Mean-Variance Portfolio Selection in Contagious Markets

Yang Shen[*]   Bin Zou[†]

Abstract

We consider a mean-variance portfolio selection problem in a financial market with contagion risk. The risky assets follow a jump-diffusion model, in which jumps are driven by a multivariate Hawkes process with mutual-excitation effect. The mutual-excitation feature of the Hawkes process captures the contagion risk in the sense that each price jump of an asset increases the likelihood of future jumps not only in the same asset but also in other assets. We apply the stochastic maximum principle, backward stochastic differential equation theory, and linear-quadratic control technique to solve the problem and obtain the efficient strategy and efficient frontier in semi-closed form, subject to a non-local partial differential equation. Numerical examples are provided to illustrate our results.

1 Introduction

Asset prices exhibit jumps, occasionally and persistently, in all financial markets across the world, which has been well documented and empirically tested in the literature. Large price movements are unlikely to be observed under standard financial models driven solely by Brownian motion(s), e.g., the Black-Scholes model. The most popular models incorporating jumps are the jump-diffusion models, with the jump part predominantly driven by a Poisson process or a more general Lévy process. Those models have enjoyed great

---

[*]School of Risk and Actuarial Studies and CEPAR, University of New South Wales, Sydney, NSW 2052, Australia. Email: y.shen@unsw.edu.au

[†]Department of Mathematics, University of Connecticut, Storrs, CT, 06269-1069, USA. Email: bin.zou@uconn.edu

popularity in option pricing, term structure and credit risk modelling, and other applications. One may refer to the survey article Kou (2007), the monographs Cont and Tankov (2003) and Øksendal and Sulem (2019), and the references therein for detailed discussions of those models.

More strikingly, empirical studies confirm that the price jumps of an asset (or a class of assets) are likely to be accompanied by more jumps, over a short time period, from not only the same (class of) asset but also different (classes of) assets, creating a contagion or clustering effect in the market (see Figure 1 for an example). This contagion effect is observed across different markets globally, in particular, during financial crises. The cascade of market declines experienced during the global financial crisis of 2007-2008 is a prominent example (see Aït-Sahalia et al. (2015) and Azizpour et al. (2018) for more examples). As pointed out in Aït-Sahalia et al. (2015), jumps under a standard Lévy jump-diffusion model are rare events, and, as a result, the clustering of large price movements cannot be properly explained using a standard jump-diffusion model. The interplay between jumps from different assets or markets is complex, leaving the *independent* increment assumption of Lévy processes vulnerable.
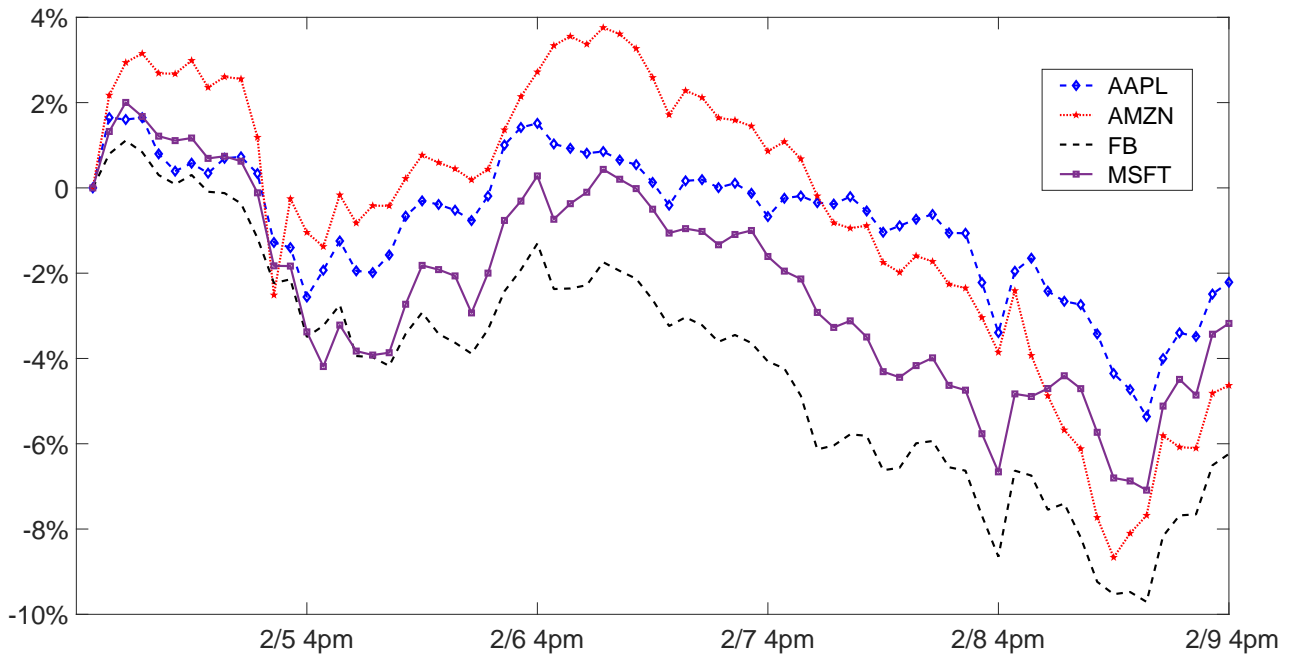

Figure 1. We plot the price changes in *percentages* for AAPL (Apple), AMZN (Amazon), FB (Facebook) and MSFT (Microsoft) from 2/5/2018 9:30 am to 2/9/2018 4:00 pm during the 2018 February sell-off in the U.S. stock market. The price data are in frequency 30m (minutes), and are taken to be the average of the highest and the lowest in each 30m window. Data are downloaded from Yahoo!Finance.

Numerous approaches and models have been proposed to account for the clustering effect of jumps (extreme events) in finance and beyond. Among them includes a successful candidate, *Hawkes processes*, first

introduced by Hawkes (1971). In a multivariate Hawkes process $N = (N_1, N_2, \ldots, N_m)^\top$, the occurrence of a jump from one component (say N_1) raises not only the jump intensity of itself (self-excitation feature) but also those of other components $N_2, \ldots, N_m$ (cross-excitation feature). The joint effect of self-excitation and cross-excitation, called *mutual-excitation*, increases the likelihood of seeing more jumps both over small time period and from different components, which provides a convincing explanation to the clustering of jumps observed in the financial markets. An early paper using Hawkes processes to model financial data is Bowsher (2007). (Its working paper version first appeared online in 2002.) More recent works that apply Hawkes processes to financial modeling include Embrechts et al. (2011), Chavez-Demoulin and McGill (2012), Aït-Sahalia et al. (2015), and Azizpour et al. (2018). Hawkes processes have also been applied to risk analysis, credit risk, derivatives pricing, and other financial topics. See Chavez-Demoulin et al. (2005), Errais et al. (2010), Dassios and Zhao (2011), Aït-Sahalia et al. (2014), and Zhang et al. (2018), among many others. In actuarial science, Dassios and Zhao (2012) and Zhu (2013) analyze ruin probabilities when the arrivals of claims are modeled by a Hawkes process. We refer to the survey papers of Bacry et al. (2015) and Hawkes (2018) for the overview and progress of Hawkes processes in finance.

In this paper, we consider the classical mean-variance (MV) portfolio selection problem (pioneered by Markowitz (1952)) in a financial market with contagion risk. The asset prices are modeled by a multi-dimensional jump-diffusion model, in which self-excited and cross-excited jumps are driven by a multivariate Hawkes process. The mutual-excitation feature of the Hawkes process captures the clustering and contagion effects in the financial market. It is well known that the MV portfolio selection problem is a *time inconsistent* control problem, in the sense that an optimal strategy found at the current time may cease to be optimal as time evolves. Two lines of research are developed to tackle such a problem. The first method is to treat the problem as a *precommitted* problem and solve for an optimal strategy; see Bajeux-Besnainou and Portait (1998), Li and Ng (2000), Zhou and Li (2000), Lim and Zhou (2002), Zhou and Yin (2003), and Yin and Zhou (2004), among many others. The second approach is to treat the problem under a game theoretical framework and seek an equilibrium strategy; see for instance Basak and Chabakauri (2010) and Björk et al. (2014). Shen and Zou (2021) provide detailed comparisons between these two approaches. We adopt the first precommitted framework in this paper, namely, we assume that once an optimal strategy is obtained, the agent will commit herself to this strategy throughout the entire investment horizon.

Although Hawkes processes have enjoyed great success in financial modeling, studies on control problems with systems driven by Hawkes processes are still rare, especially when comparing to those driven by Brownian motions and/or Poisson jump processes. To our knowledge, Aït-Sahalia and Hurd (2016) is the earliest paper which considers the portfolio selection problem in a contagious market driven by a Hawkes process. Another related paper is Bo et al. (2019), which solves a portfolio optimization problem under a defaultable market modeled by mean-reverting diffusion processes enhanced with self-excitation. Cao et al.

(2020) apply a Hawkes process to model the claim frequency for an insurer and seek the optimal strategy to the insurer's optimal reinsurance problem. Liu et al. (2021) study an optimal investment, consumption, and life insurance problem in a self-contagious market (i.e., there is only one risky asset) for a utility-maximizing household with a bequest motive.

We summarize the key contributions and findings of this paper as follows:

- This paper contributes to the stochastic control literature when the controlled system is driven by a Brownian motion and a Hawkes process. In particular, in the area of portfolio selection problems, ours shall follow the work of Aït-Sahalia and Hurd (2016) by modeling the contagion risk via a Hawkes process, but with significant differences. The objective in Aït-Sahalia and Hurd (2016) is to maximize the expected utility of consumption over an infinite time horizon, while we consider the MV problem over a finite horizon. The difference in methodology is further elaborated below.

- In all the related papers mentioned above, the standard dynamic programming principle is used and the optimal strategy is found by solving the associated Hamilton-Jacobi-Bellman (HJB) equation. In comparison, we apply the stochastic maximum principle, backward stochastic differential equation (BSDE) theory, and linear-quadratic (LQ) control technique to solve our MV portfolio selection problem. Given that the stochastic maximum principle can be applied to deal with non-Markovian[1] control systems with random coefficients, our approach has the potential to tackle portfolio selection problems under a more general setup with non-Markovian Hawkes processes (see Gao et al. (2018)).

- Numerical findings: In the numerical analysis of a univariate example, we find that as the initial intensity value, mean-reversion level, or jump size of the intensity process increases, the efficient frontier deteriorates and the MV investor is worse off. On the contrary, a larger mean-reversion speed improves the efficient frontier. Compared with the Poisson-jump-diffusion model with the same initial intensity, the Hawkes model may lead to either an improved or deteriorated efficient frontier, depending on the relative size of the initial intensity and the mean-reversion level in the Hawkes model.

The rest of the article is organized as follows. In Section 2, we introduce a contagious financial market and present the MV portfolio selection problem. In Section 3, we carry out heuristic derivations to solve an equivalent problem using the stochastic maximum principle and BSDE theory. In Section 4, we apply the completing the square technique in LQ theory and formally obtain the efficient (optimal) strategy and the efficient frontier of the MV problem. We conduct numerical analysis in Section 5 and summarize concluding remarks in Section 6. Technical proofs are collected in Appendixes A-C. Particularly, Appendix

---

[1]The control systems are (joint) Markovian in Aït-Sahalia and Hurd (2016), Cao et al. (2020), and Liu et al. (2021), while Gao et al. (2018) study a self-exciting non-Markovian Hawkes process.

C is devoted to the discussion of the connection between the stochastic maximum principle and the HJB equation approach.

2 Problem formulation

In this section, we first introduce a multi-dimensional Hawkes process to model mutually exciting jumps, which capture the contagion risk in the financial market. Then, we formulate the mean-variance portfolio selection problem in (2.13).

2.1 Probabilistic setup and Hawkes jump modeling

Let $[0, T]$ be a finite time horizon, where $T < \infty$ represents the terminal time of planning. We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t \in [0,T]}$, which is assumed to carry all the random objects considered in the sequel and satisfy the usual conditions of right-continuity and $\mathbb{P}$-completeness. The operator $\mathbb{E}$ (resp. Var) denotes taking expectation (resp. variance) under $\mathbb{P}$, and $\mathbb{E}_t := \mathbb{E}[\cdot | \mathcal{F}_t]$. On this probability space, two stochastic processes are defined: an n-dimensional standard Brownian motion $W := \{(W_1(t), W_2(t), \ldots, W_n(t))^\top\}_{t \in [0,T]}$ and an m-dimensional càdlàg (right continuous with left limits) point process $N := \{(N_1(t), N_2(t), \ldots, N_m(t))^\top\}_{t \in [0,T]}$. Here, n and m are two positive integers, and $^\top$ denotes the usual transpose operator on a vector or a matrix. Let k be another positive integer, we define three index sets by[2]

$$\mathbb{N} := \{1, 2, \ldots, n\}, \qquad \mathbb{M} := \{1, 2, \ldots, m\}, \qquad \text{and} \qquad \mathbb{K} := \{1, 2, \ldots, k\}. \tag{2.1}$$

Let us denote the intensity process of N by $\lambda := \{(\lambda_1(t), \lambda_2(t), \ldots, \lambda_m(t))^\top\}_{t \in [0,T]}$, where $\lambda_l(t)$ is the corresponding instantaneous intensity of N_l at time t for each $l \in \mathbb{M}$ and $t \in [0, T]$. Hereinafter, we work with the càglàd version of λ, i.e., λ is left continuous with right limits. Heuristically, we have (for rigorous definitions see Daley and Vere-Jones (2007))

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathbb{E}_t \left[N(t + \Delta t) - N(t) \right]}{\Delta t}, \qquad t \in [0, T].$$

In addition, we define a new process $\widetilde{N} := \{(\widetilde{N}_1(t), \widetilde{N}_2(t), \ldots, \widetilde{N}_m(t))^\top\}_{t \in [0,T]}$, commonly referred to as the compensated process of N, by

$$\widetilde{N}(t) := N(t) - \int_0^t \lambda(u) \, \mathrm{d}u, \qquad t \in [0, T].$$

It is well known that $\widetilde{N}$ is an m-dimensional $(\mathbb{F}, \mathbb{P})$-local martingale.

---

[2]Set $\mathbb{N}$ is the index set of n components of the Brownian motion W, set $\mathbb{M}$ is the index set of m components of the Hawkes process N, and $\mathbb{K}$ is the index set of k risky assets (introduced in the next subsection).

In this paper, we use a Hawkes process, introduced in Hawkes (1971), to model the point process N. In particular, we assume that the intensity process λ_l of N_l is governed by

$$d\lambda_l(t) = \alpha_l(\lambda_{l\infty} - \lambda_l(t))\,dt + \sum_{j=1}^{m} \beta_{lj}\,dN_j(t)$$

$$= \left[\alpha_l(\lambda_{l\infty} - \lambda_l(t)) + \sum_{j=1}^{m} \beta_{lj}\lambda_j(t)\right]dt + \sum_{j=1}^{m} \beta_{lj}\,d\widetilde{N}_j(t), \tag{2.2}$$

where $\lambda_l(0) = \lambda_{l0}$ and $l \in \mathbb{M}$. We assume that all the coefficients in (2.2) are constants, and $\lambda_{l0} > 0$, $\alpha_l > 0$, $\lambda_{l\infty} \geq 0$ and $\beta_{lj} \geq 0$, for all $l, j \in \mathbb{M}$. The dynamics equation given by (2.2) implies that the intensity process λ_l mean reverts to the long-term rate $\lambda_{l\infty}$ at speed α_l, and jumps up by size β_{lj} whenever process N_j jumps, where $j \in \mathbb{M}$. Alternatively, we can solve (2.2) and represent λ_l by

$$\lambda_l(t) = e^{-\alpha_l t}\lambda_{l0} + \left(1 - e^{-\alpha_l t}\right)\lambda_{l\infty} + \sum_{j=1}^{m} \int_0^{t-} \beta_{lj}e^{-\alpha_l(t-s)}dN_j(s), \qquad t \in [0,T],\, l \in \mathbb{M}.$$

Obviously, N is an m-dimensional Hawkes process with *exponential* decay.

With λ_l satisfying (2.2) for all $l \in \mathbb{M}$, we write the dynamics of λ in the following vector form

$$d\lambda(t) = \alpha(\lambda_\infty - \lambda(t))\,dt + \beta\,dN(t) = \left(\alpha\lambda_\infty + (\beta - \alpha)\lambda(t)\right)dt + \beta\,d\widetilde{N}(t), \tag{2.3}$$

where we denote

$$\begin{aligned}
\lambda(0) = \lambda_0 &:= (\lambda_{10}, \lambda_{20}, \ldots, \lambda_{m0})^\top, \qquad \lambda_\infty := (\lambda_{1\infty}, \lambda_{2\infty}, \ldots, \lambda_{m\infty})^\top, \\
\alpha &:= \mathrm{Diag}[(\alpha_1, \alpha_2, \ldots, \alpha_m)^\top], \qquad \beta := [\beta_{lj}]_{l,j \in \mathbb{M}}.
\end{aligned} \tag{2.4}$$

Here, for a vector v, $\mathrm{Diag}[v]$ denotes the square matrix, whose diagonal vector is equal to v and other entries are zero.

Remark 2.1. *As can be easily seen from* (2.3), *a jump of N_l at time t not only increases its own instantaneous intensity $\lambda_l(t)$ by β_{ll} (self-excitation) but also increases the instantaneous intensities $\lambda_j(t)$ of other processes N_j by β_{jl} (cross-excitation), where $j \neq l, j \in \mathbb{M}$. Furthermore, a jump occurring at time t has a permanent impact on the intensity process $\lambda_l(s)$ for all $s \geq t$ and $l \in \mathbb{M}$, although such an impact decays exponentially at rate α_l. As a result, the intensity process λ depends on the past history of N, making N a path-dependent process. The dynamics of λ in (2.3) indicate that (N, λ) is a joint Markov process. The above Hawkes jump model is also used in Aït-Sahalia and Hurd (2016), Cao et al. (2020), and Liu et al. (2021).*

2.2 Financial market with contagion risk

We consider a financial market that consists of one risk-free asset (e.g., savings account) and k risky assets (e.g., stocks). The risk-free asset earns interest continuously at a constant rate $r > 0$, and its price evolves according to

$$\mathrm{d}S_0(t) = rS_0(t)\,\mathrm{d}t, \qquad S_0(0) = 1.$$

The price processes of the risky assets are given by

$$\mathrm{d}S_i(t) = S_i(t-)\left[\mu_i\,\mathrm{d}t + \sum_{j=1}^{n}\sigma_{ij}\,\mathrm{d}W_j(t) + \sum_{l=1}^{m}J_{il}\Big(Z_l(t)\,\mathrm{d}N_l(t) - \mathbb{E}[Z_l(t)]\lambda_l(t)\,\mathrm{d}t\Big)\right], \qquad i \in \mathbb{K}, \quad (2.5)$$

where $S_i(0) = s_{i0} > 0$ and N_l is the lth entry of the m-dimensional Hawkes process N introduced in Section 2.1, for $l \in \mathbb{M}$. Recall set $\mathbb{K} = \{1, 2, \ldots, k\}$ (see (2.1)). The assumptions of model (2.5) will be presented shortly after a brief introduction of notations. Our model (2.5) is similar to that of Aït-Sahalia and Hurd (2016) (see Eq.(2.2) in their paper), and both consider exponential decay for the intensity process (see Eq.(2.2) or (2.3) in ours and Eq.(2.3) in theirs).

For notational simplicity, we introduce the expected return rate vector μ, the risk premium vector B, the volatility matrix (of dimension $k \times n$) σ, where

$$\mu := (\mu_1, \mu_2, \ldots, \mu_k)^\top, \quad B := (\mu_1 - r, \mu_2 - r, \ldots, \mu_k - r)^\top, \quad \sigma := [\sigma_{ij}]_{i \in \mathbb{K},\, j \in \mathbb{N}}, \qquad (2.6)$$

and the jump size matrix (of dimension $k \times m$) η, where

$$\eta(Z(t)) := [J_{il}Z_l(t)]_{i \in \mathbb{K},\, l \in \mathbb{M}}, \qquad \forall\, t \in [0, T]. \qquad (2.7)$$

We summarize below the model assumptions that will be imposed throughout the paper.

Assumption 2.1. *In the market model* (2.5)*, the drift $\mu_i > 0$, the volatility rate $\sigma_{ij} > 0$, and the scaling factor $J_{il} \in [0, 1]$ are all constants, for all $i \in \mathbb{K}$, $j \in \mathbb{N}$, and $l \in \mathbb{M}$. The jump size $Z_l = \{Z_l(t)\}_{t \in [0,T]}$ is a series of independent and identically distributed (i.i.d.) random variables, with common probability measure ν_l supported on $(-1, \infty)$ and finite second moment, for all $l \in \mathbb{M}$. Let us denote $\nu := (\nu_1, \nu_2, \ldots, \nu_l)^\top$. We assume the following non-degenerate condition holds true for both the variance-covariance matrix $\sigma\sigma^\top$ and the precision matrix $(\sigma\sigma^\top)^{-1}$ (see for instance Zhou and Yin (2003) for similar assumptions):*

$$\zeta^\top \sigma\sigma^\top \zeta \geq \epsilon|\zeta|^2, \quad \zeta^\top (\sigma\sigma^\top)^{-1}\zeta \geq \epsilon|\zeta|^2, \qquad \forall\, \zeta \in \mathbb{R}^k,$$

where ϵ is a positive constant. The jumps in asset prices are modeled by the m-dimensional Hawkes process N, whose intensity process λ is given by (2.3)*. Furthermore, we suppose that the Brownian motion W, the Hawkes process N, and the jump size random variables $Z = \{Z_l\}_{l \in \mathbb{M}}$ are stochastically independent of each other. The filtration $\mathbb{F}$ is the augmented filtration generated by W, N, and Z.*

Let u and v be two arbitrary vectors with the same dimension (say n), we define operator $\bullet$ by

$$u \bullet v := (u_1 v_1, u_2 v_2, \ldots, u_n v_n)^\top. \tag{2.8}$$

Lemma 2.1. *The generalized variance-covariance matrix, defined by*

$$\Sigma(t) := \sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z) \operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)] \eta(z)^\top \tag{2.9}$$

is non-degenerate and positive definite. Moreover, $\Sigma(t)^{-1}$ is positive semi-definite.

Proof. We first note that, in the definition of (2.9), σ and η are given by (2.6) and (2.7), ν is the distribution function of jump random variables Z, and λ is the intensity process, which is positive and governed by (2.3). $\operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]$ denotes the square $(m \times m)$ matrix with diagonal entries equal to $\lambda(t) \bullet \nu(\mathrm{d}z)$, where

$$\lambda(t) \bullet \nu(\mathrm{d}z) = \big(\lambda_1(t)\nu_1(\mathrm{d}z_1), \lambda_2(t)\nu_2(\mathrm{d}z_2), \ldots, \lambda_m(t)\nu_m(\mathrm{d}z_m)\big)^\top,$$

and all other entries are equal to 0.

By simple algebra, we obtain, for any $\zeta \in \mathbb{R}^k$, that

$$\int_{(-1,\infty)^m} \zeta^\top \eta(z) \operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)] \eta(z)^\top \zeta = \sum_{l=1}^m \int_{(-1,\infty)} \big[(\zeta^\top \eta(z))_l\big]^2 \lambda_l(t)\, \nu_l(\mathrm{d}z_l) \geq 0,$$

where $(\zeta^\top \eta(z))_l$ denotes the lth entry of the m-dimensional vector $\zeta^\top \eta(z)$, for all $l \in \mathbb{M}$. Thus, by the non-degenerate condition in Assumption 2.1, we have

$$\zeta^\top \Sigma(t)\zeta \geq \zeta^\top \sigma\sigma^\top \zeta \geq \epsilon|\zeta|^2, \qquad \forall \zeta \in \mathbb{R}^k,$$

which implies $\Sigma(t)$ is non-degenerate and positive definite. The positive semi-definiteness of $\Sigma(t)^{-1}$ also holds true because

$$\zeta^\top \Sigma(t)^{-1}\zeta = (\zeta^\top \Sigma(t)^{-1})\, \Sigma(t)\, (\Sigma(t)^{-1}\zeta) \geq \epsilon|\Sigma(t)^{-1}\zeta|^2 \geq 0, \qquad \forall \zeta \in \mathbb{R}^k.$$

The proof is then completed. $\qquad\square$

Remark 2.2. *The market model (2.5) naturally inherits both self-excitation and cross-excitation features from the Hawkes process N (together called mutual-excitation). Such a financial market is contagious in the sense that, a jump from one source will increase the jump intensity of all the sources and then likely lead to more jumps from different sources in the near future, creating contagion effects and clustering phenomena. We also point out that, by imposing $J_{il} \in [0,1]$ and $Z_l \in (-1,\infty)$, the asset price S_i stays strictly positive for all $i \in \mathbb{K}$. Aït-Sahalia and Hurd (2016) focus on negative jumps only and further assume $J_{il}Z_l \in (-1,0)$ for all $i \in \mathbb{K}$ and $l \in \mathbb{M}$.*

Moreover, we follow the standard notation of random measures (see, e.g., Eq.(1.1.2) in Øksendal and Sulem (2019)) and define (with slight abuse of notations)

$$N_l(t, \mathbb{I}) := \sum_{0 < s \leq t} \chi_{\mathbb{I}}(\Delta N_l(s)), \qquad \forall l \in \mathbb{M}, \, t \in [0, T],$$

where $\mathbb{I} \subset (-1, \infty)$, $0 \notin \bar{\mathbb{I}}$ (the closure of $\mathbb{I}$), and χ is an indicator function. Let us define

$$\widetilde{N}_l(\mathrm{d}t, \mathrm{d}z_l) := N_l(\mathrm{d}t, \mathrm{d}z_l) - \lambda_l(t)\, \nu_l(\mathrm{d}z_l)\, \mathrm{d}t, \qquad l \in \mathbb{M},$$

$$\text{and} \qquad \widetilde{N}(\mathrm{d}t, \mathrm{d}z) := \left(\widetilde{N}_1(\mathrm{d}t, \mathrm{d}z_1), \widetilde{N}_2(\mathrm{d}t, \mathrm{d}z_2), \ldots, \widetilde{N}_m(\mathrm{d}t, \mathrm{d}z_m) \right)^{\top}. \tag{2.10}$$

By using the random measure notation, we naturally have

$$\int_0^t J_{il}\left(Z_l(t)\, \mathrm{d}N_l(t) - \mathbb{E}[Z_l(t)]\lambda_l(t)\, \mathrm{d}t \right) = \int_0^t \int_{(-1,\infty)} J_{il} z_l\, \widetilde{N}_l(\mathrm{d}t, \mathrm{d}z_l), \qquad \forall t \in [0, T], \, l \in \mathbb{M}.$$

We can also rewrite the market model (2.5) in the following vector form

$$\mathrm{d}S(t) = \mathrm{Diag}[S(t-)] \left(\mu\, \mathrm{d}t + \sigma\, \mathrm{d}W(t) + \int_{(-1,\infty)^m} \eta(z)\, \widetilde{N}(\mathrm{d}t, \mathrm{d}z) \right). \tag{2.11}$$

We consider a representative agent (investor), who is a price-taker in the economy (i.e., her trading does not move the prices of the assets). In the financial market with contagion risk, as described above, the agent chooses her investment strategy $\pi := \{(\pi_1(t), \pi_2(t), \ldots, \pi_k(t))^{\top}\}_{t \in [0,T]}$, where $\pi_i(t)$ represents the *dollar amount* invested in the ith risky asset at time t. Let us denote by $X^{\pi} := \{X^{\pi}(t)\}_{t \in [0,T]}$ the wealth process associated with strategy π. We consider self-financing strategies only, that means the remaining amount of $X^{\pi}(t) - \sum_{i=1}^k \pi_i(t)$ is fully invested in the risk-free asset at time t. We shall write $X(t) := X^{\pi}(t)$ for simplicity, whenever there is no risk of confusion. Using (2.11), we obtain the dynamics of X by

$$\mathrm{d}X(t) = \left(rX(t) + \pi(t)^{\top} B \right) \mathrm{d}t + \pi(t)^{\top} \sigma\, \mathrm{d}W(t) + \int_{(-1,\infty)^m} \pi(t)^{\top} \eta(z) \widetilde{N}(\mathrm{d}t, \mathrm{d}z), \tag{2.12}$$

where $X(0) = x_0 > 0$ is the agent's initial wealth, the risk premium B and volatility matrix σ are given by (2.6), and η is the jump size matrix defined by (2.7).

Before introducing the MV problem, we define three spaces of processes that will be used in the subsequent analysis.

Definition 2.1. *Denote by $\mathcal{P}$ the $\mathbb{F}$-predictable σ-field on $\Omega \times [0, T]$. We define*

- $\mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^k)$ *as the space of all $\mathbb{R}^k$-valued, $\mathbb{F}$-predictable processes $\varphi = \{\varphi(t)\}_{t \in [0,T]}$ such that*

$$\mathbb{E}\left[\int_0^T |\varphi(t)|^2\, \mathrm{d}t \right] < \infty;$$

9

- $\mathcal{L}_{\mathcal{F}}^{2,N}(0,T;\mathbb{R}^m)$ *as the space of all $\mathbb{R}^m$-valued, $\mathcal{P} \times \mathcal{B}((-1,\infty)^m)$-measurable processes $\varphi = \{\varphi(t,z)\}$ such that*

$$\mathbb{E}\left[\int_0^T |\varphi(t,\cdot)|_N^2 \, \mathrm{d}t\right] < \infty,$$

where $\mathcal{B}$ denotes the Borel set, $\varphi(t,z) := (\varphi_1(t,z_1), \varphi_2(t,z_1), \ldots, \varphi_m(t,z_m))^\top$, with $(t,z) \in [0,T] \times (-1,\infty)^m$, and

$$|\varphi(t,\cdot)|_N^2 := \sum_{l=1}^m \int_{(-1,\infty)} |\varphi_l(t,z_l)|^2 \lambda_l(t)\nu_l(\mathrm{d}z_l);$$

- $\mathcal{S}_{\mathcal{F}}^2(0,T;\mathbb{R})$ *as the space of all real-valued, $\mathbb{F}$-adapted, càdlàg processes $\varphi = \{\varphi(t)\}_{t\in[0,T]}$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\varphi(t)|^2\right] < \infty.$$

As is standard in the literature (see Definition 2.1 in Zhou and Yin (2003)), we need to impose certain integrability conditions on the investment strategy π, which are listed in the definition below.

Definition 2.2. *An investment strategy π is said to be admissible if π is self-financing, $\pi \in \mathcal{L}_{\mathcal{F}}^2(0,T;\mathbb{R}^k)$ and $\eta^\top \pi \in \mathcal{L}_{\mathcal{F}}^{2,N}(0,T;\mathbb{R}^m)$, where spaces $\mathcal{L}_{\mathcal{F}}^2(0,T;\mathbb{R}^k)$ and $\mathcal{L}_{\mathcal{F}}^{2,N}(0,T;\mathbb{R}^m)$ are defined in Definition 2.1. Denote by $\mathcal{A}$ the set of all admissible strategies.*

Remark 2.3. *For any $\pi \in \mathcal{A}$, the stochastic differential equation (SDE) of X given by (2.12) admits a unique strong solution X, which belongs to the space $\mathcal{S}_{\mathcal{F}}^2(0,T;\mathbb{R})$ (see Definition 2.1). In consequence, we have $\mathbb{E}[X(T)] < \infty$ and $\mathbb{E}[X^2(T)] < \infty$, implying that the variance $\mathrm{Var}[X(T)]$ is well defined. However, for utility maximization problems, an additional requirement of X being bounded from below (e.g., $X > 0$ a.s.) is often imposed to exclude certain strategies (e.g., doubling strategies) and make the problems well posed (see Karatzas and Shreve (1998)[Definition 3.2]).*

We now proceed to formulate the MV portfolio selection problem that is of pivotal interest to this paper.

Problem 2.1. *The agent seeks to solve the following mean-variance portfolio selection problem:*

$$\begin{cases} \min_{\pi\in\mathcal{A}} \ J(x_0,\lambda_0;\pi) := \min_{\pi\in\mathcal{A}} \ \mathbb{E}[(X(T)-\xi)^2] \\ \\ subject \ to \ \begin{cases} \mathbb{E}[X(T)] = \xi, \\ X \text{ is the solution to (2.12)}, \end{cases} \end{cases} \qquad (2.13)$$

where x_0 is the agent's initial wealth and λ_0 is the initial value of the intensity process λ given by (2.3). We call a solution π^* to the above problem an optimal (investment) strategy or an efficient strategy. To avoid trivial cases, we set $\xi \geq x_0 e^{rT}$.

Remark 2.4. *The formulation of problem* (2.13) *is the same as that in Lim and Zhou (2002) and Zhou and Yin (2003). In Zhou and Li (2000), an alternative formulation of the MV portfolio selection problem is considered:*

$$\min_{\pi \in \mathcal{A}} \left\{ -\mathbb{E}[X(T)] + \gamma \mathrm{Var}(X(T)) \right\}, \qquad \gamma > 0.$$

Due to the variance term, this alternative one is not a standard control problem, although it does not involve expectation constraint. A standard approach to overcoming the difficulty caused by the variance term is the embedding technique developed in Li and Ng (2000)[Theorem 2] and Zhou and Li (2000)[Theorem 3.1].

Notice that problem (2.13) is a constrained stochastic optimization problem, with an equality constraint. To solve this problem, we consider a modified objective J_1, defined by

$$J_1(x_0, \lambda_0; \pi, \theta) := \mathbb{E}\big[(X(T) - \xi)^2\big] + 2\theta\big(\mathbb{E}[X(T)] - \xi\big) = \mathbb{E}\big[(X(T) - (\xi - \theta))^2\big] - \theta^2,$$

where $\theta \in \mathbb{R}$ is the Lagrange multiplier. Please refer to Lim and Zhou (2002) and Zhou and Yin (2003) for similar ideas. As discussed in these two papers, it follows from the Lagrangian duality theorem that solving the MV portfolio selection problem (2.13) is equivalent to solving the following max-min problem:

$$\begin{cases} \max\limits_{\theta \in \mathbb{R}} \min\limits_{\pi \in \mathcal{A}} \ J_1(x_0, \lambda_0; \pi, \theta) = \max\limits_{\theta \in \mathbb{R}} \min\limits_{\pi \in \mathcal{A}} \left\{ \mathbb{E}\big[(X(T) - (\xi - \theta))^2\big] - \theta^2 \right\} \\ \text{subject to } X \text{ given by (2.12).} \end{cases} \tag{2.14}$$

To solve problem (2.14), we first solve the quadratic-loss minimization problem:

$$\begin{cases} \min\limits_{\pi \in \mathcal{A}} \ J_2(x_0, \lambda_0; \pi, c) := \min\limits_{\pi \in \mathcal{A}} \ \mathbb{E}\big[(X(T) - c)^2\big] \\ \text{subject to } X \text{ given by (2.12),} \end{cases} \tag{2.15}$$

where $c \in \mathbb{R}$ is a free parameter in problem (2.15). Let us denote by $\mathcal{J}_2(x_0, \lambda_0; c)$ the value function of problem (2.15), i.e., $\mathcal{J}_2(x_0, \lambda_0; c) := J_2(x_0, \lambda_0; \pi_c^*, c)$, where π_c^* is an optimizer to problem (2.15) associated with the free parameter c. Once problem (2.15) is solved, we are able to solve problem (2.14) and the primal problem (2.13) by considering

$$\max_{\theta \in \mathbb{R}} \ J_3(x_0, \lambda_0; \theta) := \max_{\theta \in \mathbb{R}} \left\{ \mathcal{J}_2(x_0, \lambda_0; \xi - \theta) - \theta^2 \right\}, \tag{2.16}$$

where ξ is the target expected terminal wealth in problem (2.13). To be precise, given the solution θ^* of (2.16), $\pi^* := \pi_c^*|_{c=\xi-\theta^*}$ is the efficient strategy to problem (2.13), and $(\mathbb{E}[(X^*(T) - \xi)^2], \xi)$ is the corresponding efficient frontier, where $X^*(T)$ is the terminal wealth associated with π^* and $\xi \geq x_0 e^{rT}$.

3 Heuristic derivations

As explained above, the key to solving the (constrained) mean-variance problem (2.13) is to analyze the (unconstrained) quadratic-loss minimization problem (2.15). This section is devoted to the investigations of problem (2.15). We obtain a candidate solution of π_c^* to problem (2.15) in (3.22).

Notice that the objective functional J_2 of problem (2.15) is the expected value of a quadratic function and no longer has the "annoying" equality constraint, which is different from the objective J of the original MV problem (2.13). Namely, problem (2.15) is a standard stochastic control problem. As such, several approaches are available to solve such a problem, including the well-known Hamilton-Jacobi-Bellman (HJB) method (via the application of the dynamic programming principle). Here we apply the stochastic maximum principle (see, e.g., Yong and Zhou (1999)[Chapter 3] for standard theory) to tackle problem (2.15). One may refer to Appendix C for the connection between the stochastic maximum principle approach and the HJB equation approach. One of the advantages of using the stochastic maximum principle is that it is naturally related to BDSE that is pivotal to discuss the solvability of (3.24) and design the algorithm for the numerical analysis in Section 5.

Since the standard maximum principle only serves as a necessary condition for optimality, the analysis in this section is heuristic and π_c^* in (3.22) is a *candidate* for the optimal solution to problem (2.15). We do not attempt to verify that π_c^* given by (3.22) solves problem (2.15), since our ultimate goal is to solve the MV problem (2.13). (Under additional assumptions, the stochastic maximum principle becomes sufficient and indeed leads to an optimal strategy; see for instance Øksendal and Sulem (2019)[Theorem 5.4] and Yong and Zhou (1999)[Chapter 3, Section 5].) Instead, the main results of this section will pave the way to Theorem 4.1 in the next section, which eventually provides a complete solution to the MV problem (2.13). The candidate solution π_c^* in (3.22) to problem (2.15) has a similar representation form as the efficient strategy π^* in (4.1) to problem (2.13). The relation between π_c^* and π^* will be established in Section 4 by using the Lagrangian technique as in Lim and Zhou (2002) and Zhou and Yin (2003).

In what follows, we carry out heuristic derivations in *four* steps to solve the quadratic-loss minimization problem (2.15). The main arguments are similar to those in Shen and Zeng (2014)[Lemma 3.1], where the authors apply the stochastic maximum principle to solve a mean-variance investment and reinsurance problem with delay. Please refer to Øksendal and Sulem (2019)[Section 5.2.5] for the application of the stochastic maximum principle to a standard MV portfolio selection problem under the Black-Scholes model *without* contagion risk. Note that the results in Øksendal and Sulem (2019)[Section 5.2.5] are also heuristic or at least informal, as they do not verify the optimality and admissibility of the obtained candidate strategy. Finally, we remark that, since the Hawkes intensity process λ is unbounded, the existing results of the stochastic maximum principle for jump-diffusion control problems cannot be directly applied to verify the optimality, which will be deferred to the next section by using the completing the square technique in LQ

control theory.

Step 1. Write down the associated Hamiltonian system.

We define a new process $\widehat{X} := \{\widehat{X}(t)\}_{t\in[0,T]}$ by

$$\widehat{X}(t) := X(t) - ce^{-r(T-t)}, \tag{3.1}$$

where $c \in \mathbb{R}$ is the same free parameter in problem (2.15). Notice that $\widehat{X}$ has the same flavor as the auxiliary process y in Zhou and Li (2000). It is clear that $\widehat{X}$ is defined for any $\pi \in \mathcal{A}$, and here we suppress such dependency for notational brevity. By applying Itô's formula to $\widehat{X}$, we obtain

$$\mathrm{d}\widehat{X}(t) = \left(r\widehat{X}(t) + \pi(t)^\top B\right)\mathrm{d}t + \pi(t)^\top \sigma \, \mathrm{d}W(t) + \int_{(-1,\infty)^m} \pi(t)^\top \eta(z)\widetilde{N}(\mathrm{d}t, \mathrm{d}z), \tag{3.2}$$

with the initial value $\widehat{X}(0) = x_0 - ce^{-rT}$. The objective functional J_2 of problem (2.15) can be rewritten as

$$J_2(x_0, \lambda_0; \pi, c) = \mathbb{E}\left[(X(T) - c)^2\right] = \mathbb{E}\left[\widehat{X}^2(T)\right].$$

We treat process $\widehat{X}$ as the state process, and its dynamics equation (3.2) suggests that the Hamiltonian of problem (2.15) reads as

$$\mathcal{H}(t, x, \pi, p, q, u) := \left(rx + \pi^\top B\right)p + \pi^\top \sigma q + \int_{(-1,\infty)^m} \pi^\top \eta(z) \, \mathrm{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)] \, u(t, z), \tag{3.3}$$

where operator $\bullet$ is defined by (2.8). The triplet (p, q, u) is the so-called *adjoint process*. Notice that $p := \{p(t)\}_{t\in[0,T]} \in \mathbb{R}$, $q := \{q(t)\}_{t\in[0,T]} \in \mathbb{R}^n$, and $u := \{u(t, z)\}_{(t,z)\in[0,T]\times(-1,\infty)^m} \in \mathbb{R}^m$.

Step 2. Find the adjoint process (p, q, u).

According to the standard control theory (see, e.g., Yong and Zhou (1999) and Øksendal and Sulem (2019)), the adjoint process (p, q, u) satisfies the following adjoint equation:

$$\begin{cases} \mathrm{d}p(t) = -rp(t)\,\mathrm{d}t + q(t)^\top \, \mathrm{d}W(t) + \int_{(-1,\infty)^m} u(t, z)^\top \widetilde{N}(\mathrm{d}t, \mathrm{d}z), \\ p(T) = 2\widehat{X}(T) = 2(X(T) - c), \end{cases} \tag{3.4}$$

where r is the risk-free interest rate, W is the n-dimensional Brownian motion, and $\widetilde{N}$ is the compensated random measure defined in (2.10). In fact, the drift term in (3.4) is found by

$$-\mathcal{H}_x(t, \widehat{X}(t), \pi(t), p(t), q(t), u(t)) = -rp(t).$$

To solve the adjoint equation (3.4), we try an *ansatz* for p in the form of

$$p(t) = Y(t)\widehat{X}(t), \tag{3.5}$$

13

where Y is the first component of the solution (Y, V) to the following backward stochastic differential equation (BSDE):

$$\begin{cases} \mathrm{d}Y(t) = -f(t, Y(t), V(t))\,\mathrm{d}t + V(t)^\top\,\mathrm{d}\widetilde{N}(t), \\ Y(T) = 2. \end{cases} \tag{3.6}$$

In (3.6), the process V is not a priori known. Instead, Y and V need to be solved at the same time. The driver f of the BSDE (3.6) is unknown at the moment, but will be determined later. Notice that the BSDE (3.6) is driven by the jump process $\widetilde{N}$ only.

Using (3.2) and (3.6), and the stochastic product rule on $p = Y\widehat{X}$, we get

$$\begin{aligned} \mathrm{d}p(t) &= \widehat{X}(t-)\,\mathrm{d}Y(t) + Y(t-)\,\mathrm{d}\widehat{X}(t) + \mathrm{d}[Y(t), \widehat{X}(t)] \\ &= \Bigg[-f(t, Y(t), V(t))\widehat{X}(t) + Y(t)(r\widehat{X}(t) + \pi(t)^\top B) \\ &\quad + \sum_{i=1}^k \sum_{l=1}^m \int_{(-1,\infty)} \pi_i(t)\eta_{il}(z_l)V_l(t)\lambda_l(t)\nu_l(\mathrm{d}z_l) \Bigg]\mathrm{d}t + Y(t)\pi(t)^\top\sigma\,\mathrm{d}W(t) \\ &\quad + \sum_{l=1}^m \int_{(-1,\infty)} \Bigg[\widehat{X}(t-)V_l(t) + \sum_{i=1}^k \pi_i(t)\eta_{il}(z_l)\left(Y(t-) + V_l(t)\right) \Bigg] \widetilde{N}_l(\mathrm{d}t, \mathrm{d}z_l). \end{aligned} \tag{3.7}$$

By comparing the dynamics of p in (3.4) and (3.7), we obtain

$$\begin{aligned} -rY(t)\widehat{X}(t) &= -f(t, Y(t), V(t))\widehat{X}(t) + Y(t)(r\widehat{X}(t) + \pi(t)^\top B) \\ &\quad + \sum_{i=1}^k \sum_{l=1}^m \int_{(-1,\infty)} \pi_i(t)\eta_{il}(z_l)V_l(t)\lambda_l(t)\nu_l(\mathrm{d}z_l), \end{aligned} \tag{3.8}$$

$$q(t) = \sigma^\top\pi(t)Y(t), \tag{3.9}$$

$$u_l(t, z_l) = \widehat{X}(t-)V_l(t) + \sum_{i=1}^k \pi_i(t)\eta_{il}(z_l)\left(Y(t-) + V_l(t)\right), \qquad \forall\, l \in \mathbb{M}. \tag{3.10}$$

In the above system of equations (3.8), (3.9) and (3.10), the unknowns are Y, $V = (V_1, \ldots, V_m)^\top$, f, q, and $u = (u_1, \cdots, u_m)^\top$. The three equations are not enough to fully solve all the five unknowns. At this stage, the adjoint equation (3.4) is partially solved subject to Y and V satisfying the BSDE (3.6).

Step 3. Find a candidate for the optimal solution to problem (2.14).

Under the maximum principle, the candidate for the optimal solution is given by

$$\pi_c^*(t) = \underset{\pi \in \mathbb{R}^k}{\arg\max}\ \mathcal{H}\left(t, \widehat{X}^*(t), \pi, p^*(t), q^*(t), u^*(t, \cdot)\right),$$

where $\mathcal{H}$ is defined in (3.3) and $\widehat{X}^*$ and (p^*, q^*, u^*) are the wealth process and the adjoint process associated with strategy π_c^*. The subscript in π_c^* indicates that problem (2.15) is parameterized by $c \in \mathbb{R}$, and note

14

that the same parameter c also enters the definition of $\widehat{X}$ in (3.1). As a result, we obtain π_c^* via the first-order condition

$$p(t)B + \sigma q(t) + \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]\, u(t,z) = 0, \tag{3.11}$$

where B is the risk premium vector of dimension $k \times 1$, σ is the volatility matrix of dimension $k \times n$, and η is the jump size matrix of dimension $k \times m$ (see their definitions in (2.6) and (2.7)).

Substituting (3.5), (3.9) and (3.10) into the above equation (3.11), we get

$$\pi_c^*(t) = -\Theta(t, Y(t-), V(t))^{-1}\mathcal{Z}(t, Y(t-), V(t))\,\widehat{X}^*(t-), \tag{3.12}$$

where the $k \times k$ matrix Θ and the k-dimensional vector $\mathcal{Z}$ are defined by

$$\Theta(t, Y(t-), V(t)) := Y(t-)\left(\sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]\eta(z)^\top\right)$$

$$+ \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[V(t) \bullet \lambda(t) \bullet \nu(\mathrm{d}z)]\,\eta(z)^\top, \tag{3.13}$$

$$\text{and} \quad \mathcal{Z}(t, Y(t-), V(t)) := BY(t-) + \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]\,V(t). \tag{3.14}$$

We end this step by noting that π_c^* given by (3.12) is in closed form, once Y and V are identified, which will be achieved in the next step.

Step 4. Solve the BSDE (3.6) to obtain Y and V.

We first plug π_c^*, given by (3.12), back to (3.8), and identify the driver f of the BSDE (3.6) as follows:

$$f(t, Y(t), V(t)) = \left(2r - B^\top\Theta^{-1}\mathcal{Z}\right)Y(t) - \int_{(-1,\infty)^m} V(t)^\top \operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]\eta(z)^\top\Theta^{-1}\mathcal{Z}$$

$$= 2rY(t) - \mathcal{Z}(t, Y(t), V(t))^\top\Theta(t, Y(t), V(t))^{-1}\mathcal{Z}(t, Y(t), V(t)), \tag{3.15}$$

where we have suppressed the arguments of both Θ and $\mathcal{Z}$ in the first equality above.

Now with the driver f given by (3.15), the BSDE (3.6) is fully specified. We try the following ansatz for the solution (Y, V) to (3.6):

$$Y(t) = 2e^{2r(T-t)+g(t,\lambda(t))}, \tag{3.16}$$

$$V_l(t) = 2e^{2r(T-t)+g(t,\lambda(t))}\left(e^{g(t,\lambda(t)+\beta_{(l)})-g(t,\lambda(t))} - 1\right), \qquad l \in \mathbb{M}, \tag{3.17}$$

where $g(t,\lambda)$ is yet to be determined and $\beta_{(l)}$ is the lth column of the matrix $\beta = [\beta_{jl}]_{j,l\in\mathbb{M}}$, i.e.,

$$\beta_{(l)} := (\beta_{1l}, \beta_{2l}, \ldots, \beta_{ml})^\top.$$

Recall that β captures the impact of mutual-excitation on the intensity process λ; see (2.3) and (2.4).

Applying Itô's formula to Y in (3.16), we have

$$dY(t) = Y(t)\left[-2r + g_t + g_\lambda^\top\left(\alpha\lambda_\infty + (\beta - \alpha)\lambda(t)\right) + \sum_{l=1}^m\left(e^{g(t,\lambda(t)+\beta_{(l)})-g(t,\lambda(t))} - 1\right)\right.$$

$$\left. - \sum_{j=1}^m \beta_{jl}g_{\lambda_l}\right)\lambda_l(t)\right]dt + Y(t-)\sum_{l=1}^m\left(e^{g(t,\lambda(t)+\beta_{(l)})-g(t,\lambda(t))} - 1\right)d\widetilde{N}_l(t),\tag{3.18}$$

where $g_.$ denotes the partial derivative of g with respect to the corresponding argument (e.g., $g_t = \frac{\partial g}{\partial t}(t,\lambda)$), and the arguments of g are suppressed in its partial derivatives.

To simplify notations, let us introduce

$$U(t,\lambda) := \left(e^{g(t,\lambda+\beta_{(1)})-g(t,\lambda)} - 1, \ldots, e^{g(t,\lambda+\beta_{(m)})-g(t,\lambda)} - 1\right)^\top,\tag{3.19}$$

$$\Gamma(t,\lambda) := \frac{1}{Y(t-)}\,\Theta(t,Y(t-),V(t))$$

$$= \sigma\sigma^\top + \int_{(-1,\infty)^m}\eta(z)\,\mathrm{Diag}[(U(t,\lambda) + \mathbf{1}_m)\bullet\lambda\bullet\nu(dz)]\,\eta(z)^\top,\tag{3.20}$$

$$\widehat{\mathcal{Z}}(t,\lambda) := \frac{1}{Y(t-)}\,\mathcal{Z}(t,Y(t-),V(t))$$

$$= B + \int_{(-1,\infty)^m}\eta(z)\,\mathrm{Diag}[\lambda\bullet\nu(dz)]\,U(t,\lambda),\tag{3.21}$$

where Θ and $\mathcal{Z}$ are defined respectively by (3.13) and (3.14), and $\mathbf{1}_m$ is the m-dimensional vector with all elements equal to 1. It is important to notice that, with the ansatz of Y and V in (3.16) and (3.17), Γ and $\widehat{\mathcal{Z}}$ defined above are indeed functions of t and λ only, and are independent of Y and V. By the above definitions, we easily see

$$V(t) = Y(t)\,U(t,\lambda(t)),\quad \Theta(t,Y(t-),V(t)) = Y(t-)\,\Gamma(t,\lambda(t)),\quad \mathcal{Z}(t,Y(t-),V(t)) = Y(t-)\widehat{\mathcal{Z}}(t,\lambda(t)).$$

In addition, we reorganize π_c^* in (3.12) and obtain

$$\pi_c^*(t) = -\Gamma(t,\lambda(t))^{-1}\,\widehat{\mathcal{Z}}(t,\lambda(t))\,\widehat{X}^*(t-).\tag{3.22}$$

Using U and Γ in (3.19) and (3.20), we rewrite the BSDE (3.6) in the differential form

$$dY(t) = -Y(t)\left[2r - \widehat{\mathcal{Z}}(t,\lambda(t))^\top\Gamma(t,\lambda(t))^{-1}\widehat{\mathcal{Z}}(t,\lambda(t))\right]dt + Y(t-)U(t,\lambda(t))^\top\,d\widetilde{N}(t).\tag{3.23}$$

By matching the drift terms in (3.18) and (3.23), we obtain the following non-local partial differential equation of $g : [0,T]\times\mathbb{R}_+^m \to \mathbb{R}$

$$\begin{cases} g_t(t,\lambda) + g_\lambda(t,\lambda)^\top\alpha(\lambda_\infty - \lambda) + U(t,\lambda)^\top\lambda = \widehat{\mathcal{Z}}(t,\lambda)^\top\Gamma(t,\lambda)^{-1}\widehat{\mathcal{Z}}(t,\lambda), & t < T, \\ g(t,\cdot) = 0, & t = T. \end{cases}\tag{3.24}$$

Recall matrix α and vector λ_∞ are defined by (2.4), while U, Γ and $\widehat{\mathcal{Z}}$ are defined respectively by (3.19), (3.20) and (3.21). In (3.24), U has a non-local dependence on g (see (3.19)), which originates from the mutual-excitation feature of the Hawkes process. This completes the 4-step procedure.

We next present a technical result regrading the bounds of $e^{g(t,\lambda)}$, where g solves (3.24).

Lemma 3.1. *Let* $g(\cdot,\cdot)$ *be the solution to* (3.24). *Then there exists* $\epsilon > 0$ *such that*

$$\epsilon \leq e^{g(t,\lambda(t))} \leq 1, \qquad \forall t \in [0,T], \qquad and \qquad \epsilon \leq e^{g(0,\lambda_0)} < 1.$$

where $\lambda(\cdot)$ *is the intensity process satisfying the Hawkes model* (2.3).

Proof. See Appendix A. $\qquad\qquad\square$

We provide the solvability results of (3.24) in the following lemma.

Lemma 3.2. *The non-local partial differential equation* (3.24) *admits a unique solution.*

Proof. See Appendix B. $\qquad\qquad\square$

We end this section with several important remarks on the above heuristic analysis of problem (2.15) and explain how the analysis here will help solve the MV problem (2.13) in Section 4.

- Notice that problem (2.15) aims to minimize the expectation of a convex function of $\widehat{X}(T)$ and the admissible set $\mathcal{A}$ is also convex, as easily seen from Definition 2.2. In consequence, we can also apply the HJB equation approach to solve problem (2.15); see Appendix C. Indeed, the solution obtained by the HJB equation approach matches that obtained by the stochastic maximum principle approach, which is due to the inherent relationship between the two approaches. One can refer to Øksendal and Sulem (2019) for such relationship under the Poisson jump case. If the jump process is not a Hawkes process but a standard Poisson process, one can follow the standard procedure to verify that π_c^* given by (3.22) is indeed the optimal solution to problem (2.15). We refer interested readers to Yong and Zhou (1999)[Chapter 3, Section 5] for the sufficient conditions of optimality on general stochastic control problems without jumps. However, jumps in our model are derived from a multivariate Hawkes process and, due to the unboundedness of its intensity process λ, existing results cannot be readily applied to confirm the optimality of π_c^*.

- Given that the wealth SDE (2.12) is linear and the objective J_2 is quadratic, problem (2.15) is a linear-quadratic (LQ) control problem. By LQ theory, we naturally conjecture that $Y\widehat{X}^2$ is a submartingale for any $\pi \in \mathcal{A}$ and a martingale under π_c^*. Assume for now that such a conjecture is correct, which will be verified in the next section, then we immediately have

$$\frac{1}{2}\mathbb{E}\Big[Y(T)\big(\widehat{X}^*(T)\big)^2\Big] = \frac{1}{2}Y(0)\big(\widehat{X}^*(0)\big)^2 \leq \frac{1}{2}\mathbb{E}\Big[Y(T)\big(\widehat{X}(T)\big)^2\Big],$$

where $\widehat{X}^*$ denotes the wealth process under the "optimal" strategy π_c^*. As a result, we observe after noticing $Y(T) = 2$ that

$$\mathcal{J}_2(x_0, \lambda_0; c) = \min_{\pi \in \mathcal{A}} \frac{1}{2} \mathbb{E}\Big[Y(T) \big(\widehat{X}(T) \big)^2 \Big] = e^{2rT + g(0, \lambda_0)} \big(x_0 - c\, e^{-rT} \big)^2, \tag{3.25}$$

where we have used (3.1) and (3.16).

- Once the value function to problem (2.15) is found via (3.25), we next solve problem (2.16), a simple quadratic problem of θ (i.e., the Lagrangian multiplier), and obtain its solution θ^*. Now applying the arguments at the end of Section 2 (see also Lim and Zhou (2002) and Zhou and Yin (2003)), we conclude that π_c^* with c replaced by $\xi - \theta^*$ is the efficient strategy to the MV problem (2.13), upon a successful verification of $\pi_{\xi - \theta^*}^* \in \mathcal{A}$.

4 Efficient strategy and efficient frontier

In this section, we provide a complete solution to the mean-variance portfolio selection problem (2.13). Theorem 4.1 obtains the efficient strategy and the efficient frontier to problem (2.13), both in semi-closed form. To prove this theorem, we follow the key steps outlined in the above remarks in the previous section. In particular, to show $Y\widehat{X}^2$ is a martingale under the optimal strategy π_c^*, we apply the idea of "completing the square" in LQ control theory (see Yong and Zhou (1999) and Zhou and Li (2000)). A technical difficulty in the proof of Theorem 4.1 is to verify that the candidate efficient strategy is admissible.

Theorem 4.1. *Let $g(\cdot, \cdot)$ be the solution to (3.24), and $\Gamma(\cdot, \cdot)$ and $\widehat{\mathcal{Z}}(\cdot, \cdot)$ be defined by (3.20) and (3.21), respectively. The efficient strategy $\pi^* = \{\pi^*(t)\}_{t \in [0, T]}$ of the mean-variance portfolio selection problem (2.13) is given by*

$$\pi^*(t) = -\Gamma(t, \lambda(t))^{-1}\, \widehat{\mathcal{Z}}(t, \lambda(t)) \left(X^*(t-) - (\xi - \theta^*) e^{-r(T-t)} \right), \tag{4.1}$$

and the efficient frontier is given by

$$\mathrm{Var}[X^*(T)] = \frac{e^{g(0, \lambda_0)}}{1 - e^{g(0, \lambda_0)}} \big(x_0 e^{rT} - \xi \big)^2, \qquad \xi \geq x_0 e^{rT}, \tag{4.2}$$

where the corresponding Lagrange multiplier θ^* is obtained by

$$\theta^* = \frac{e^{g(0, \lambda_0)}}{1 - e^{g(0, \lambda_0)}} \big(x_0 e^{rT} - \xi \big). \tag{4.3}$$

Here, X^* is the solution to the SDE (2.12) under strategy π^*, and is called the optimal wealth process.

Proof. We divide the proof into three steps.

Step 1: Apply the "completing the square" to obtain π_c^*.

18

Recall that the dynamics of $\widehat{X}$ is given by (3.2) and that of Y in (3.23) from Section 3. By applying Itô's formula to $\widehat{X}^2$ and $Y\widehat{X}^2$, we obtain

$$d\widehat{X}^2(t) = \left(2r\widehat{X}^2(t) + 2\widehat{X}(t)\pi(t)^\top B + \pi(t)^\top \sigma\sigma^\top \pi(t) + \int_{(-1,\infty)^m} \pi(t)^\top \eta(z)\,\mathrm{Diag}[\lambda(t)\bullet\nu(dz)]\,\eta(z)^\top \pi(t)\right)dt$$

$$+ 2\widehat{X}(t)\pi(t)^\top \sigma\,dW(t) + \int_{(-1,\infty)^m} 2\widehat{X}(t-)\pi(t)^\top \eta(z)\widetilde{N}(dt,dz)$$

$$+ \int_{(-1,\infty)^m} \pi(t)^\top \eta(t,z)\,\mathrm{Diag}[\widetilde{N}(dt,dz)]\,\eta(z)^\top \pi(t),$$

and

$$dY(t)\widehat{X}^2(t) = Y(t)\left[\pi(t) + \Gamma(t,\lambda(t))^{-1}\widehat{\mathcal{Z}}(t,\lambda(t))\widehat{X}(t)\right]^\top \Gamma(t,\lambda(t))\underbrace{\left[\pi(t) + \Gamma(t,\lambda(t))^{-1}\widehat{\mathcal{Z}}(t,\lambda(t))\widehat{X}(t)\right]}_{:=\mathcal{U}(t,\lambda(t))}dt$$

$$+ 2Y(t)\widehat{X}(t)\pi(t)^\top \sigma\,dW(t) + \int_{(-1,\infty)^m} Y(t-)\pi(t)^\top \eta(z)\,\mathrm{Diag}[U(t,\lambda(t))\bullet\widetilde{N}(dt,dz)]\,\eta(z)^\top \pi(t)$$

$$+ \int_{(-1,\infty)^m} Y(t-)\widehat{X}(t-)\big(2\pi(t)^\top \eta(z) + \widehat{X}(t-)U(t,\lambda(t))\big)\widetilde{N}(dt,dz). \tag{4.4}$$

Define a sequence of stopping times $\{\tau_i\}_{i=1,2,\dots}$ by

$$\tau_i := \inf\left\{t \geq 0 \,\Big|\, |Y(t)\widehat{X}(t)| > i\right\}, \qquad i = 1, 2, \dots.$$

It is easily seen that $\tau_i \uparrow \infty$ and $(T\wedge\tau_i) \uparrow T$, as $i \to \infty$. We apply the localization technique, i.e., integrating from 0 to $T\wedge\tau_i$ and taking expectations on both sides of (4.4), and eventually get

$$\frac{1}{2}\mathbb{E}\big[Y(T\wedge\tau_i)(X(T\wedge\tau_i) - c)^2\big] - \frac{1}{2}Y(0)\left(x_0 - ce^{-rT}\right)^2 = \frac{1}{2}\mathbb{E}\left[\int_0^{T\wedge\tau_i} Y(t)\mathcal{U}(t,\lambda(t))^\top \Gamma(t,\lambda(t))\mathcal{U}(t,\lambda(t))dt\right],$$
$$\tag{4.5}$$

where we have used the integrability conditions in Definition 2.2. Note that $X(\cdot) \in \mathcal{S}_{\mathcal{F}}^2(0,T;\mathbb{R})$, $Y(\cdot)$ is bounded, and the integrand on the right-hand of the above equality is positive. Then by the dominated and monotone convergence theorems to the left-hand side and the right-hand side of (4.5), we obtain, when sending i to ∞, that

$$\mathbb{E}\big[(X(T) - c)^2\big] - \frac{1}{2}Y(0)\left(x_0 - ce^{-rT}\right)^2 = \frac{1}{2}\mathbb{E}\left[\int_0^T Y(t)\mathcal{U}(t,\lambda(t))^\top \Gamma(t,\lambda(t))\mathcal{U}(t,\lambda(t))dt\right],$$

where we have used the terminal value $Y(T) = 2$.

19

Therefore, by setting $\mathcal{U}(t, \lambda(t)) = 0$, we find the (candidate) optimal control by

$$\pi_c^*(t) = -\Gamma(t, \lambda(t))^{-1} \underbrace{\left(B + \int_{(-1,\infty)^m} \eta(z) \operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)] U(t, \lambda(t)) \right)}_{=\widehat{\mathcal{Z}}(t, \lambda(t))} \underbrace{\left(X^*(t-) - c e^{-r(T-t)} \right)}_{=\widehat{X}^*(t-)}, \quad (4.6)$$

which coincides with the one in (3.22).

We end this step with an important observation. In Section 3, we obtain π_c^* in (3.22) by the stochastic maximum principle as a necessary condition for optimality. Here by the so-called "completing the square" technique, we easily see that π_c^* given by (4.6) is indeed optimal to the quadratic-loss minimization problem (2.15) (a sufficient condition for optimality). Note that sufficiency is guaranteed by the facts that $Y > 0$ and Γ is positive definite; see (3.16) and (3.20).

Step 2: Solve problem (2.16) to get θ^* and obtain the efficient strategy and the efficient frontier.

Recall that problem (2.16) is to maximize $\mathcal{J}_2(x_0, \lambda_0; \xi - \theta) - \theta^2$ over all $\theta \in \mathbb{R}$, where $\mathcal{J}_2$ is the value function to problem (2.15) and is obtained by (see (3.25))

$$\mathcal{J}_2(x_0, \lambda_0; c) = \mathbb{E}\left[(\widehat{X}^*(T))^2\right] = \mathbb{E}\left[(X^*(T) - c)^2\right] = \frac{1}{2} Y(0) \left(x_0 - c e^{-rT}\right)^2. \quad (4.7)$$

Now noting $c = \xi - \theta$, we obtain θ^* by

$$\begin{aligned}
\theta^* &= \arg\max_{\theta \in \mathbb{R}} \left\{ \frac{1}{2} Y(0) \left(x_0 - (\xi - \theta) e^{-rT}\right)^2 - \theta^2 \right\} \\
&= \frac{e^{-2rT} Y(0)}{2 - e^{-2rT} Y(0)} \left(x_0 e^{rT} - \xi\right),
\end{aligned}$$

which reduces to (4.3) by using (3.16). The result of Lemma 3.1 guarantees that the first-order condition to problem (2.16) is also sufficient and the above θ^* is indeed the optimal solution to problem (2.16).

With θ^* obtained as in (4.3), we immediately get the efficient strategy π^* by using $\pi^* = \pi_{\xi-\theta^*}^*$ and the general expression of π_c^* in (4.6). Using the definition of $\widehat{X}$ in (3.1) again then leads to the result of π^* in (4.1). Finally, by the Lagrangian duality theorem, we obtain the efficient frontier in (4.2) via

$$\operatorname{Var}[X^*(T)] = \mathbb{E}\left[\left(X^*(T) - (\xi - \theta^*)\right)^2\right] - (\theta^*)^2.$$

Step 3: Verify that π^* given by (4.1) is admissible (i.e., $\pi^* \in \mathcal{A}$).

The previous two steps have justified all the results in Theorem 4.1, only contingent on $\pi^* \in \mathcal{A}$. Hence, in the final step, we verify that the efficient strategy π^* in (4.1) is admissible.

To start, by using (4.7) and $e^{g(0, \lambda_0)} < 1$ from Lemma 3.1, we have

$$\mathbb{E}\left[(X^*(T) - (\xi - \theta^*))^2\right] = \operatorname{Var}[X^*(T)] + (\theta^*)^2 = \frac{e^{g(0, \lambda_0)}}{(1 - e^{g(0, \lambda_0)})^2} \left(x_0 e^{rT} - \xi\right)^2 < \infty.$$

Thus, $\widehat{X}^*(T) = X^*(T) - (\xi - \theta^*)$ is square integrable. This fact implies that the adjoint equation (3.4), associated with the efficient strategy π^*, admits a unique solution such that $(p^*, q^*, u^*) \in \mathcal{S}_{\mathcal{F}}^2(0, T; \mathbb{R}) \times \mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^n) \times \mathcal{L}_{\mathcal{F}}^{2,N}(0, T; \mathbb{R}^m)$.

Next we proceed to show that $Y(t)$ is bounded below from 0. That is, there exists a positive constant δ such that $Y(t) \geq \delta$. Since the efficient strategy does not correspond to a risk-free investment, $e^{g(0,\lambda_0)}$ must be positive. Otherwise, it can be seen from (4.2) that the minimum variance is always zero regardless of the expected return, i.e., $\mathrm{Var}[X^*(T)] = 0$, for any $\xi \in (x_0 e^{rT}, +\infty)$, which is unreasonable. Therefore, using (A.3), we obtain

$$Y(t) = 2e^{2r(T-t)+g(t,\lambda(t))} \geq 2e^{2rT} \cdot e^{g(0,\lambda_0)} := \delta > 0. \tag{4.8}$$

It then follows from (3.5) and $p^* \in \mathcal{S}_{\mathcal{F}}^2(0, T; \mathbb{R})$ that

$$\widehat{X}^*(t) = \frac{p^*(t)}{Y(t)} \leq \frac{p^*(t)}{\delta} \quad \text{for all } t \in [0, T] \quad \Rightarrow \quad \widehat{X}^* \in \mathcal{S}_{\mathcal{F}}^2(0, T; \mathbb{R})$$

and from (A.3) that V_l is bounded, for all $l \in \mathbb{M}$. Note

$$\sum_{i=1}^k \pi_i^*(t)\eta_{il}(z_l) Y(t) = \sum_{i=1}^k \pi_i^*(t)\eta_{il}(z_l) \left(Y(t-) + V_l(t) \right) = u_l(t, z_l) - \widehat{X}^*(t-)V_l(t).$$

Thus, we have

$$Y\eta^\top \pi^* = u - \widehat{X}^* V \in \mathcal{L}_{\mathcal{F}}^{2,N}(0, T; \mathbb{R}^m)$$

$$\text{and} \quad \eta^\top \pi^* = \frac{u - \widehat{X}^* V}{Y} \leq \frac{u - \widehat{X}^* V}{\delta} \in \mathcal{L}_{\mathcal{F}}^{2,N}(0, T; \mathbb{R}^m).$$

Combining the above results, the non-degenerate condition in Lemma 2.1 and $q^* \in \mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^n)$ leads to

$$\epsilon \delta^2 \, \mathbb{E}\left[\int_0^T |\pi^*(t)|^2 \, \mathrm{d}t \right] \leq \mathbb{E}\left[\int_0^T |Y(t)|^2 \pi^*(t)^\top \Sigma(t) \pi^*(t) \, \mathrm{d}t \right]$$

$$= \mathbb{E}\left[\int_0^T |q^*(t)|^2 \, \mathrm{d}t + \int_0^T |Y(t)\eta(z)^\top \pi^*(t)|_N^2 \, \mathrm{d}t \right] < \infty.$$

That is, $\pi^* \in \mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^k)$. All the conditions in Definition 2.2 are now confirmed, and the proof is completed. $\qquad \square$

For comparison, let us consider a standard *Poisson-jump-diffusion* market model without contagion risk, in which jumps of asset prices are modeled by an m-dimensional Poisson process N_P with a *deterministic* intensity λ_P. More specifically, the asset prices follow the SDE:

$$\mathrm{d}S(t) = \mathrm{Diag}[S(t-)]\left(\mu \, \mathrm{d}t + \sigma \, \mathrm{d}W(t) + \int_{(-1,\infty)^m} \eta(z) \, \widetilde{N}_P(\mathrm{d}t, \mathrm{d}z) \right), \tag{4.9}$$

21

where $\widetilde{N}_P$ is defined similarly as $\widetilde{N}$ in (2.10), but with the stochastic intensity λ replaced by the deterministic intensity λ_P. For a given investment strategy π, the investor's wealth process evolves according to

$$\mathrm{d}X(t) = \left(rX(t) + \pi(t)^\top B\right)\mathrm{d}t + \pi(t)^\top \sigma\, \mathrm{d}W(t) + \int_{(-1,\infty)^m} \pi(t)^\top \eta(z)\widetilde{N}_P(\mathrm{d}t, \mathrm{d}z).$$

The MV problems with Poisson jumps are considered by Framstad et al. (2004) and Shen and Siu (2013). We directly provide solutions to the MV problem (2.13) in the Poisson-jump-diffusion model (4.9) as a corollary, which recovers the results obtained in these papers.

Corollary 4.1. *In the standard Poisson-jump-diffusion model, described in (4.9), we obtain the efficient strategy to problem (2.13) by*

$$\pi_P^*(t) = -\Gamma_P(\lambda_P(t))^{-1} B \left(X^*(t-) - (\xi - \theta_P^*)e^{-r(T-t)}\right), \tag{4.10}$$

where

$$\Gamma_P(\lambda) := \sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z)\, \mathrm{Diag}[\lambda \bullet \nu(\mathrm{d}z)]\, \eta(z)^\top, \tag{4.11}$$

and

$$\theta_P^* = \frac{e^{-\int_0^T B^\top \Gamma_P(\lambda_P(t))^{-1} B\, \mathrm{d}t}}{1 - e^{-\int_0^T B^\top \Gamma_P(\lambda_P(t))^{-1} B\, \mathrm{d}t}} \left(x_0 e^{rT} - \xi\right).$$

The efficient frontier is given by

$$\mathrm{Var}[X_P^*(T)] = \frac{e^{-\int_0^T B^\top \Gamma_P(\lambda_P(t))^{-1} B\, \mathrm{d}t}}{1 - e^{-\int_0^T B^\top \Gamma_P(\lambda_P(t))^{-1} B\, \mathrm{d}t}} \left(x_0 e^{rT} - \xi\right)^2, \qquad \xi \geq x_0 e^{rT}.$$

Remark 4.1. *In Corollary 4.1, the intensity λ_P of the Poisson process is deterministic. In comparison, the intensity λ of the Hawkes process N is stochastic, given by (2.2). However, if we set $\beta_{lj} = 0$ for all $l, j \in \mathbb{M}$, the intensity λ in (2.2) reduces to the deterministic case, denoted by λ_P, and the non-local partial differential equation (3.24) becomes an ODE given as follows:*

$$\frac{\mathrm{d}g(t)}{\mathrm{d}t} = B^\top \left(\sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z)\, \mathrm{Diag}[\lambda_P(t) \bullet \nu(\mathrm{d}z)]\, \eta(z)^\top\right)^{-1} B = B^\top \Gamma_P(\lambda_P(t))^{-1} B,$$

where Γ_P is defined in (4.11). Solving the above ODE yields the following closed-form solution

$$g(t) = -\int_t^T B^\top \Gamma_P(\lambda_P(s))^{-1} B\, \mathrm{d}s. \tag{4.12}$$

By comparing (4.1) with (4.10), we conclude that the efficient strategy π_P^* in the Poisson-jump-diffusion model can be seen as a special case of the efficient strategy π^* in the Hawkes-jump-diffusion model. In

(4.10), *the term* $\Gamma_P(\lambda_P(t))^{-1} B$ *represents the product of the precision matrix and the risk premium vector,* *which is standard in the literature on portfolio selection problems, while in* (4.1), $\Gamma(t, \lambda(t))^{-1} \widehat{\mathcal{Z}}(t, \lambda(t))$ *is* *the counterpart term, in which both* $\Gamma(t, \lambda(t))^{-1}$ *and* $\widehat{\mathcal{Z}}(t, \lambda(t))$ *are related to the precision matrix* $\Sigma(t)^{-1}$ *and the risk premium vector* B, *but adjusted by the non-local terms in* $U(t, \lambda(t))$. *Please refer to equations* (3.19)-(3.21) *for the definitions of* $U(\cdot, \cdot)$, $\Gamma(\cdot, \cdot)$, *and* $\widehat{\mathcal{Z}}(\cdot, \cdot)$.

5   Numerical analysis

In this section, we conduct numerical analysis to illustrate the theoretical findings of Theorem 4.1. The main objective is to obtain *qualitative* results from the rather abstract expressions of the efficient frontier in Theorem 4.1.

In the numerical analysis, we study a univariate example with $m = k = 1$ (i.e., there is only one risky asset whose price jumps are modeled by a one-dimensional Hawkes process). We follow Aït-Sahalia et al. (2015) and Liu et al. (2021) to assign the model parameters for this example as in Table 1.

r	μ	σ	J	$\mathbb{E}[Z]$	$\mathbb{E}[Z^2]$	α	β	λ_∞	T	x_0
2%	9%	20%	1	-2%	6%	5	0.1	0.48	2	1

Table 1: Model parameters in the numerical example

5.1   The solution to (3.24)

Recall from Theorem 4.1 that both the efficient strategy (4.1) and the efficient frontier (4.2) depend on $g(0, \lambda_0)$, where $g(\cdot, \cdot)$ solves the non-local partial differential equation (3.24) and $\lambda_0 = \lambda(0)$ is the initial intensity value of the Hawkes process N. Equation (3.24) does not have an analytic solution and is difficult to solve numerically as well, due to the presence of the non-local component. In particular, Equation (3.24) dictates that the partial derivatives of g at (t, λ) relate to not only its value at the same point, $g(t, \lambda)$, but also its value at the corresponding post-jump points, $g(t, \lambda + \beta_{(l)})$, where $l = 1, \ldots, m$. (Here in this example $\beta_{(l)} = \beta$ since $m = 1$.) In order to solve g from (3.24), we apply an exponential transform and study $\widetilde{g}(t, \lambda) = e^{g(t,\lambda)}$. Based on the analysis in Appendix B, we obtain $\widetilde{g}(t, \lambda)$ in (B.1), which can be seen as the Feynman-Kac representation of $\widetilde{g}(t, \lambda)$. We then use this result to numerically solve for $\widetilde{g}(t, \lambda)$ via Monte Carlo simulation, which is briefly discussed below.

Given $t \in [0, T]$ and $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, we choose time step Δ_t and space step Δ_λ, and partition $[0, T] \times [\underline{\lambda}, \bar{\lambda}]$ to obtain discrete grid points (t_i, λ_j), where $i = 1, 2, \ldots, 1 + \mathcal{N}_T$ (i.e., $\mathcal{N}_T = T/\Delta_t$) and $j = 1, 2, \ldots, 1 + \mathcal{N}_\lambda^{(i)}$ (e.g., $\mathcal{N}_\lambda^{(1)} = (\bar{\lambda} - \underline{\lambda})/\Delta_\lambda$) with Δ_t and Δ_λ properly chosen so that $\mathcal{N}_T$ and $\mathcal{N}_\lambda^{(i)}$ are integers. Note that the spatial dimension $1 + \mathcal{N}_\lambda^{(i)}$ varies as time moves forward; see Remark 5.1 for details. The pseudo-algorithm

23

for computing $\left(\widetilde{g}(t_1, \lambda_j)\right)_{j=1,2,\ldots,\mathcal{N}_\lambda^{(1)}}$ is given below:

1. Initialize $\widetilde{g}(t_{1+\mathcal{N}_T}, \lambda_j) = 1$ for all $j = 1, 2, \ldots, \mathcal{N}_\lambda^{(1+\mathcal{N}_T)}$, since $\widetilde{g}(T, \lambda) = 1$ for all $\lambda > 0$.

2. With all $\{\widetilde{g}(t_k, \cdot)\}_{k \geq i+1}$ computed, we calculate $\{\widetilde{g}(t_i, \cdot)\}$ at time t_i by the following two steps:

   (i) Simulate M different paths for the intensity process between t_i and T via (2.3) starting with $\lambda(t_i) = \lambda_j$, denoted by $\{\lambda^{(\mathfrak{m})}(t_k) : k = i, i+1, \ldots, 1 + \mathcal{N}_T \text{ with } \lambda(t_i) = \lambda_j\}$, where $\mathfrak{m} = 1, \ldots, M$;

   (ii) For each path, compute the right hand side of (B.1), denoted by $\widetilde{g}^{(\mathfrak{m})}(t_i, \lambda_j)$, and obtain $\widetilde{g}(t_i, \lambda_j) = \frac{1}{M} \sum_{\mathfrak{m}=1}^M \widetilde{g}^{(\mathfrak{m})}(t_i, \lambda_j)$.

With the above algorithm in hand, we set $T = 2$, $\underline{\lambda} = 0.1$ and $\bar{\lambda} = 2$ (the rest are given in Table 1), and plot $\widetilde{g}$ over $[0, 2] \times [0.1, 2]$ in Figure 1. It can be seen from Figure 1 that $\widetilde{g}$ (or g) is an increasing function of both time argument t and spatial argument λ. Such an observation in comparative statics is consistent with the result derived from the Poisson model in Corollary 4.1. To be precise, the corresponding g under the Poisson model is obtained in (4.12), from which we can easily see that g is increasing with respect to t. By writing $\lambda_P(s) = \lambda + \int_t^s \lambda'(u) \, du$ (here we assume the deterministic intensity $\lambda_P(s)$ is differentiable over $[0, T]$) and using the definition of Γ_P in (4.11), we obtain that Γ_P is an increasing function of $\lambda = \lambda_P(t)$ and so is g.
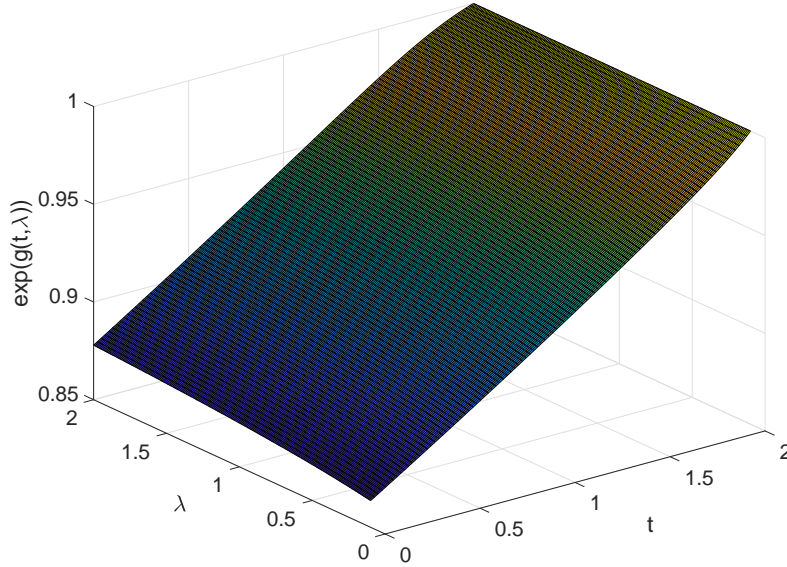

Figure 1: Graph of $\widetilde{g}(t, \lambda) = e^{g(t,\lambda)}$ over $(t, \lambda) \in [0, 2] \times [0.1, 2]$

Remark 5.1. *We have mentioned in the above algorithm that the dimension of spatial points* varies *along the time grid, and here we explain why this is the case for the non-local partial differential equation* (3.24).

24

Given Δ_λ, the spatial dimension at time t_1 ($t_1 = 0$) is easily obtained via $1 + \mathcal{N}_\lambda^{(1)} = 1 + (\bar{\lambda} - \underline{\lambda})/\Delta_\lambda$. Now we investigate what spatial points must be included at time t_2 ($t_2 = \Delta_t$) in order to calculate $\widetilde{g}(t_1, \lambda(t_1))$, where $\lambda(t_1) \in [\underline{\lambda}, \bar{\lambda}]$. To compute $\widetilde{g}(t_1, \lambda(t_1))$ via (B.1), we need $\widetilde{g}(t_2, \lambda(t_2))$ and $\widetilde{g}(t_2, \lambda(t_2) + \beta)$ for all possible $\lambda(t_2)$ at t_2 that can be reached from $\lambda(t_1)$. Given any $\lambda(t_1) \in [\underline{\lambda}, \bar{\lambda}]$ on the spatial grid, by discretization of (2.3), $\lambda(t_2)$ takes two possible values: (1) $\lambda(t_2) = \lambda(t_1) + \alpha(\lambda_\infty - \lambda(t_1))\Delta_t$ (no jump case); and (2) $\lambda(t_2) = \lambda(t_1) + \alpha(\lambda_\infty - \lambda(t_1))\Delta_t + \beta$ (jump case). In addition, we also need the values $\widetilde{g}(t_2, \lambda(t_2) + \beta)$, which are evaluated at $\lambda(t_2) + \beta$, i.e., the post-jump point of $\lambda(t_2)$. In summary, not only the dynamics of the intensity process and but also the non-local equation (3.24) (see U in (3.19)) involve jumps, both of which may lead to the change of the spatial dimension along time.

5.2   Sensitivity analysis

In this subsection, we focus on the sensitivity analysis of the efficient frontier under the Hawkes-jump-diffusion model (hereafter Hawkes model), which is given by (4.2) in Theorem 4.1. We are particularly interested in how the parameters of the Hawkes intensity process affect the efficient frontier. In all the analysis below, we only allow one parameter to vary at different levels and fix the rest as in Table 1.

We first investigate how the initial intensity value λ_0 and the mean-reversion level λ_∞ affect the efficient frontier under the Hawkes model (see (2.4) for their definitions). The results are plotted in Figure 2. In the left panel of Figure 2, we consider three levels for the initial intensity λ_0: (1) low level $\lambda_0 = 0.1$; (2) mean-reversion level $\lambda_0 = \lambda_\infty = 0.48$; and (3) high level $\lambda_0 = 1.9$. We observe from the graphs that the efficient frontier moves downward (i.e., it deteriorates) as λ_0 increases. To see this result, notice when the current intensity level increases, all the subsequent intensity increases as well (although the impact of the current increment decays exponentially), which makes the MV investor worse off. In the right panel of Figure 2, we fix $\lambda_0 = \lambda_\infty$ and consider three levels for the mean-reversion level λ_∞: $\lambda_\infty = 0.3, 0.48, 0.8$. The graphs show that the MV investor benefits when λ_∞ decreases. To understand this finding, note that the higher the λ_∞, the higher the intensity λ on average, as λ mean reverts to λ_∞. A comparison between the two panels indicates that λ_∞ has a more significant impact on the efficient frontier than λ_0.

We next study the impact of α and β on the efficient frontier. Recall from (2.3) that α is the mean-reversion speed and β is the jump size of the intensity process λ. We plot the results in Figure 3. In the left panel of Figure 3, we consider three different levels for α: $\alpha = 1, 2, 5$. When α increases, the intensity process mean reverts to λ_∞ faster, which improves the efficient frontier. The reason for such a result is that with a higher α, the intensity λ spends less time at an abnormal level and is more likely to be around λ_∞, which reduces the uncertainty of the risky asset. In the right panel of Figure 3, we consider three different levels for β: $\beta = 0.1, 0.5, 2$. It is readily seen that when β increases, the efficient frontier worsens. Recall β is the jump size of the intensity λ upon a jump of the Hawkes process N. Therefore, given a
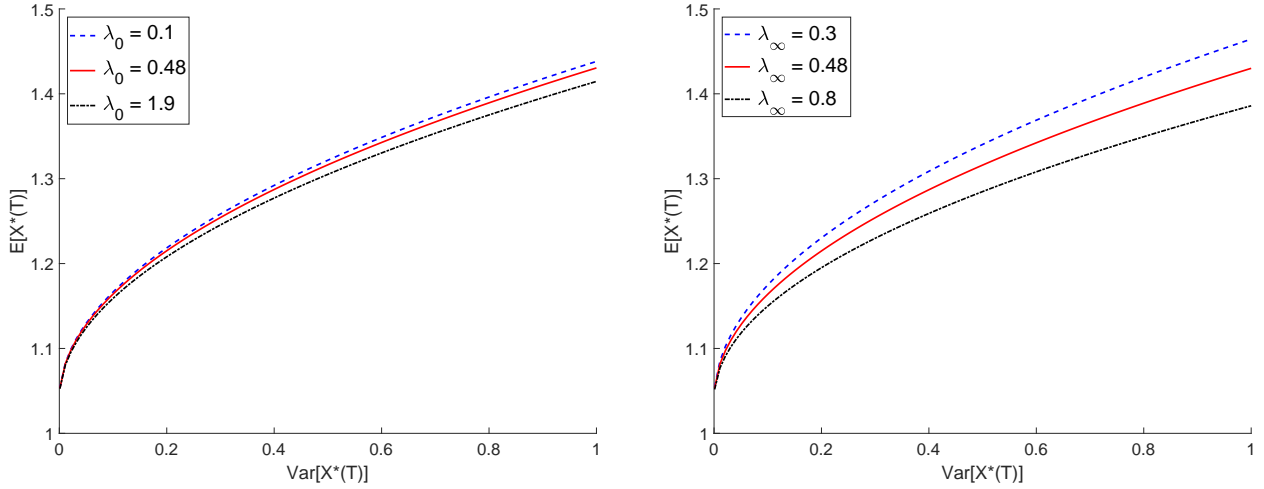
Figure 2: Impact of λ_0 (Left) and λ_∞ (Right) on the efficient frontier under the Hawkes model

bigger β, the increment of λ upon a jump of N is larger, which further "excites" more jumps in the near future and increases the variance in the efficient strategy. For both panels in Figure 3, we set $\lambda_0 = 1$, so that the difference among the three efficient frontiers is more visible. We emphasize that all the findings remain the same for a smaller λ_0.
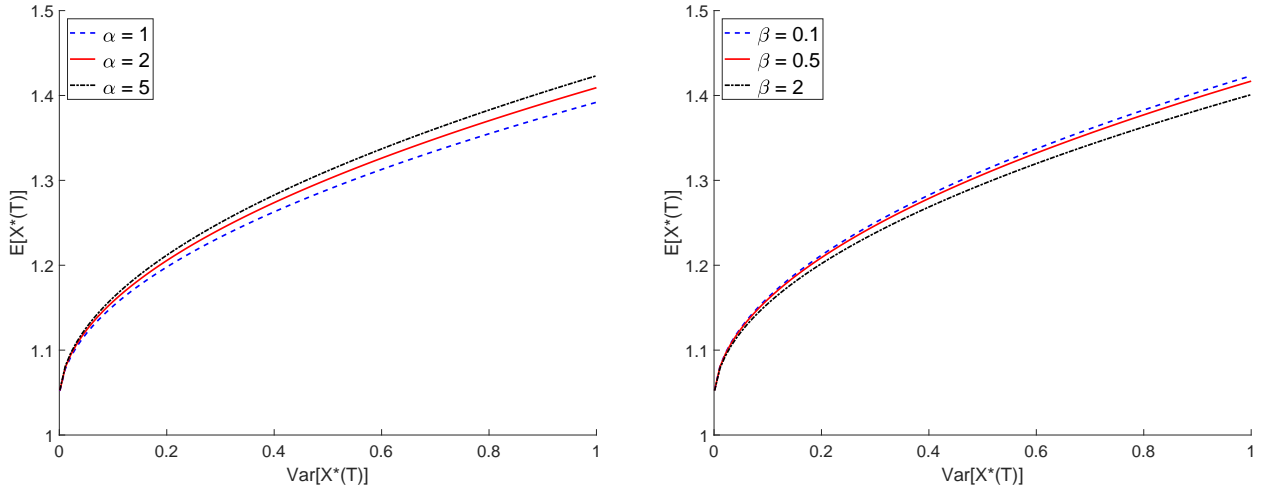

Figure 3: Impact of α (Left) and β (Right) on the efficient frontier under the Hawkes model

5.3   Comparisons with the Poisson model

In the last part of the numerical analysis, we conduct a comparison study between the Hawkes model (2.11) and the Poisson-jump-diffusion model (hereafter, Poisson model) (4.9). In the first study, we consider a relatively small value for the initial intensity ($\lambda_0 = 0.3$) in the Hawkes model, and three different levels

for the *constant* intensity ($\lambda_P = 0.3, 0.4, 0.5$) in the Poisson model. The corresponding efficient frontiers are plotted in the left panel of Figure 4. In the second study, we consider a relatively large value for the initial intensity ($\lambda_0 = 0.7$) in the Hawkes model, and three different levels for the *constant* intensity ($\lambda_P = 0.4, 0.7, 0.8$) in the Poisson model. The corresponding efficient frontiers are plotted in the right panel of Figure 4. We observe from Figure 4 that if $\lambda_0 < \lambda_\infty = 0.48$ (resp. $\lambda_0 > \lambda_\infty = 0.48$), the Hawkes model with the initial intensity λ_0 yields a worse (resp. better) efficient frontier than the Poisson model with the same constant intensity $\lambda_P = \lambda_0$. To understand this finding, let us recall from Table 1 that $\lambda_\infty = 0.48$ (mean-reversion level of the Hawkes intensity λ), $\alpha = 5$ (mean-reversion speed), $\beta = 0.1$ (jump size) and $T = 2$ (investment horizon). In consequence, under the given parameters, the mean-reversion effect dominates the jump effect in the dynamics of the Hawkes intensity process λ given by (2.2). Therefore, if $\lambda_0 < \lambda_\infty$ (resp. $\lambda_0 > \lambda_\infty$), setting $\lambda_P = \lambda_0$ underestimates (resp. overestimates) the overall jump intensity, while the increase of the (stochastic or deterministic) jump intensity deteriorates the efficient frontier (see both Figures 3 and 4), which together explain the comparative finding of Figure 4.
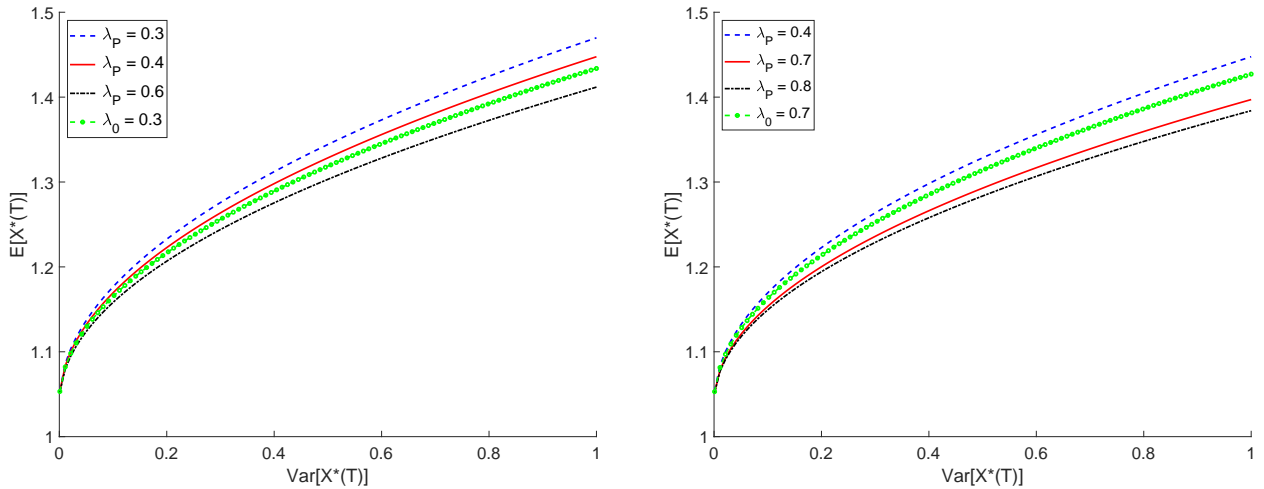

Figure 4: Comparisons of the efficient frontiers under the Hawkes and Poisson models ($\beta = 0.1$)

When a more subtle analysis of Figure 4 is performed, we find that the efficient frontier of the Hawkes model with $\lambda_0 = 0.3$ (resp. 0.7) is very close to that of the Poisson model with $\lambda_P = 0.47$ (resp. 0.51). Again recall from Table 1 that $\lambda_\infty = 0.48$. It then seems necessary to investigate the case of $\lambda_0 = \lambda_\infty = 0.48$ to further compare the Hawkes model with the Poisson model. To see how the self-excitation effect of the Hawkes process affects the efficient frontier, we consider a larger jump size with $\beta = 2$. (Each jump of the Hawkes process $dN_t = 1$ "excites" its own intensity by β.) We plot the results in Figure 5. A clear message from Figure 5 is that the self-excitation feature of Hawkes processes increases the variance of the terminal wealth under the efficient strategy (for any chosen expectation target $\xi \geq x_0 e^{rT}$) and thus leads to a worsened efficient frontier, comparing to the Poisson model with intensity $\lambda_P = \lambda_0 = \lambda_\infty$, where λ_0

(resp. λ_∞) is the initial value (resp. long-term value) of the Hawkes intensity process. (We comment that the finding of Figure 5 remains the same when $\beta = 0.1$ from Table 1 is used, though one needs to zoom in to observe the difference between the two efficient frontiers.)
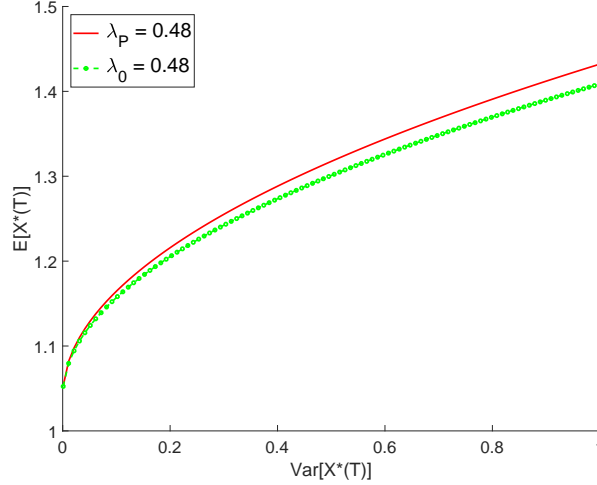

Figure 5: Comparisons of the efficient frontiers under the Hawkes and Poisson models ($\beta = 2$)

Remark 5.2. *In this remark, we offer some discussions on the difficulty of solving* (3.24) *numerically when the spatial dimension is greater than 1 (often known as the "curse of dimensionality"). For this purpose, we first review the computational complexity of the one-dimensional case ($m = k = 1$) in Subsections 5.1-5.3. To solve* (3.24) *numerically in the one-dimensional case, we choose $\Delta_t = 0.02$ and $T = 2$, implying the time dimension is 101, and $\Delta_\lambda = 0.01$ and $\lambda_0 \in [0.1, 2]$, implying the spatial dimension at time 0 is 191. Recall from Remark 5.1 that the spatial dimension is varying as time evolves, and the maximum spatial dimension number is 239 over the total 101 time grid points, corresponding to $\max \lambda = 2.48$, under the given parameters. The computation algorithm moves backward in time with the terminal values $\widetilde{g}(T = t_{101}, \lambda_j) = 1$, and, at each grid point (t_i, λ_j), simulates M (chosen to be 5,000) paths to compute $\widetilde{g}(t_i, \lambda_j)$, yielding a complexity of approximately $100 \times 235 \times 5000 = 1.175 \times 10^8$ (the mean of the spatial dimensions is about 235). We also mention that computing the integrand in (B.1) is also time consuming when time t is getting close to 0.*

Now we are ready to discuss the simplest high dimensional case with $m = 2$ and $k = 2$ (i.e., there are two risky assets whose jumps are derived from a bivariate Hawkes process). Assume the self-excitation effect parameters $\beta_{11} = \beta_{22} = 0.1$ (i.e., a jump of $N_i(t)$ at time t "excites" its own intensity $\lambda_i(t)$ by 0.1, where $i = 1, 2$) and the cross-excitation effect parameters $\beta_{12} = \beta_{21} = 0.08$ (i.e., a jump of $N_i(t)$ at time t "excites" the other's intensity $\lambda_j(t)$ by 0.08). Due to the additional cross-excitation effect, the mean of the spatial dimensions in the bivariate case is about 300 (recall the number is 235 in the univariate case).

As such, the total number of simulations quickly grows to $100 \times 300^2 \times 5000 = 4.5 \times 10^{10}$. In addition, computing the integrand in (B.1) now involves finding the inverse of a matrix within each time step Δ_t. In fact, the computation is so complex that we cannot even obtain $\widetilde{g}(0, \lambda_0)$ for a single set of parameter configuration in three days on a personal laptop (MacBook Pro 2017 with 3.1G Dual Core i5 CPU, 8GB memory, and 512G SSD). This is not entirely surprising though, as solving a multi-dimensional partial differential equation with a non-local component remains an open question to our knowledge, and there are no universally efficient numerical methods to tackle such a problem.

6 Conclusion

In this paper, we consider a mean-variance portfolio selection problem in a contagious financial market, where the prices of risky assets are subject to mutually exciting jumps of a multivariate Hawkes process. The contagion risk is then captured by the fact that a price jump of an asset will increase the jump intensities of both the same asset and all the other assets in the market. Aït-Sahalia and Hurd (2016) study portfolio selection problems under the utility maximization criterion in a similar market model with contagion risk, and apply the standard HJB equation method to obtain optimal investment strategies. Here we obtain semi-explicit solutions to the efficient strategy and the efficient frontier via the stochastic maximum principle, BSDE theory, and LQ control technique. Our paper is among one of the few papers concerning optimal control of dynamical systems with Hawkes-type jumps in the literature of mathematical finance, and we hope that our work could motivate further research in this area. An interesting future research direction is to consider non-Markovian Hawkes processes with non-exponential decay functions. In that case, the HJB equation approach will not be applicable, while the stochastic maximum principle together with BSDEs still works.

Acknowledgments

A   Proof to Lemma 3.1

Proof to Lemma 3.1. We start by rewriting the dynamics of Y in (3.23) by

$$\mathrm{d}Y(t) = Y(t)\left(-2r + \widehat{\mathcal{Z}}(t,\lambda(t))^\top \Gamma(t,\lambda(t))^{-1}\widehat{\mathcal{Z}}(t,\lambda(t))\right)\mathrm{d}t + Y(t-)U(t,\lambda(t))^\top \mathrm{d}\widetilde{N}(t),$$

which implies that $\{e^{2rt}Y(t)\}_{t\in[0,T]}$ is a sub-martingale, and thus

$$e^{2rt}Y(t) \leq \mathbb{E}_t\left[e^{2rT}Y(T)\right], \qquad \forall\, t \in [0,T].$$

Recall $Y(T) = 2$ and (3.16), we have

$$2e^{2r(T-t)} \cdot e^{g(t,\lambda(t))} = Y(t) \leq 2e^{2r(T-t)}, \quad \forall t \in [0,T]$$

and thereby

$$0 \leq e^{g(t,\lambda(t))} \leq 1, \quad \forall t \in [0,T]. \tag{A.1}$$

Moreover, the lower bound of Y established in (4.8) leads to a strictly positive lower bound for $e^{g(t,\lambda(t))}$, that is,

$$e^{g(t,\lambda(t))} \geq \frac{\delta}{2}e^{-2rT} := \epsilon, \quad \forall t \in [0,T]. \tag{A.2}$$

This completes the proof of $\epsilon \leq e^{g(t,\lambda(t))} \leq 1$, for any $t \in [0,T]$.

Next we focus on the case $t = 0$. Note that (A.2) holds for any $t \in [0,T]$. Hence, it remains to show that $e^{g(0,\lambda_0)} < 1$. To that end, we recall (3.24) and apply Itô's formula to $e^{g(t,\lambda(t))}$:

$$\mathrm{d}e^{g(t,\lambda(t))} = e^{g(t,\lambda(t))}\left[g_t(t,\lambda(t)) + g_\lambda^\top(t,\lambda(t))\,\alpha(\lambda_\infty - \lambda(t)) + U(t,\lambda(t))^\top \lambda(t)\right]\mathrm{d}t + e^{g(t,\lambda(t))}U(t,\lambda(t))^\top \mathrm{d}\widetilde{N}(t)$$

$$= e^{g(t,\lambda(t))}\widehat{\mathcal{Z}}(t,\lambda(t))^\top \Gamma(t,\lambda(t))^{-1}\widehat{\mathcal{Z}}(t,\lambda(t))\mathrm{d}t + e^{g(t,\lambda(t))}U(t,\lambda(t))^\top \mathrm{d}\widetilde{N}(t).$$

Since $e^{g(t,\lambda(t))}U(t,\lambda(t))^\top$ is bounded, the jump component in the above equation is a martingale. Thus, integrating from t_1 to t_2 and conditioning on t_1, where $0 \leq t_1 \leq t_2 \leq T$, we have

$$0 \leq e^{g(t_2,\lambda(t_2))} - e^{g(t_1,\lambda(t_1))} = \mathbb{E}_{t_1}\left[\int_{t_1}^{t_2} e^{g(s,\lambda(s))}\widehat{\mathcal{Z}}(s,\lambda(s))^\top \Gamma(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}(s,\lambda(s))\mathrm{d}s\right] \leq 1. \tag{A.3}$$

Hence, we have $e^{g(t,\lambda(t))} \geq e^{g(0,\lambda_0)}$, for any $t \in [0,T]$. If $e^{g(0,\lambda_0)} = 1$, it then follows from (A.1) that $e^{g(t,\lambda(t))} \equiv 1$, for any $t \in [0,T]$. In that case, from (A.3), we obtain

$$\mathbb{E}\left[\int_0^T \widehat{\mathcal{Z}}(s,\lambda(s))^\top \Gamma(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}(s,\lambda(s))\mathrm{d}s\right] = 0.$$

On the other hand, if $e^{g(t,\lambda(t))} \equiv 1$, then

$$U(t,\lambda(t)) \equiv -\mathbf{1}_m = (-1,\ldots,-1)^\top \in \mathbb{R}^m, \quad \Gamma(t,\lambda(t)) \equiv \sigma\sigma^\top, \qquad \forall t \in [0,T].$$

Recall that the non-degeneracy of the precision matrix $(\sigma\sigma^\top)^{-1}$ is postulated. The above results lead to

$$\widehat{\mathcal{Z}}(t,\lambda(t)) = B - \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[\lambda(t) \bullet \nu(\mathrm{d}z)]\,\mathbf{1}_m = \mathbf{0}_k, \qquad \mathbb{P}\text{-a.s.}, \quad \forall t \in [0,T],$$

which cannot hold for all t since λ is a stochastic process. Therefore, $e^{g(0,\lambda_0)} < 1$ must hold true. $\qquad\square$

B   Proof of Lemma 3.2

Proof to Lemma 3.2. To analyze (3.24), we apply exponential transformation to study $\widetilde{g}(t,\lambda) := e^{g(t,\lambda)}$. Recalling (A.3), we have

$$\begin{aligned}
\widetilde{g}(t,\lambda) = \ e^{g(t,\lambda)} &= 1 - \mathbb{E}_{t,\lambda}\left[\int_t^T e^{g(s,\lambda(s))}\widehat{\mathcal{Z}}(s,\lambda(s))^\top \Gamma(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}(s,\lambda(s))\,\mathrm{d}s\right]\\
&= 1 - \mathbb{E}_{t,\lambda}\left[\int_t^T \widetilde{g}(s,\lambda(s))\widehat{\mathcal{Z}}^{\widetilde{g}}(s,\lambda(s))^\top \Gamma^{\widetilde{g}}(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}^{\widetilde{g}}(s,\lambda(s))\,\mathrm{d}s\right], \qquad (\text{B.1})
\end{aligned}$$

where we use $\mathbb{E}_{t,\lambda}$, instead of $\mathbb{E}_t$, to emphasize that the expectation is take under the condition $\lambda(t) = \lambda$, and the functions in the integrand are defined by

$$\begin{aligned}
\Gamma^{\widetilde{g}}(t,\lambda) &:= \sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[(U^{\widetilde{g}}(t,\lambda) + \mathbf{1}_m) \bullet \lambda \bullet \nu(\mathrm{d}z)]\eta(z)^\top,\\
\widehat{\mathcal{Z}}^{\widetilde{g}}(t,\lambda) &:= B + \int_{(-1,\infty)^m} \eta(z)\operatorname{Diag}[\lambda \bullet \nu(\mathrm{d}z)]\,U^{\widetilde{g}}(t,\lambda),\\
U^{\widetilde{g}}(t,\lambda) &:= \left(\frac{\widetilde{g}(t,\lambda + \beta_{(1)})}{\widetilde{g}(t,\lambda)} - 1, \cdots, \frac{\widetilde{g}(t,\lambda + \beta_{(m)})}{\widetilde{g}(t,\lambda)} - 1\right)^\top.
\end{aligned}$$

Next we apply Schauder's fixed point theorem to prove the existence of a solution to (B.1). To that end, we consider a Banach space $\mathcal{X} := \mathcal{S}_{\mathcal{F},L}(0,T;\mathbb{R})$ equipped with the sup norm $\|\cdot\|_\mathcal{X}$

$$\|\widetilde{g}\|_\mathcal{X} := \mathbb{E}\left[\sup_{t \in [0,T]} e^{-L(T-t)}\big|\widetilde{g}(t,\lambda(t))\big|\right].$$

where L is a positive constant.

Define a map $\mathcal{T}$ from $\mathcal{M} := \mathcal{S}_{\mathcal{F},L}(0,T;[\epsilon,1])$ onto itself as follows:

$$(\mathcal{T}\widetilde{g})(t) := 1 - \mathbb{E}_{t,\lambda}\left[\int_t^T \widetilde{g}(s,\lambda(s))\widehat{\mathcal{Z}}^{\widetilde{g}}(s,\lambda(s))^\top \Gamma^{\widetilde{g}}(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}^{\widetilde{g}}(s,\lambda(s))\,\mathrm{d}s\right].$$

Obviously, $\mathcal{M} = \mathcal{S}_{\mathcal{F},L}(0,T;[\epsilon,1])$ is a non-empty, convex, and compact space.

31

In what follows, we show that the map $\mathcal{T} : \mathcal{M} \subset \mathcal{X} \to \mathcal{M}$ is continuous. For such a purpose, let us choose any $\widetilde{g}_1, \widetilde{g}_2 \in \mathcal{M}$. Then, we have

$$
\|(\mathcal{T}\widetilde{g}_1) - (\mathcal{T}\widetilde{g}_2)\|_{\mathcal{X}}
$$

$$
= \mathbb{E}\left\{ \sup_{t\in[0,T]} e^{-L(T-t)} \left| \mathbb{E}_{t,\lambda}\left[\int_t^T \widetilde{g}_1(s,\lambda(s))\widehat{\mathcal{Z}}^{\widetilde{g}_1}(s,\lambda(s))^\top \Gamma^{\widetilde{g}_1}(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}^{\widetilde{g}_1}(s,\lambda(s))\mathrm{d}s \right. \right. \right.
$$

$$
\left. \left. \left. - \int_t^T \widetilde{g}_2(s,\lambda(s))\widehat{\mathcal{Z}}^{\widetilde{g}_2}(s,\lambda(s))^\top \Gamma^{\widetilde{g}_2}(s,\lambda(s))^{-1}\widehat{\mathcal{Z}}^{\widetilde{g}_2}(s,\lambda(s))\mathrm{d}s \right] \right| \right\}
$$

$$
\leq K_1 \mathbb{E}\left\{ \sup_{t\in[0,T]} e^{-L(T-t)}\mathbb{E}_{t,\lambda}\left[\int_t^T |\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))||\lambda(s)|^2 \mathrm{d}s \right] \right\}
$$

$$
\leq K_1 \mathbb{E}\left\{ \sup_{t\in[0,T]} e^{-L(T-t)}\mathbb{E}_{t,\lambda}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))| \int_t^T e^{L(T-s)}|\lambda(s)|^2 ds \right] \right\}
$$

$$
\leq K_1 \mathbb{E}\left\{ \sup_{t\in[0,T]} e^{-L(T-t)}\left\{ \mathbb{E}_{t,\lambda}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^2 \right] \right\}^{\frac{1}{2}}\left\{ \mathbb{E}_{t,\lambda}\left[\left(\int_t^T e^{L(T-s)}|\lambda(s)|^2 ds \right)^2 \right] \right\}^{\frac{1}{2}} \right\}
$$

$$
\leq K_1 \mathbb{E}\left\{ \sup_{t\in[0,T]} e^{-L(T-t)}\left\{ \mathbb{E}_{t,\lambda}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^2 \right] \right\}^{\frac{1}{2}}\left\{ \mathbb{E}_{t,\lambda}\left[\sup_{s\in[t,T]} |\lambda(s)|^4 \right] \right\}^{\frac{1}{2}}\left(\int_t^T e^{L(T-s)} ds \right) \right\}
$$

$$
\leq \frac{K_1}{L}(1 - e^{-LT})\left\{ \mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}_{t,\lambda}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^2 \right]^{\frac{\mathfrak{p}}{2}} \right] \right\}^{\frac{1}{\mathfrak{p}}}\left\{ \mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}_{t,\lambda}\left[\sup_{s\in[t,T]} |\lambda(s)|^4 \right]^{\frac{\mathfrak{q}}{2}} \right] \right\}^{\frac{1}{\mathfrak{q}}},
$$

where $\mathfrak{p}$ and $\mathfrak{q}$ are two positive constants such that $1/\mathfrak{p} + 1/\mathfrak{q} = 1$ and $\mathfrak{p} > 2$.

Next we apply the Doob martingale inequality and derive

$$
\left\{ \mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}_{t,\lambda}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^2 \right]^{\frac{\mathfrak{p}}{2}} \right] \right\}^{\frac{1}{\mathfrak{p}}}
$$

$$
\leq K_2 \left\{ \mathbb{E}\left[\mathbb{E}_T\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^2 \right]^{\frac{\mathfrak{p}}{2}} \right] \right\}^{\frac{1}{\mathfrak{p}}}
$$

$$
\leq K_2 \left\{ \mathbb{E}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))|^{\mathfrak{p}} \right] \right\}^{\frac{1}{\mathfrak{p}}}
$$

$$
\leq K_2 \left\{ \mathbb{E}\left[\sup_{s\in[0,T]} e^{-L(T-s)}|\widetilde{g}_1(s,\lambda(s)) - \widetilde{g}_2(s,\lambda(s))| \right] \right\}^{\frac{1}{\mathfrak{p}}}.
$$

On the other hand, since the SDE (2.3) for the Hawkes process satisfies the Lipschitz and linear growth conditions, from the moment estimates in the standard theory of SDEs (see, e.g., Theorem 1.9.4 in Platen and Bruti-Liberati (2010)), we can derive

$$
\left\{ \mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}_{t,\lambda}\left[\sup_{s\in[t,T]} |\lambda(s)|^4 \right]^{\frac{\mathfrak{q}}{2}} \right] \right\}^{\frac{1}{\mathfrak{q}}} \leq K_3 \left\{ \mathbb{E}\left[\sup_{t\in[0,T]} (1 + |\lambda(t)|^4)^{\frac{\mathfrak{q}}{2}} \right] \right\}^{\frac{1}{\mathfrak{q}}}
$$

$$
\leq K_3 \left\{ 1 + \mathbb{E}\left[\sup_{t\in[0,T]} |\lambda(t)|^4 \right] \right\}^{\frac{1}{2}}
$$

$$\leq K_3 \big\{ 1 + |\lambda_0|^4 \big\}^{\frac{1}{2}}.$$

Note that in all the above derivations, K_1, K_2, and K_3 are generic positive constants and their values may vary from line to line.

Therefore, combining the above derivations, we have

$$\|(\mathcal{T}\widetilde{g}_1) - (\mathcal{T}\widetilde{g}_2)\|_{\mathcal{X}} \leq C \{ \|\widetilde{g}_1 - \widetilde{g}_2\|_{\mathcal{X}} \}^{\frac{1}{p}},$$

where C is a positive constant depending on K_1, K_2, K_3, L, and λ_0. This indicates that $\mathcal{T}$ is a continuous map. Using Schauder's fixed point theorem, we can conclude that (B.1) admits a solution $\widetilde{g} \in \mathcal{S}_{\mathcal{F},L}(0, T; [\epsilon, 1])$. From the relationship between g and $\widetilde{g}$, we can confirm that (3.24) admits a solution $g := \log[\widetilde{g}]$.

Next we show the uniqueness of the solution $\widetilde{g}$. For that purpose, recall that the quadratic-loss minimization problem (2.15) belongs to the class of stochastic LQ control problems, which admits a unique optimal control/strategy (see Zhang et al. (2020)). Specifically, for (2.15), the unique optimal strategy π_c^* is given by (4.6). As in Step 3 of the proof for Theorem 4.1, we can show that (i) $\pi_c^* \in \mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^k)$; (ii) the corresponding wealth process, denoted by $\widehat{X}^*$, is square-integrable at the terminal time, i.e., $\mathbb{E}[(\widehat{X}^*(T))^2] < \infty$.

Combining Assertion (i) and the uniqueness of π_c^*, we have that the wealth equation (3.2) associated with π_c^* admits a unique solution $\widehat{X}^*(\cdot) \in \mathcal{S}_{\mathcal{F}}^2(0, T; \mathbb{R})$. On the other hand, combining Assertion (ii) and the uniqueness of $\widehat{X}^*$ implies that the adjoint equation (3.4) admits a unique solution such that $(p^*, q^*, u^*) \in \mathcal{S}_{\mathcal{F}}^2(0, T; \mathbb{R}) \times \mathcal{L}_{\mathcal{F}}^2(0, T; \mathbb{R}^n) \times \mathcal{L}_{\mathcal{F}}^{2,N}(0, T; \mathbb{R}^m)$.

Therefore, (3.24) must have a unique solution. Otherwise, suppose that there exist two solutions g_1 and g_2 to (3.24). Denote by $Y_i(t) := 2e^{2r(T-t)+g_i(t, \lambda(t))}$, for $i = 1, 2$, as defined in (3.16). Following the derivations in Section 3, we can show that both $p_1^* := Y_1 \widehat{X}^*$ and $p_2^* := Y_2 \widehat{X}^*$ are the solutions to the adjoint equation (3.4). This is a contradiction and violates the uniqueness of p^*. This completes the proof of the uniqueness. $\qquad \square$

C   Connection with the HJB equation approach

In this appendix, we briefly discuss the connection between the stochastic maximum approach and the HJB equation approach, since both approaches can be applied to solve the quadratic-loss minimization problem (2.15). We refer interested readers to Aït-Sahalia and Hurd (2016), Cao et al. (2020), and Liu et al. (2021) for the applications of the HJB approach to optimal investment problems under the Hawkes jump models.

Let us start by considering the dynamic version of the quadratic-loss minimization problem (2.15) at (t, x, λ), defined as follows:

$$\mathcal{V}(t, x, \lambda) := \min_{\pi \in \mathcal{A}} \mathbb{E}\big[\widehat{X}^2(T) \big| \widehat{X}(t) = x, \, \lambda(t) = \lambda \big], \tag{C.1}$$

where the dynamics of $\widehat{X}$ is given by (3.2).

Firstly, we write down the HJB equation for problem (2.15) (or problem (C.1) to be precise):

$$\inf_{\pi \in \mathbb{R}} \mathcal{L}^\pi[\mathcal{V}(t,x,\lambda)] = 0, \tag{C.2}$$

with the terminal condition $\mathcal{V}(T,x,\lambda) = x^2$. The partial differential operator $\mathcal{L}^\pi$ acting on any smooth function φ and admissible π is defined by

$$\begin{aligned}
\mathcal{L}^\pi[\varphi(t,x,\lambda)] &:= \varphi_t(t,x,\lambda) + \big(rx + \pi^\top B\big)\varphi_x(t,x,\lambda) + \big(\alpha\lambda_\infty + (\beta - \alpha)\lambda\big)^\top \varphi_\lambda(t,x,\lambda) \\
&\quad + \frac{1}{2}\pi^\top \sigma\sigma^\top \pi \varphi_{xx}(t,x,\lambda) + \sum_{j=1}^m \int_{(-1,\infty)} \big[\varphi(t, x + (\pi^\top \eta(z))_{(j)}, \lambda + \beta_{(j)}) \\
&\quad - \varphi(t,x,\lambda) - \varphi_x(t,x,\lambda)(\pi^\top \eta(z))_{(j)} - \beta_{(j)}^\top \varphi_\lambda(t,x,\lambda)\big]\lambda_j \nu_j(\mathrm{d}z_j)
\end{aligned}$$

with $(\pi^\top \eta(z))_{(j)}$ being the jth component of the vector $\pi^\top \eta(z)$ and dependent on only the jth coordinate of z (i.e., z_j), and $\beta_{(j)}$ being the jth column vector of the matrix β.

We try the following *ansatz* for the value function $\mathcal{V}$ in (C.1):

$$\mathcal{V}(t,x,\lambda) = e^{2r(T-t)+\kappa(t,\lambda)}x^2,$$

with the terminal condition $\kappa(T,\lambda) = 0$. Substituting this into (C.2) and after some algebraic manipulation, we obtain

$$\inf_{\pi \in \mathbb{R}} \left\{ 2x\pi^\top B + \pi^\top \sigma\sigma^\top \pi + \sum_{j=1}^m \int_{(-1,\infty)} \big[(\pi^\top \eta(z))_{(j)}^2 e^{\kappa(t,\lambda+\beta_{(j)})-\kappa} + 2x(\pi^\top \eta(z))_{(j)}(e^{\kappa(t,\lambda+\beta_{(j)})-\kappa} - 1)\big]\lambda_j \nu_j(\mathrm{d}z_j) \right\}$$

$$+ x^2 \left\{ \kappa_t + \kappa_\lambda^\top \big(\alpha\lambda_\infty + (\beta - \alpha)\lambda\big) + \sum_{j=1}^m \int_{(-1,\infty)} \big[e^{\kappa(t,\lambda+\beta_{(j)})-\kappa} - 1 - \beta_{(j)}^\top \kappa_\lambda\big]\lambda_j \nu_j(\mathrm{d}z_j) \right\} = 0,$$

where we have suppressed the arguments (t,λ) for κ and its partial derivatives, except the post-jump ones.

The first-order condition to the above minimization problem yields

$$2x\,\widehat{\mathcal{Z}}^\kappa(t,\lambda) + 2\Gamma^\kappa(t,\lambda)\pi = \mathbf{0_m},$$

where

$$\Gamma^\kappa(t,\lambda) := \sigma\sigma^\top + \int_{(-1,\infty)^m} \eta(z)\,\mathrm{Diag}[(U^\kappa(t,\lambda) + \mathbf{1}_m) \bullet \lambda \bullet \nu(\mathrm{d}z)]\,\eta(z)^\top,$$

$$\widehat{\mathcal{Z}}^\kappa(t,\lambda) := B + \int_{(-1,\infty)^m} \eta(z)\,\mathrm{Diag}[\lambda \bullet \nu(\mathrm{d}z)]\,U^\kappa(t,\lambda),$$

34

are defined similarly as (3.20) and (3.21) but with $U(t, \lambda)$ in (3.19) replaced by

$$U^\kappa(t, \lambda) := \left(e^{\kappa(t, \lambda+\beta_{(1)})-\kappa(t, \lambda)} - 1, \cdots, e^{\kappa(t, \lambda+\beta_{(m)})-\kappa(t, \lambda)} - 1 \right)^\top.$$

Thus, the optimal strategy can be represented in the following feedback form:

$$\pi_c^*(t, x, \lambda) = -\Gamma^\kappa(t, \lambda)^{-1} \, \widehat{\mathcal{Z}}^\kappa(t, \lambda) \, x.$$

Plugging $\pi_c^*(t, x, \lambda)$ into (C.2) gives

$$\kappa_t(t, \lambda) + \kappa_\lambda(t, \lambda)^\top \alpha(\lambda_\infty - \lambda) + U^\kappa(t, \lambda)^\top \lambda = \widehat{\mathcal{Z}}^\kappa(t, \lambda)^\top \Gamma^\kappa(t, \lambda)^{-1} \widehat{\mathcal{Z}}^\kappa(t, \lambda). \tag{C.3}$$

Indeed, the above equation (C.3) is exactly the same as (3.24). It then follows from the existence and uniqueness result in Appendix B that $\kappa(t, \lambda) = g(t, \lambda)$, for any $(t, \lambda) \in [0, T] \times \mathbb{R}_+^m$. Therefore, the stochastic maximum principle approach and the HJB equation approach lead to the same solution for problem (2.15).

Indeed, such a conclusion is guaranteed by the relationship between the two approaches. That is, the value function and the adjoint process are related as follows:

$$p^*(t) = 2e^{2r(T-t)+g(t,\lambda(t))} \widehat{X}^*(t) = \mathcal{V}_x(t, \widehat{X}^*(t), \lambda(t)),$$
$$q^*(t) = \sigma^\top \pi_c^*(t) Y(t) = \sigma^\top \pi_c^*(t) \mathcal{V}_{xx}(t, \widehat{X}^*(t), \lambda(t)),$$
$$u_l^*(t, z_l) = \widehat{X}^*(t-) V_l(t) + \sum_{i=1}^k \pi_{ci}^*(t) \eta_{il}(z_l) \left(Y(t-) + V_l(t) \right)$$
$$= \mathcal{V}_x(t, \widehat{X}^*(t-) + (\pi_c^*(t)^\top \eta(z))_{(l)}, \lambda(t) + \beta_{(l)}) - \mathcal{V}_x(t, \widehat{X}^*(t-), \lambda(t)).$$

Moreover, substituting $p^*(t) = \mathcal{V}_x(t, \widehat{X}^*(t), \lambda(t))$ into the adjoint equation (3.4) and matching the drift gives the relationship between the Hamiltonian and the value function as below:

$$\mathcal{L}^{\pi_c^*}[\mathcal{V}_x(t, \widehat{X}^*(t), \lambda(t))] = -\mathcal{H}_x(t, \widehat{X}^*(t), \pi_c^*(t), p^*(t), q^*(t), u^*(t)).$$

One can refer to Theorem 3.1 in Framstad et al. (2004) and Theorem 5.6 in Øksendal and Sulem (2019) for the relationship between the stochastic maximum principle and dynamic programming principle in general jump-diffusion control systems.

References

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.

Aït-Sahalia, Y. and Hurd, T. R. (2016). Portfolio choice in markets with contagion. *Journal of Financial Econometrics*, 14(1):1–28.

Aït-Sahalia, Y., Laeven, R. J., and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183(2):151–167.

Azizpour, S., Giesecke, K., and Schwenkler, G. (2018). Exploring the sources of default clustering. *Journal of Financial Economics*, 129(1):154–183.

Bacry, E., Mastromatteo, I., and Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005.

Bajeux-Besnainou, I. and Portait, R. (1998). Dynamic asset allocation in a mean-variance framework. *Management Science*, 44(11-part-2):S79–S95.

Basak, S. and Chabakauri, G. (2010). Dynamic mean-variance asset allocation. *Review of Financial Studies*, 23(8):2970–3016.

Björk, T., Murgoci, A., and Zhou, X. Y. (2014). Mean–variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance*, 24(1):1–24.

Bo, L., Capponi, A., and Chen, P.-C. (2019). Credit portfolio selection with decaying contagion intensities. *Mathematical Finance*, 29(1):137–173.

Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912.

Cao, J., Landriault, D., and Li, B. (2020). Optimal reinsurance-investment strategies for dynamic contagion claims. *Insurance: Mathematics and Economics*, 93:206–215.

Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. (2005). Estimating value-at-risk: A point process approach. *Quantitative Finance*, 5(2):227–234.

Chavez-Demoulin, V. and McGill, J. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance*, 36(12):3415–3426.

Cont, R. and Tankov, P. (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Daley, D. J. and Vere-Jones, D. (2007). *An Introduction to the Theory of Point Processes, Volume II: General Theory and Structure*. Springer Science & Business Media.

Dassios, A. and Zhao, H. (2011). A dynamic contagion process. *Advances in Applied Probability*, 43(3):814–846.

Dassios, A. and Zhao, H. (2012). Ruin by dynamic contagion claims. *Insurance: Mathematics and Economics*, 51(1):93–106.

Embrechts, P., Liniger, T., and Lin, L. (2011). Multivariate Hawkes processes: An application to financial data. *Journal of Applied Probability*, 48(A):367–378.

Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.

Framstad, N. C., Øksendal, B., and Sulem, A. (2004). Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *Journal of Optimization Theory and Applications*, 121(1):77–98.

Gao, X., Zhou, X., and Zhu, L. (2018). Transform analysis for Hawkes processes with applications in dark pool trading. *Quantitative Finance*, 18(2): 265–282.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.

Hawkes, A. G. (2018). Hawkes processes and their applications to finance: A review. *Quantitative Finance*, 18(2):193–198.

Karatzas, I. and Shreve, S. E. (1998). *Methods of Mathematical Finance*, volume 39. Springer.

Kou, S. G. (2007). Jump-diffusion models for asset pricing in financial engineering. *Handbooks in Operations Research and Management Science*, 15:73–116.

Li, D. and Ng, W.-L. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, 10(3):387–406.

Lim, A. E. and Zhou, X. Y. (2002). Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27(1):101–120.

Liu, G., Jin, Z., and Li, S. (2021). Household lifetime strategies under a self-contagious market. *European Journal of Operational Research*, 288(3):935–952.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Øksendal, B. K. and Sulem, A. (2019). *Applied Stochastic Control of Jump Diffusions*, 3rd edition. Springer.

Platen, E., and Bruti-Liberati, N. (2010). *Numerical Solution of Stochastic Differential Equations with Jumps in Finance (Vol. 64)*. Springer Science & Business Media.

Shen, Y. and Siu, T. K. (2013). The maximum principle for a jump-diffusion mean–field model and its application to the mean–variance problem. *Nonlinear Analysis: Theory, Methods & Applications*, 86:58–73.

Shen, Y. and Zeng, Y. (2014). Optimal investment–reinsurance with delay for mean–variance insurers: A maximum principle approach. *Insurance: Mathematics and Economics*, 57:1–12.

Shen, Y. and Zou, B. (2021). Mean-variance investment and risk control strategies – A time-consistent approach via a forward auxiliary process. *Insurance: Mathematics and Economics*, 97:68–80.

Yin, G. and Zhou, X. Y. (2004). Markowitz's mean-variance portfolio selection with regime switching: From discrete-time models to their continuous-time limits. *IEEE Transactions on Automatic Control*, 49(3):349–360.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*, volume 43. Springer Science & Business Media.

Zhang, F., Dong, Y. and Meng, Q. (2020). Backward stochastic Riccati equation with jumps associated with stochastic linear quadratic optimal control with jumps and random coefficients. *SIAM Journal on Control and Optimization*, 58(1):393–424.

Zhang, X., Xiong, J. and Shen, Y. (2018). Bond and option pricing for interest rate model with clustering effects. *Quantitative Finance*, 18(6):969–981.

Zhou, X. Y. and Li, D. (2000). Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42(1):19–33.

Zhou, X. Y. and Yin, G. (2003). Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization*, 42(4):1466–1482.

Zhu, L. (2013). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance: Mathematics and Economics*, 53(3):544–550.