \frac{\mu_2-(\sigma_2\rho\mu_1)/\sigma_1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix}.$$

Let us fix a constant $\gamma \in (-\infty, 0) \cup (0, 1)$, whose meaning is explained below. We let $\mathcal{X}$ consist of all pairs (ξ, τ), such that τ is any stopping time with values in $[0, T]$ and $\xi, \xi^\gamma \in L^1 \cap L^0_+(\mathcal{F}_\tau)$. For any $(\xi, \tau) \in \mathcal{X}$, we define $\mathcal{A}(\xi, \tau)$ as the set of all locally integrable processes π, such that the resulting $X^{\pi,\xi,\tau}$ is strictly positive and

$$\mathbb{E} \sup_{t\in[\tau,T]} X_t^{\pi,\xi,\tau} + \mathbb{E} \sup_{t\in[\tau,T]} \left(X_t^{\pi,\xi,\tau}\right)^\gamma < \infty.$$

Next, consider an investor who is looking to hire an agent to manage her initial capital X_0. As discussed in the introduction, we assume that the contract between the agent and the investor is designed by a third party, referred to as the principal (e.g., it can be a regulator, the board of directors of a mutual fund, etc.). The principal chooses an optimal contract using the following individual objective:

$$J(\pi) = \frac{1}{\gamma} \left(X_T^{\pi,X_0,0}\right)^\gamma \mathbf{1}_{\{\pi^2\equiv 0\}} - \infty \cdot (1 - \mathbf{1}_{\{\pi^2\equiv 0\}}), \tag{4.1}$$

where π is a random field and $X^{\pi,X_0,0}$ is the associated wealth. The rationale behind this choice is as follows. The principal assumes (e.g., based on her estimates) that a typical investor uses power utility, with the relative risk aversion $1 - \gamma$, and she adds the constraint that no investment can be made in S^2, as the latter asset is deemed inappropriate (e.g., immoral, subject to sanctions, etc.).

Note that the investor may not be interested in the constraint $\pi^2 \equiv 0$ being met: e.g., in accordance with the assumption of the principal, she may aim to optimize the expected power utility, without the constraint. After the contract is initiated, the investor may have an opportunity to increase the size of her investment, at some stopping time τ, to a random level ξ. As the investor may not care about the constraint $\pi^2 \equiv 0$, a priori, her capital injection may encourage the agent to violate this constraint. Neither the principal nor the agent are aware of the probabilistic properties of (ξ, τ) (i.e., we take the approach of Knightian uncertainty with regards to the opportunities of capital injections). In particular, after any capital injection (ξ, τ), the agent maximizes the expected value of the worst-case future scenario, which corresponds to no future capital injections, since he can always choose to keep the additional capital in cash, which does not decrease his objective value, compared to the case of no additional investment, provided his contract is non-decreasing in terminal wealth (which is the case, as follows from (3.14)). Thus, after every capital injection $(\xi, \tau) \in \mathcal{X}$, the agent solves[11]

$$\max_{\pi\in\mathcal{A}(\xi,\tau)} \mathbb{E}\left(C(X_T^{\pi,\xi,\tau}) \mid \mathcal{F}_\tau\right). \tag{4.2}$$

The regulator's task is two-fold. First, she needs to ensure that the investor is as happy with the contract as possible, given the constraint $\pi^2 \equiv 0$. Namely, the contract should be such that every optimal strategy of the agent maximizes the expectation of (4.1) less the expected payment to the agent, even in the presence of capital injections by the investor. Since these injections are not known to the regulator, she aims to maximize the worst case scenario for the investor, which is the case of no future opportunities for capital injections (as the investor can always choose not to use such an opportunity). This leads to the following objective for the regulator: find admissible contract C^*, such that, for any C^*-optimal π^*, (C^*,π^*) maximizes

$$\mathbb{E}\left[J(\pi) - C(X_T^{\pi,X_0,0})\right] \tag{4.3}$$

among all pairs (C,π) with admissible C and C-optimal π. It is easy to see that, if C^* is an optimal contract, in the sense of Definition 2.2, then it solves the first task of the regulator. The second task of the regulator is to ensure that the investor will not encourage the agent to invest in the second asset by her capital injections. This task is resolved by the admissibility property of an optimal contract C^*: cf. Definitions 2.1 and 2.2. Indeed, the admissibility implies that, after each capital injection, it is still optimal for the agent to follow the optimal strategy (understood as a random field) computed under the assumption of no capital injections. The latter strategy does not invest in S^2, as the pair (C^*,π^*) maximizes the objective (4.3). In the following subsections, we construct an optimal contract C^* explicitly.

4.1. Principal's optimal strategy. Following the solution approach outlined at the beginning of Section 3, we, first, search for a random field π^{*1}, such that

$$(\pi_t^{*1}(X_t^*)) \in \text{argmax}\frac{1}{\gamma}\mathbb{E}\left(X_T^{\pi,X_0,0}\right)^\gamma,$$

where X^* is the associated optimal wealth (starting from X_0 at time zero), and the supremum is taken over all processes π^1, such that $\pi = (\pi^1,0)^T \in \mathcal{A}(X_0,0)$. The wealth process, in this case, satisfies

$$X_0^{\pi,X_0,0} = X_0 \in \mathbb{R}, \quad dX_s^{\pi,X_0,0} = \pi_s^1\sigma_1\lambda_1 ds + \pi_s^1\sigma_1 dW_s^1, \quad s\in[0,T].$$

The solution to the above optimal investment problem is well known, but we briefly outline it here, for the sake of completeness. The associated HJB equation for the value function V is

$$\partial_t V + \max_{\pi^1}(\pi_s^1\sigma_1\lambda_1\partial_x V + \frac{1}{2}(\pi_s^1)^2\sigma_1^2\partial_{xx}^2 V) = 0, \quad x>0,\ s\in(0,T), \quad V(T,x) = x^\gamma/\gamma.$$

[11] We do not assume that only one capital injection may occur. In particular, we allow for an arbitrary finite number of capital injections $\{(\xi_i,\tau_i)\}$. Since, by our modeling assumption, the agent does not anticipate (i.e., does not have a model for) the capital injections, at any given time, he maximizes his expected profits using the worst-case scenario with respect to the injections. It is easy to see that, due to monotonicity of the contract in x, the worst-case scenario is always the one with no future injections. Hence, after each injection, the agent will choose the strategy that solves (4.2), even if more injections will follow (as he is not anticipating them).

This yields

$$V(t,x) = \frac{x^\gamma}{\gamma} \exp\left((T-t)\frac{\lambda_1^2\gamma}{2(1-\gamma)}\right), \quad \pi_t^{*1}(x) = \frac{\lambda_1}{\sigma_1(1-\gamma)}x, \tag{4.4}$$

$$X_0^* = X_0 > 0, \quad dX_s^* = \frac{\lambda_1^2}{1-\gamma}X_s^* ds + \frac{\lambda_1}{1-\gamma}X_s^* dW_s^1, \quad s \in [0,T]. \tag{4.5}$$

A standard verification argument shows that, indeed, V is the value function of the optimization problem, $(\pi_t^{*1}(X_t^*))$ is the optimal policy, and X^* is the optimal wealth (note that X^* is a geometric Brownian motion, hence, $(\pi_t^*(X_t^*)) \in \mathcal{A}(X_0, 0)$). In particular, it follows that

$$J(\pi) \leq J(\pi^*),$$

for any π that is C-optimal for some admissible contract C, with J given by (4.1).

4.2. Fake optimal contracts. Recall that the notion of optimal contract used herein (cf. Definition 2.2) is stronger than usual. The main additional requirement of the present definition is that the contract is robust with respect to capital injections. In this subsection, we show how to construct a (trivial) contract that does not possess this feature, to illustrate the differences.

Recall the optimal wealth process of the principal, X^*, given by (4.5), and consider the following contract:

$$\hat{C}(x) := u_0 \mathbf{1}_{\{X_T^*\}}(x) \tag{4.6}$$

Note that, as long as X_T^* is attainable from the current wealth level, the agent will always aim for X_T^* as the terminal wealth, according to such contract. From the non-degeneracy of the market (i.e. the columns of σ are linearly independent), it follows that the agent will keep following the prescribed strategy $(\pi_t^*(X_t^*))$, given by (4.4), as this is the only strategy that generates X_T^*. As a result, the contract $\hat{C}$ leaves both the principal and the agent satisfied. In fact, the above construction is well known in the optimal contract theory, and it always works for the first-best (risk-sharing) problems. However, the resulting contract $\hat{C}$ is not robust with respect to capital injections. Indeed, if the current wealth level is perturbed, the new set of attainable terminal wealth values may not include X_T^* anymore. In this case, it is not clear which strategy the agent will choose: in fact, in the case of a positive capital injection, the contract will actually provide an incentive for the agent to "lose" (or steal) funds (which, strictly speaking, is not allowed in the model, but can certainly happen in practice). In particular, there is no guarantee that the agent will follow a strategy that is best for the principal after a capital injection is made. One can modify the definition of the "fake" optimal contract (4.6), by using functions other than indicator, and, e.g., obtain contracts that are non-decreasing in the terminal wealth. Nevertheless, such modifications will not resolve the main problem: the agent is not guaranteed to follow the prescribed strategy (viewed as a random field) after a capital injection is made. Fundamentally, this lack of robustness is due to the fact that the contract is not allowed to depend on the capital injections themselves. If such dependence were allowed, the problem would reduce to a standard first-best setting, but at the cost of making it much more difficult to implement the resulting optimal contract.

To conclude this subsection, we illustrate the importance of the fact that the individual objective of the principal, J, given by (4.1), depends on the strategy π in a more general way than through the terminal wealth X_T^π alone. Recall that the principal needs to ensure that the agent's strategy satisfies the constraint $\pi^2 \equiv 0$ (this is what we call an endogenous constraint). Then, if the principal's individual objective were a deterministic function of terminal wealth, e.g.,

$$\tilde{J}(X^\pi) = \frac{1}{\gamma}\,(X_T^\pi)^\gamma\,,$$

we could maximize the expectation of this objective, to obtain an optimal strategy $\tilde{\pi}^*$ (viewed as a random field), and choose the contract

$$\tilde{C}(x) := \tilde{J}(x)\frac{u_0}{\mathbb{E}\,\tilde{J}(X_T^*)}.$$

Note that $\mathbb{E}\,\tilde{J}(X_T^\pi)$ is indeed maximized by the desired optimal strategy π^*. The dynamic programming principle also implies that π^* (as a random field) remains optimal for the agent, for any initial wealth level, and at any starting time. Thus, $\tilde{C}$ would be a (trivial) optimal contract, in the sense of Definition 2.2. Nevertheless, this construction is only possible if the individual objective of the principal depends on π through X_T^π only. Recall, however, that, in the present formulation, $J(\pi)$ depends directly on π, via the constraint $\pi^2 \equiv 0$. Hence, if we use $\mathbb{E}\,(X_T^\pi)^\gamma$ as the objective in the unconstrained problem, faced by the agent, it may not yield the same optimal strategy π^*. Indeed, the optimal contract constructed explicitly in the next subsection does not coincide with the power function with exponent γ; in fact, it becomes a random function of terminal wealth.

4.3. Optimal contract. Recall that $\pi_t^*(x) = (\pi^{*1}x, 0)^T$, with

$$\pi^{*1} = \frac{\lambda_1}{\sigma_1(1-\gamma)},$$

maximizes the individual objective of the principal. Following Proposition 3.2 and Theorem 3.3, we start by solving the SPDE (3.12), which, in the present case, becomes

$$\begin{aligned}dR_t(x) = &\frac{1}{2}\left[\sigma_1^2(\pi^{*1})^2x^2\partial_{xx}^2R_t(x) + 4\sigma_1^2(\pi^{*1})^2x\partial_xR_t(x) + 2\sigma_1^2(\pi^{*1})^2R_t(x)\right]dt\\ &- \left[\sigma_1\pi^{*1}x\partial_xR_t(x) + \left(\lambda_1+\sigma_1\pi^{*1}\right)R_t(x)\right]dW_t^1 - \lambda_2R_t(x)dW_t^2,\end{aligned}$$

With the ansatz $R_t(x) = R(t, x, -W_t^1, -W_t^2)$, the above becomes

$$(\partial_tR + \frac{1}{2}\partial_{yy}^2R + \frac{1}{2}\partial_{zz}^2R)dt - \partial_yRdW_t^1 - \partial_zRdW_t^2$$

$$= \frac{1}{2}\left[\sigma_1^2(\pi^{*1})^2x^2\partial_{xx}^2R + 4\sigma_1^2(\pi^{*1})^2x\partial_xR + 2\sigma_1^2(\pi^{*1})^2R\right]dt$$

$$- \left[\sigma_1 \pi^{*1} x \partial_x R + \left(\lambda_1 + \sigma_1 \pi^{*1}\right) R\right] dW_t^1 - \lambda_2 R dW_t^2,$$

which is equivalent to

$$\partial_t R + \frac{1}{2}\partial_{yy}^2 R + \frac{1}{2}\partial_{zz}^2 R = \frac{1}{2}\sigma_1^2 (\pi^{*1})^2 x^2 \partial_{xx}^2 R + 2\sigma_1^2 (\pi^{*1})^2 x \partial_x R + \sigma_1^2 (\pi^{*1})^2 R,$$

$$\partial_y R = \sigma_1 \pi^{*1} x \partial_x R + \left(\lambda_1 + \sigma_1 \pi^{*1}\right) R, \quad \partial_z R = \lambda_2 R.$$

The following specification solves the above system:

$$R(t,x,y,z) = \tilde{R}(t, \sigma_1 \pi^{*1} y + \log x) e^{(\lambda_1 + \sigma_1 \pi^{*1}) y + \lambda_2 z},$$

$$\partial_t \tilde{R} + A \partial_x \tilde{R} + (A+B)\tilde{R} = 0,$$

$$A := \frac{1}{2}\left(2\lambda_1 \sigma_1 \pi^{*1} - \sigma_1^2 (\pi^{*1})^2\right), \quad B := \frac{1}{2}\left(\lambda_1^2 + \lambda_2^2\right).$$

A specific solution to the above equation is given by

$$\tilde{R}(t,x) = \exp\left(-(B - \varepsilon A)t - (1+\varepsilon)x\right),$$

$$R(t,x,y,z) = \exp\left(-(B - \varepsilon A)t - (1+\varepsilon)\log x + \left(\lambda_1 - \varepsilon \sigma_1 \pi^{*1}\right) y + \lambda_2 z\right),$$

with any $\varepsilon \in (0,1)$. Then,

$$R_t(x) = \frac{1}{x^{1+\varepsilon}} Q_t,$$

where

$$Q_t = \exp\left(-(B - \varepsilon A)t - \left(\lambda_1 - \varepsilon \sigma_1 \pi^{*1}\right) W_t^1 - \lambda_2 W_t^2\right).$$

Let us fix any $X_0^* > 0$, and note that λ, σ, U_0, and π^*, satisfy the assumptions of Theorem 3.3. To complete the construction, we choose $\bar{x} = 1$ and

$$a_t^1(\bar{x}) = -\frac{\lambda_1 - \varepsilon \sigma_1 \pi^{*1}}{\varepsilon(1-\varepsilon)} Q_t, \quad a_t^2(\bar{x}) = -\frac{\lambda_2}{\varepsilon(1-\varepsilon)} Q_t,$$

to obtain

$$U_t(x) = \zeta_t + \int_1^x \int_y^\infty R_t(z)dzdy = \zeta_t + Q_t \frac{1}{\varepsilon}\int_1^x y^{-\varepsilon} dy = \zeta_t + Q_t \frac{1}{\varepsilon(1-\varepsilon)}(x^{1-\varepsilon} - 1), \tag{4.7}$$

with $\zeta_0 = 1$ and

$$d\zeta_t = -\frac{1}{2}\sigma_1^2 (\pi^{*1})^2 Q_t dt - \frac{\lambda_1 - \varepsilon \sigma_1 \pi^{*1}}{\varepsilon(1-\varepsilon)} Q_t dW_t^1 - \frac{\lambda_2}{\varepsilon(1-\varepsilon)} Q_t dW_t^2 = \frac{1}{\varepsilon(1-\varepsilon)} dQ_t.$$

Then

$$U_t(x) = Q_t \frac{1}{\varepsilon(1-\varepsilon)} x^{1-\varepsilon}, \quad C^*(x) = u_0 \left(\frac{x}{X_0}\right)^{1-\varepsilon} Q_T.$$

Notice that such choice of ζ ensures that $U_t(x) \geq 0$, for all $x > 0$ and all (t, ω), thus, satisfying the limited liability condition. In addition, we can express Q_t and, hence, $U_t(x)$, as deterministic functions of the returns of the two assets, S^1 and S^2, at time t:

$$W_t^1 = \frac{1}{\sigma_1} \log(S_t^1/S_0^1) - \lambda_1 t + \frac{\sigma_1}{2} t,$$

$$W_t^2 = \frac{1}{\sigma_2\sqrt{1-\rho^2}} \log(S_t^2/S_0^2) - \frac{\rho}{\sigma_1\sqrt{1-\rho^2}} \log(S_t^1/S_0^1) + \left(\frac{\sigma_2}{2\sqrt{1-\rho^2}} - \frac{\sigma_1 \rho}{2\sqrt{1-\rho^2}} - \lambda_2 \right) t,$$

$$Q_t = \exp\left(\left(\frac{1}{2}(\lambda_1^2 + \lambda_2^2) - \lambda_1 \frac{\sigma_1}{2} + \varepsilon \pi^{*1} \frac{\sigma_1^2}{2}(1 - \pi^{*1}) - \lambda_2 \frac{\sigma_2 - \sigma_1 \rho}{2\sqrt{1-\rho^2}} \right) t \right)$$

$$\times \left(\frac{S_t^1}{S_0^1} \right)^{\varepsilon \pi^{*1} + \frac{\rho \lambda_2}{\sigma_1 \sqrt{1-\rho^2}} - \frac{\lambda_1}{\sigma_1}} \left(\frac{S_t^2}{S_0^2} \right)^{-\frac{\lambda_2}{\sigma_2\sqrt{1-\rho^2}}} := \widehat{Q}\left(t, S_t^2/S_0^2, S_t^3/S_0^3\right).$$

To conclude that C^* is an optimal contract, it remains to verify that the assumptions of the last statement of Theorem 3.3 are satisfied. Note that $U \geq 0$. Part 1 of Theorem 3.3 implies that, for any $(\xi, \tau) \in \mathcal{X}$ and any $\pi \in \mathcal{A}(\xi, \tau)$, the process $\left(U_t\left(X_t^{\pi,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a local supermartingale. As it is nonnegative, and

$$U_\tau(\xi) = \text{const} \cdot Q_\tau \, \xi^{1-\varepsilon} \in L^1,$$

(which follows form Hölder inequality), an application of Fatou's lemma yields that it is a true supermartingale. Next, Part 2 of Theorem 3.3 implies that $\left(U_t\left(X_t^{*,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a local martingale. As it is also positive, we have

$$\mathbb{E} \sup_{t\in[0,T]} \left| U_t\left(X_t^{*,\xi,\tau}\right) \right| \leq \text{const} \cdot \mathbb{E}\left(\xi^{1-\varepsilon} \sup_{t\in[0,T]} \left(Q_t \left(X_t^{*,1,\tau} \right)^{1-\varepsilon} \right) \right) < \infty,$$

which follows, again, from Hölder inequality, by observing that the expression inside the supremum is a geometric Brownian motion. The above inequality implies that $\left(U_t\left(X_t^{*,\xi,\tau}\right)\right)_{t\in[\tau,T]}$ is a true martingale and completes the proof of the fact that C^* is an optimal contract (by Theorem 3.3).

Notice that the optimal contract C^* is given by a power function of terminal wealth multiplied by a random scalar. This is in contrast to the individual objective of the principal, which is a deterministic function of terminal wealth. The random scalar, Q_T, itself, is a power function of the returns generated by the two assets available in the market. Thus, effectively, the optimal contract measures the terminal wealth generated by the agent relative to the performance of the available assets. Note also that the exponents in the latter power functions depend on the characteristics of the assets, such as the market price of risk. Recall also that the optimal contract is nonnegative, thus, satisfying the limited liability condition.

Note also that, as $\varepsilon \approx 0$, the optimal contract converges to u_0 multiplied by

$$\frac{x/X_0}{\widehat{Q}\left(T, S_T^2/S_0^2, S_T^3/S_0^3\right)}.$$

The above ratio measures the return of the fund relative to the returns of the two assets, the latter being captured by $\widehat{Q}$. If this ratio exceeds one (i.e. if the fund outperformance the benchmark), the manager's fee exceeds its initially expected value u_0 (i.e. he receives a bonus). Otherwise, his payment drops below u_0 (i.e. he is penalized).

Finally, it is worth mentioning that the optimal contract C^* is a deterministic function of the terminal values of the wealth process and of the tradable assets. Hence, it is particularly easy to implement.

To conclude this section, it is worth mentioning that the explicit form of the optimal contract derived herein is due to the very simple structure of the target strategy π^* (i.e., it is linear in x) and of the underlying (Black-Scholes) model. To date, we do not have a complete understanding of what is the class of Markovian models in which the SPDE (3.12) has a finite-dimensional realization – i.e., a solution that can be represented as a deterministic function of x and of a finite number of stochastic factors. We leave this important question for future investigation.

5. Appendix A. In this section, we present a summary of the relevant results from [8] on the existence and uniqueness of solutions to linear SPDEs. Many of the expressions and statements are, in fact, more specific corollaries of the results of [8], as we do not require the full power of the latter results in the present paper.

Consider an SPDE of the form

$$du(t,\omega,x) = \Big(\sum_{j=1}^{d_2} \mathcal{M}_j^2 u(t,\omega,x) + \frac{1}{2}\sum_{i=1}^{d_1} \mathcal{N}_i^2 u(t,\omega,x) + \sum_{i=1}^{d_1} \mathcal{N}_i g_i(t,\omega,x) \tag{5.1}$$

$$+\mathcal{M}_0 u(t,\omega,x) + f(t,\omega,x)\Big)\, dt \;+ \sum_{i=1}^{d_1} (\mathcal{N}_i u(t,\omega,x) + g_i(t,\omega,x))\, dW^i(t),$$

$$u(0,\omega,x) = u_0(\omega,x), \tag{5.2}$$

where $(W^1,\dots,W^{d_1})$ is a standard Brownian motion on a stochastic basis $(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\geq 0})$, and $\mathcal{M}_j$, $\mathcal{N}_i$ are first-order differential operators in the space variable $x \in \mathbb{R}$. In particular, we assume:

$$\mathcal{M}_j = \hat{b}_j(t,\omega,x)\frac{\partial}{\partial x} + b_j^0(t,\omega,x), \quad \mathcal{N}_i = \hat{c}_i(t,\omega,x)\frac{\partial}{\partial x} + c_i^0(t,\omega,x).$$

(In the notation of [8], the above corresponds to choosing $d = 1$ and $\Gamma = 1$.)

Consider a function $r : \mathbb{R} \to \mathbb{R}$, satisfying the assumption $(\widetilde{W})$: for any nonnegative integer i,

$$\sup_{x\in\mathbb{R}} \left| r^{-1}(x)\, \partial_x^i\, r(x) \right| < \infty.$$

Example 2.2 in [8] contains a list of functions r that satisfy the condition $(\widetilde{W})$ (note that we are interested in the cases corresponding to $\Gamma = 1$). It contains r given by (3.13), with $\eta = 1$. It is easy to check, however, that $(\widetilde{W})$ is satisfied for any $\eta > 1$.

Next, we introduce assumption (E), stated for any given integer $m \geq 1$.

- The functions $(f(t,\cdot))_{t\geq 0}$, $(g_i(t,\cdot))_{t\geq 0}$ are progressively measurable processes taking values in $\mathbb{W}^m$ and $\mathbb{W}^{m+1}$, respectively, with $\mathbb{W}^m$ being the weighted Sobolev space with norm $\|\cdot\|_m$, defined directly below equation (3.13).
- The function u_0 is an $\mathcal{F}_0$-measurable random element of $\mathbb{W}^m$.
- The coefficients $\hat{b}_i$, $\hat{c}_i$, $\hat{c}_i^0$ are differentiable in $x \in \mathbb{R}$ up to the order $m+1$, and the coefficient $\hat{b}_j^0$ is differentiable up to the order m.
- There exists a progressively measurable stochastic process (ξ_t), such that, for all $\gamma \leq m+1$ and $\beta \leq m$:

$$|\partial_x^\gamma \hat{b}_j|^2 \leq \xi, \quad |\partial_x^\beta \hat{b}_j^0|^2 \leq \xi, \quad j = 1, \ldots, d_2,$$

$$|\partial_x^\gamma \hat{c}_i|^2 \leq \xi, \quad |\partial_x^\gamma \hat{c}_i^0|^2 \leq \xi, \quad i = 1, \ldots, d_1,$$

$$\|f\|_m^2 \leq \xi, \quad \|g_i\|_{m+1}^2 \leq \xi, \quad \|u_0\|_m^2 \leq \xi_0,$$

for all $(t,\omega) \in [0,T] \times \Omega$ and all $x \in \mathbb{R}$.

Theorem 5.1. *(Corollary of Theorem 2.5 in [8]) Assume $(\widetilde{W})$ and (E) (with some integer $m \geq 1$). Then there exists a unique generalized solution u to (5.1)–(5.2). Moreover, $(u(t))_{t\in[0,T]}$ is a weakly continuous $\mathbb{W}^m$-valued process, and it is strongly continuous as a $\mathbb{W}^{m-1}$-valued process.*

REFERENCES

[1] S. Basak, A. Pavlova, and A. Shapiro, *Optimal asset allocation and risk shifting in money management*, The Review of Financial Studies, 20(5) (2007), pp. 1583 – 1621.

[2] A. Cadenillas, J. Cvitanić, and F. Zapatero, *Optimal risk-sharing with effort and project choice*, Journal of Economic Theory, 133(1) (2007), pp. 403 – 440.

[3] J. Cvitanić, D. Possamai, and N. Touzi, *Dynamic programming approach to principal-agent problems*, Finance and Stochastics, 22(1) (2018), pp. 1 – 37.

[4] J. Cvitanić and J. Zhang, *Contract theory in continuous-time models*, Springer, 2012.

[5] B. Holmstrom and B. Milgrom, *Aggregation and linearity in the provision of intertemporal incentivese*, Econometrica, 55(2) (1987), pp. 303 – 328.

[6] N. E. Karoui and M. M'rad, *Stochastic utilities with a given optimal portfolio: approach by stochastic flows*, Preprint, available at arXiv:1004.5192, (2010).

[7] N. E. Karoui and M. M'rad, *An exact connection between two solvable SDEs and a nonlinear utility stochastic PDE*, SIAM Journal on Financial Mathematics, 4 (2013), pp. 697–736.

[8] N. Krylov and I. Gyöngy, *On stochastic partial differential equations with unbounded coefficients*, Potential Analysis, 1 (1992), pp. 233 – 256.

[9] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge University Press, 1990.

[10] M. Musiela and T. Zariphopoulou, *Stochastic partial differential equations in portfolio choice*, in Contemporary Quantitative Finance, C. Chiarella and A. Novikov, eds., Springer-Verlag Berlin Heidelberg, 2010.

[11] S. Nadtochiy and M. Tehranchi, *Optimal investment for all time horizons and Martin boundary of space-time diffusions*, Mathematical Finance, 27(2) (2017), pp. 438 – 470.

[12] H. Ou-Yang, *Optimal contracts in a continuous-time delegated portfolio management problem*, The Review of Financial Studies, 16(1) (2003), pp. 173 – 208.

[13] P. Protter, *Stochastic Integration and Differential Equations*, Springer Verlag, 2nd ed., 2004.

[14] A. Riedl and P. Smeets, *Why do investors hold socially responsible mutual funds?*, The Journal of Finance, 72(6) (2017), pp. 2505 – 2550.

[15] Y. Sannikov, *A continuous-time version of the principal-agent problem*, The Review of Economic Studies, 75(3) (2008), pp. 957 – 984.

[16] L. T. Starks, *Performance incentive fees: An agency theoretic approach*, The Journal of Financial and Quantitative Analysis, 22(1) (1987), pp. 17 – 32.

[17] N. M. Stoughton, *Moral hazard and the portfolio management problem*, The Journal of Finance, 48(5) (1993), pp. 2009 – 2028.

[18] T. Zariphopoulou, *Optimal asset allocation in a stochastic factor model - an overview and open problems*, Advanced Financial Modelling, Radon Series in Computational and Applied Mathematics, 8 (2009), pp. 427 – 453.