



SOLVING STOCHASTIC OPTIMAL CONTROL VIA THE MUTUAL EXCITATION MALLIAVIN CALCULUS METHOD AND ITS APPLICATIONS IN INSURANCE

CHENHAN LIU, MENGRO LIU, JICHAO ZHANG*

School of Mathematics and Statistics, Beihua University, Jilin, Jilin

Abstract. This paper solves an optimal reinsurance-investment problem for insurers under partial information with Hawkes-driven claim clustering. Since the true risk intensity is unobservable, a Gamma projection filter is used to estimate its conditional distribution. Methodologically, Malliavin calculus is introduced to derive a stochastic Hamilton-Jacobi-Bellman (HJB) equation that directly handles the non-Markovianity induced by partial-information filtering. This extends classical verification theory to non-Markovian settings. A verification theorem is established, linking smooth solutions of the stochastic HJB equation to optimal admissible strategies. Numerical experiments validate our theoretical results.

Keywords. Gamma projection filter; Hawkes process; Malliavin calculus; optimal reinsurance; Stochastic HJB equation.

2020 MSC. 34B15, 34B18, 47H10.

1. INTRODUCTION

In the fields of mathematics and actuarial science, the optimal investment and reinsurance problem remains a core research topic. Its primary objective is to achieve effective risk hedging and robust asset growth through dynamic adjustment of reinsurance retention and asset allocation. The expected utility maximization criterion is widely adopted, reflecting the risk preferences of decision-makers. However, real-world risk evolution exhibits complex characteristics. On one hand, insurance claim risks demonstrate significant clustering and contagion effects, such as chain claims triggered by catastrophes. Traditional Poisson processes, lacking a self-exciting mechanism, cannot capture this. On the other hand, the true risk intensity is often unobservable; insurers can only access partial information (claim times, types, and losses). This incompleteness poses a significant challenge for optimal strategy construction. To capture claim clustering, the Hawkes process has received attention. Hawkes [1] reviewed its applications in finance, confirming its superiority over Poisson processes in capturing jump clustering. Ji et

*Corresponding author.

E-mail address: zhangjichao87@163.com (J. Zhang).

Received May 19, 2026; Accepted September 12, 2026.

al. [2] derived explicit optimal investment-reinsurance solutions for loss-averse insurers. Bernis et al. [3] characterized optimal reinsurance contracts with a three-piece linear structure. Liu et al. [4] incorporated Hawkes processes into the Gamma–Omega model. Zhou et al. [5] studied a reinsurance-investment game with delay under asymmetric information. These studies demonstrate that Hawkes processes effectively characterize claim clustering, but most assume full observability of risk intensity. Under partial information, filtering becomes a core challenge for optimal control. Early works include Sass et al. [6], who used unnormalized filtering and Malliavin derivatives to solve a terminal wealth maximization problem. Peng et al. [7] studied optimal proportional reinsurance under partial information using Malliavin calculus. Jang et al. [8] compared optimal reinsurance under partial vs. complete information. Brachetta et al. [9] characterized optimal reinsurance with claim clustering via Kushner–Stratonovich filtering and BSDEs. While theoretically rigorous, the Kushner–Stratonovich equation is infinite-dimensional, making direct solution computationally challenging. Projection filtering addresses this by assuming the posterior belongs to a parametric family. This paper adopts the Gamma distribution, which maintains non-negativity and achieves moment closure. In solving the non-Markovian optimal control problem resulting from filtering, Malliavin calculus provided a set of powerful tools. Chen et al. [10] employed a general jump-diffusion process and used Malliavin calculus to solve the mean-variance optimal reinsurance and investment problem, extending research methods for non-Markovian problems in actuarial science. Bai et al. [11] studied the mean-variance optimal reinsurance-investment problem with a no-shorting constraint using viscosity solutions and a verification theorem. Colaneri et al. [12] constructed a stochastic factor model within an exponential forward utility framework to solve the optimal investment-reinsurance strategy. These works demonstrate the potential of Malliavin methods, yet they typically used Malliavin derivatives primarily for representation or necessary condition derivation, rather than as a core tool to reconstruct the HJB equation itself under partial information. Specifically, let $D_{t,0}$ and $D_{t,y}^{(k)}$ denote the Malliavin derivatives with respect to the observed Brownian motion and the observed Poisson random measure, respectively. These operators allow us to differentiate the value function with respect to randomness embedded in the observation flow. By applying the Clark–Ocone representation to the value process and matching martingale components, we derive a stochastic HJB equation (Theorem 3.1) that explicitly incorporates Malliavin correction terms. This equation naturally accounts for the non-Markovianity induced by filtering, which is impossible in the classical HJB framework. A key theoretical contribution of this paper is a verification theorem (Theorem 3.2) for this stochastic HJB equation. While classical verification theorems require the state process to be Markovian, our theorem applies directly to non-Markovian settings. It establishes that any sufficiently smooth solution to the stochastic HJB equation, together with the associated admissible control, coincides with the value function of the original partial-information problem. This extends the scope of stochastic control theory beyond the traditional Markovian paradigm. Unlike existing approaches that either approximate the filter as Markovian or solve infinite-dimensional Kushner–Stratonovich equations, our Malliavin-based stochastic HJB method preserves the original non-Markovian structure and yields explicit solutions under exponential utility. More broadly, this framework is generalizable to other non-Markovian control problems with delayed information, model uncertainty, or path-dependent criteria. In summary, existing research focuses separately on Hawkes modeling, partial information filtering, or Malliavin control methods, but fails to integrate these

three to address the complex decision-making problem where claim clustering and information incompleteness coexist (see also Liu et al. [13] for related work on contagious claims in a game-theoretic setting). The main contributions of this paper are fourfold. First, we construct a bivariate marked Hawkes claim intensity model incorporating internal and external excitation mechanisms. Second, we adopt a Gamma projection filter for partial information, derive the filtering equations, and prove existence and uniqueness. Third, we introduce Malliavin calculus to construct a stochastic HJB equation and establish a verification theorem for non-Markovian settings, obtaining explicit optimal strategies via separation of variables. Fourth, we show that the optimal investment strategy is independent of the filtering state (reducing to the Merton formula), while the optimal reinsurance strategy exhibits a three-piece structure determined by the safety loading.

The structure of this paper is as follows: Section 2 presents the model setup, including the Hawkes claim process, financial market, and wealth dynamics. Section 3 develops the Gamma projection filter for partial information estimation. Section 4 introduces Malliavin calculus and derives the stochastic HJB equation, followed by a verification theorem and explicit optimal strategies. Section 5 provides numerical experiments. Section 6 concludes.

2. PRELIMINARIES

Consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a right-continuous filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions, where $\mathbb{P}$ is the real-world probability measure. On this space, the following mutually independent processes are defined: $\{W_t\}_{t \geq 0}$ is a standard Brownian motion, and two independent marked point processes $N^{(1)}$ and $N^{(2)}$, corresponding to internally-excited claim events and externally-excited claim events, respectively. The total claim intensity λ_t is given by the following equation:

$$\lambda_t = \beta + (\lambda_0 - \beta)e^{-\alpha t} + \sum_{j=1}^{N_t^{(1)}} e^{-\alpha(t-T_j^{(1)})} \ell(Z_j^{(1)}) + \sum_{j=1}^{N_t^{(2)}} e^{-\alpha(t-T_j^{(2)})} Z_j^{(2)}. \quad (2.1)$$

Its stochastic differential form is

$$d\lambda_t = \alpha(\beta - \lambda_t)dt + \int_0^\infty \ell(z) \tilde{N}^{(1)}(dt, dz) + \int_0^\infty z \tilde{N}^{(2)}(dt, dz), \quad \lambda_0 > 0, \quad (2.2)$$

where $\alpha > 0$ is the decay rate, $\beta > 0$ is the long-term mean; $\ell : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is the internal excitation function, typically taken as $\ell(z) = \eta z$, $\eta > 0$; $Z_j^{(k)}$ is the mark (shock size) of the j -th event of type k , following distribution $F_{Z^{(k)}}(dz)$; $\tilde{N}^{(k)}(dt, dz) := N^{(k)}(dt, dz) - \mathbf{v}_t^{(k)}(dz)dt$ is the compensated random measure; the stochastic intensity measure is $\mathbf{v}_t^{(k)}(dz) = \lambda_t \cdot \kappa^{(k)} \cdot F_{Z^{(k)}}(dz)$, and $\kappa^{(1)} + \kappa^{(2)} = 1$. The insurance company cannot directly observe the true claim intensity process λ_t and can only access partial information, i.e., it can only observe: the actual occurrence times of claim events $\{T_j\}$; the type attribute of each claim event (internally-excited type ($k = 1$) or externally-excited type ($k = 2$)); and the loss amount of each claim Y_j , which may be related to the mark $Z_j^{(k)}$. Under partial information, all strategies that are $\mathcal{F}_t$ -measurable are infeasible. The insurance company must use strategies based on the observed information $\mathcal{F}_t^o$. The observable information flow $\mathbb{F}^o = \{\mathcal{F}_t^o\}_{t \geq 0}$ is defined as: $\mathcal{F}_t^o = \sigma(\{N_s, \{T_j \leq s\}, \{K_j\}, \{Y_j\} : j \geq 1, s \leq t\})$, where $K_j \in \{1, 2\}$ represents the type of the j -th claim event. Consider a financial market consisting of one risk-free asset and one risky

asset. The price of the risk-free asset B_t satisfies: $dB_t = rB_t dt$, $B_0 = 1$, where $r > 0$ is the constant risk-free interest rate. The price of the risky asset S_t follows a geometric Brownian motion: $dS_t = \mu S_t dt + \sigma S_t dW_t$, $S_0 = s_0 > 0$, where $\mu > r$ and $\sigma > 0$ are constants. We assume that the Brownian motion W_t driving the financial market and the Hawkes process driving the claims are mutually independent. The insurance company adopts proportional reinsurance (with retention proportion $q_t \in [0, 1]$) and invests its wealth in the financial market (with investment amount θ_t). The premium income rate is c , and the reinsurance premium rate is calculated based on the expected value principle using the filter estimate. The wealth process X_t satisfies:

$$\begin{aligned} dX_t = & \left[rX_t + \theta_t(\mu - r) + c - (1 + \theta_R)(1 - q_t)\hat{\lambda}_t \mathbb{E}[Y] \right] dt \\ & + \sigma \theta_t dW_t - q_t \int_0^\infty y \tilde{N}^o(dt, dy), \quad X_0 = x_0, \end{aligned} \quad (2.3)$$

where $\tilde{N}^o(dt, dy) = N(dt, dy) - \hat{\lambda}_t \phi_Y(dy)dt$ is the compensated claim measure under the observed information, $\phi_Y(dy) = \sum_{k=1}^2 \kappa^{(k)} f_{Y|k}(y)dy$. The control problem of this paper is to find an $\mathcal{F}_t^o$ -admissible strategy $\pi_t = (q_t, \theta_t)$ to maximize the terminal expected exponential utility:

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}[U(X_T)], \quad U(x) = -\frac{1}{\gamma} e^{-\gamma x}, \quad \gamma > 0. \quad (2.4)$$

3. THEORETICAL FRAMEWORK OF PROJECTION FILTERING

Under partial information, we need to compute the conditional posterior distribution of the intensity λ_t based on the observation filtration $\mathbb{F}^o = \{\mathcal{F}_t^o\}$. Since the Kushner-Stratonovich equation is infinite-dimensional, this study employs the projection filtering method. Its core idea is to assume that the posterior distribution belongs to a parametric family of distributions, and derive the evolution equations of the parameters by forcing the approximate distribution to be consistent with the "projection" of the true posterior dynamics on key statistics.

Definition 3.1. Assume that, for given $\mathcal{F}_t^o$, the conditional distribution of λ_t approximately follows a Gamma distribution:

$$p(\lambda_t | \mathcal{F}_t^o) \approx q_t(\lambda) = \frac{b_t^{a_t}}{\Gamma(a_t)} \lambda^{a_t-1} e^{-b_t \lambda}, \quad \lambda > 0, \quad (3.1)$$

where $a_t > 0$ and $b_t > 0$ are $\mathcal{F}_t^o$ -adapted stochastic processes. Under this assumption,

$$\hat{\lambda}_t = \mathbb{E}[\lambda_t | \mathcal{F}_t^o] = \frac{a_t}{b_t}, \quad P_t = \text{Var}(\lambda_t | \mathcal{F}_t^o) = \frac{a_t}{b_t^2}. \quad (3.2)$$

For any smooth function $\varphi(\lambda)$, define the conditional expectation $\pi_t(\varphi) = \mathbb{E}[\varphi(\lambda_t) | \mathcal{F}_t^o]$. By nonlinear filtering theory, $\pi_t(\varphi)$ satisfies the Kushner-Stratonovich equation:

$$d\pi_t(\varphi) = \pi_t(\mathcal{L}\varphi)dt + \sum_{k=1}^2 \int_0^\infty \left[\frac{\pi_{t-}(\varphi \lambda \kappa^{(k)} f_{Z^{(k)}})}{\pi_{t-}(\lambda \kappa^{(k)})} - \pi_{t-}(\varphi) \right] \left(dN_t^{(k)} - \pi_{t-}(\lambda \kappa^{(k)})dt \right), \quad (3.3)$$

where $\mathcal{L}$ is the infinitesimal generator of λ_t . For λ_t given in Equation (2.2),

$$\mathcal{L}\varphi(\lambda) = \alpha(\beta - \lambda)\varphi'(\lambda) + \lambda \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \left[\varphi(\lambda + \Delta^{(k)}(z)) - \varphi(\lambda) \right] F_{Z^{(k)}}(dz). \quad (3.4)$$

Lemma 3.1. Suppose that the evolution of filtering parameters (a_t, b_t) is given by the following system of stochastic differential equations:

$$\begin{cases} da_t = A_t dt + \sum_{k=1}^2 \int_0^\infty A_t^{(k)}(z) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \\ db_t = B_t dt + \sum_{k=1}^2 \int_0^\infty B_t^{(k)}(z) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \end{cases} \quad (3.5)$$

where $A_t, B_t, A_t^{(k)}(z), B_t^{(k)}(z)$ are adapted and satisfy certain integrability conditions. Then, under the Gamma distribution assumption $q_t(\lambda) = \text{Gamma}(\lambda; a_t, b_t)$, the stochastic differentials of the conditional expectations of sufficient statistics $\varphi_1 = \ln \lambda$ and $\varphi_2 = \lambda$ are:

$$\begin{cases} d\mathbb{E}_{q_t}[\lambda] = \left(\frac{A_t}{b_t} - \frac{a_t B_t}{b_t^2} \right) dt + \sum_{k=1}^2 \int_0^\infty J_\lambda^{(k)}(z) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \\ d\mathbb{E}_{q_t}[\ln \lambda] = \left[\psi'(a_t) A_t - \frac{1}{b_t} B_t \right] dt + \sum_{k=1}^2 \int_0^\infty J_{\ln}^{(k)}(z) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \end{cases} \quad (3.6)$$

where $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$ is the digamma function, $\psi'(x)$ is its derivative (the trigamma function), and the jump coefficients are

$$J_\lambda^{(k)}(z) = \frac{A_t^{(k)}(z)}{b_{t-}} - \frac{a_{t-} B_t^{(k)}(z)}{b_{t-}^2} + \frac{a_{t-} (B_t^{(k)}(z))^2}{b_{t-}^3}, \quad (3.7)$$

$$J_{\ln}^{(k)}(z) = \ln \left(1 + \frac{A_t^{(k)}(z)}{a_{t-}} \right) - \ln \left(1 + \frac{B_t^{(k)}(z)}{b_{t-}} \right) - \frac{A_t^{(k)}(z)}{a_{t-}} + \frac{B_t^{(k)}(z)}{b_{t-}}. \quad (3.8)$$

Proof. For $f(a, b) = a/b = \mathbb{E}_{q_t}[\lambda]$, partial derivatives are computed as $\frac{\partial f}{\partial a} = \frac{1}{b}$, $\frac{\partial f}{\partial b} = -\frac{a}{b^2}$, and $\frac{\partial^2 f}{\partial b^2} = \frac{2a}{b^3}$. Since da_t and db_t have no continuous diffusion components, the continuous quadratic variation terms vanish. Applying Itô's formula for jump processes,

$$d \left(\frac{a_t}{b_t} \right) = \frac{\partial f}{\partial a} da_t + \frac{\partial f}{\partial b} db_t + \sum_{s \leq t} \left[\Delta f(a_s, b_s) - \frac{\partial f}{\partial a} \Delta a_s - \frac{\partial f}{\partial b} \Delta b_s \right].$$

For a jump of type k with mark z , $\Delta f = \frac{a_{t-} + A_t^{(k)}(z)}{b_{t-} + B_t^{(k)}(z)} - \frac{a_{t-}}{b_{t-}}$. The continuous drift term is obtained via second-order Taylor expansion of the jump correction in $B_t^{(k)}(z)$, the jump coefficient $J_\lambda^{(k)}(z)$ given in (3.7). For $g(a, b) = \psi(a) - \ln b = \mathbb{E}_{q_t}[\ln \lambda]$, we compute $\frac{\partial g}{\partial a} = \psi'(a)$, $\frac{\partial g}{\partial b} = -\frac{1}{b}$, and $\frac{\partial^2 g}{\partial b^2} = \frac{1}{b^2}$. Applying Itô's formula:

$$d(\psi(a_t) - \ln b_t) = \psi'(a_t) da_t - \frac{1}{b_t} db_t + \sum_{s \leq t} \left[\Delta g(a_s, b_s) - \psi'(a_{s-}) \Delta a_s + \frac{1}{b_{s-}} \Delta b_s \right],$$

where

$$\Delta g = \left[\psi(a_{t-} + A_t^{(k)}(z)) - \ln(b_{t-} + B_t^{(k)}(z)) \right] - \left[\psi(a_{t-}) - \ln b_{t-} \right].$$

Note that $\psi(a + A) \approx \psi(a) + \psi'(a)A + \frac{1}{2}\psi''(a)A^2 + \dots$, $\ln(b + B) \approx \ln b + \frac{B}{b} - \frac{B^2}{2b^2} + \dots$. Thus $J_\lambda^{(k)}(z)$ and $J_{\ln}^{(k)}(z)$ in (3.8). This completes the proof.

Remark 3.1. The jump coefficient $J_{\text{In}}^{(k)}(z)$ in (3.9) is the exact jump correction obtained from Itô's formula within the Gamma parametric family. This is a purely mechanical consequence of the assumed parameter dynamics (3.5). It is important to note that the final projection filtering equations presented below in (3.14) are not obtained by substituting a particular choice of $A_t^{(k)}(z)$ and $B_t^{(k)}(z)$ into this general expression. As we can see later the conjugate Gamma update rule makes $J_{\text{In}}^{(k)}(z)$ vanish identically; the actual jump term in the projected filter is instead derived by matching the true filtering dynamics, as shown in Lemma 2.3. This two-step procedure — first deriving general moment formulas, then projecting the true filter onto the parametric family — is the essence of the projection filtering methodology. To close the projection filtering equations, it is necessary to specify how the parameters (a_t, b_t) evolve upon the occurrence of a claim event. The following sublemma provides the theoretical justification for the conjugate-type update rule adopted in this paper.

Sublemma 3.1. (Conjugate Gamma jump update) *Assume that the posterior distribution of λ_t is approximated by a Gamma distribution $\text{Gamma}(a_{t-}, b_{t-})$ just before a claim event of type k with mark z . Suppose that the jump in the sufficient statistic $\ln \lambda$ induces a parameter update $(a_{t-}, b_{t-}) \mapsto (a_t, b_t)$. If the update is required to satisfy the following two consistency conditions:*

- (i) *Mean preservation: the conditional mean $\mathbb{E}_{q_t}[\lambda]$ does not undergo an instantaneous jump due solely to the parameter update, i.e., $\frac{a_t}{b_t} = \frac{a_{t-}}{b_{t-}}$;*
- (ii) *Conjugate scaling: the parameter update preserves the Gamma family structure and is consistent with the likelihood increment proportional to λ , as in the standard Bayesian conjugate update,*

then the unique update is given by $a_t = a_{t-}e^{\Delta^{(k)}(z)/b_{t-}}$ and $b_t = b_{t-}e^{\Delta^{(k)}(z)/b_{t-}}$, where $\Delta^{(k)}(z)$ is the jump size in the intensity induced by the mark z .

Proof. Condition (i) directly implies $a_t/b_t = a_{t-}/b_{t-}$. Hence $a_t = ca_{t-}$ and $b_t = cb_{t-}$ for some common scaling factor $c > 0$. Under the Gamma conjugate framework, the likelihood of observing a claim event is proportional to λ , which induces an additive update of the shape parameter by 1 in the standard setting. In the present marked Hawkes setting, the mark z enters through the intensity jump $\Delta^{(k)}(z)$. The scaling factor is therefore determined by matching the first-order moment of the log-likelihood increment, yielding $c = e^{\Delta^{(k)}(z)/b_{t-}}$. Substituting back gives the claimed update.

Remark 3.2. This update rule is not an ad hoc assumption but a direct consequence of requiring that the projected filter remain within the Gamma family while preserving the mean of the sufficient statistic $\ln \lambda$. It is exactly the conjugate updating rule for the Gamma distribution, which is standard in projection filtering (see, e.g., [14]). The same jump form appears in the filtering equations (3.9) below, ensuring moment closure and consistency with the true Kushner–Stratonovich dynamics.

The jump amplitudes are then obtained as follows:

$$A_t^{(k)}(z) = a_{t-} \left(e^{\Delta^{(k)}(z)/b_{t-}} - 1 \right), \quad B_t^{(k)}(z) = b_{t-} \left(e^{\Delta^{(k)}(z)/b_{t-}} - 1 \right). \quad (3.9)$$

Consequently, this yields the following key identity:

$$\frac{A_t^{(k)}(z)}{a_{t-}} = \frac{B_t^{(k)}(z)}{b_{t-}} = e^{\Delta^{(k)}(z)/b_{t-}} - 1. \quad (3.10)$$

Substituting (3.9) into the general jump coefficients (3.7) and (3.8), the following result is then obtained: $J_\lambda^{(k)}(z) = 0$ and $J_{\ln}^{(k)}(z) = 0$. That is, within the Gamma parametric family, the conjugate jump update produces no jump contribution to either $\mathbb{E}_{q_t}[\lambda]$ or $\mathbb{E}_{q_t}[\ln \lambda]$ through the internal moment formulas. This cancellation property is a distinctive feature of the Gamma conjugate prior: the logarithmic sufficient statistic absorbs the jump entirely into the parameter scaling, leaving the moments unchanged at the instant of the jump. The cancellation observed above does not imply that the projected filter's moments should have no jump terms. On the contrary, the true posterior distribution (governed by the Kushner–Stratonovich equation) does exhibit jump updates upon claim arrivals. The projection filter must match these true moment evolutions, not merely its own internal moment formulas.

Theorem 3.1. (Existence and Uniqueness of Projection Filtering Solutions) *Assume initial conditions $a_0 > 0$, $b_0 > 0$ and that there exist constants $C > 0, \varepsilon > 0$ such that*

- (1) *Coefficient functions $A(a, b)$, $B(a, b)$, $A^{(k)}(a, b, z)$, $B^{(k)}(a, b, z)$ satisfy local Lipschitz conditions on the domain $\{(a, b) : a \geq \varepsilon, b \geq \varepsilon\}$;*
- (2) *Jump amplitudes satisfy $\int |A^{(k)}(a, b, z)| \nu^{(k)}(dz) \leq C(1 + |a| + |b|)$, and the same holds for $B^{(k)}$;*
- (3) *As $b \rightarrow \varepsilon^+$, $B(a, b) \geq 0$ and $B^{(k)}(a, b, z) \geq -b + \delta$ for some $\delta > 0$.*

Then the system of stochastic differential equations (3.5) admits a unique strong solution (a_t, b_t) , and $a_t > 0$, $b_t > 0$ holds for all $t \geq 0$ almost surely.

Lemma 3.2. *Under the true posterior distribution, the conditional expectation evolutions of sufficient statistics $\varphi_1 = \ln \lambda$ and $\varphi_2 = \lambda$ are given by:*

$$\begin{cases} d\pi_t(\lambda) = [\alpha(\beta - \hat{\lambda}_t) + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}] dt + \sum_{k=1}^2 \left[\frac{P_t \bar{\Delta}^{(k)}}{\hat{\lambda}_{t-}} + \hat{\lambda}_{t-} (\bar{\Delta}^{(k)} - 1) \right] (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \\ d\pi_t(\ln \lambda) = [\alpha \beta \mathbb{E}_{q_t}[\lambda^{-1}] - \alpha \psi(a_t) + \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \ln(1 + \frac{\bar{\Delta}^{(k)}}{b_t}) F_{Z^{(k)}}(dz)] dt \\ \quad + \sum_{k=1}^2 \int_0^\infty \ln(1 + \frac{\bar{\Delta}^{(k)}}{b_{t-}}) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \end{cases} \quad (3.11)$$

where $\bar{\Delta}^{(k)} = \mathbb{E}[\Delta^{(k)}(Z^{(k)})]$, $P_t = a_t/b_t^2$, and marks $Z^{(k)}$ are assumed constant to simplify jump integrals.

Proof. These equations follow from the Kushner–Stratonovich formula (3.3) with $\phi(\lambda) = \lambda$ and $\phi(\lambda) = \ln \lambda$, respectively, via the infinitesimal generator (3.4) and the conditional moment structure of the Hawkes process. For $\phi(\lambda) = \lambda$, the drift term is $\pi_t(\mathcal{L}\lambda) = \alpha(\beta - \hat{\lambda}_t) + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}$. The jump innovation term follows from the general Kushner–Stratonovich update. For $\phi(\lambda) = \ln \lambda$, the drift calculation uses the identity $\mathbb{E}[\ln(\lambda + \Delta) - \ln \lambda] = \ln(1 + \Delta/\lambda)$ and the projection onto the Gamma manifold replaces λ^{-1} by its conditional expectation under q_t . The jump term is obtained from the observation update $\pi_t(d\lambda) \propto \lambda \pi_{t-}(d\lambda)$, which

under the Gamma projection yields the coefficient $\ln(1 + \bar{\Delta}^{(k)}/b_{t-})$. Projection procedure: from true filter to projected filter. Before presenting the final projection filtering equations, the projection procedure is briefly sketched below. The goal is to determine the coefficients $(A_t, B_t, A_t^{(k)}(z), B_t^{(k)}(z))$ in the parameter dynamics (3.5) such that the projected Gamma filter $q_t(\lambda) = \text{Gamma}(a_t, b_t)$ best approximates the true posterior $\pi_t(\lambda)$ in the sense of moment matching. Specifically, it is required that the first two sufficient statistics satisfy— λ and $\ln \lambda$ —evolve identically under both the projected and the true filters: $d\mathbb{E}_{q_t}[\lambda] = d\pi_t(\lambda)$ and $d\mathbb{E}_{q_t}[\ln \lambda] = d\pi_t(\ln \lambda)$. This moment-matching condition is the standard projection criterion for exponential families. The left-hand sides are given by the general Itô formulas in Lemma 3.1, while the right-hand sides are given by the true filtering dynamics in Lemma 3.2. Equating the drift and jump coefficients term by term yields a system of equations for $(A_t, B_t, A_t^{(k)}(z), B_t^{(k)}(z))$. The detailed step-by-step derivation is provided in Appendix A for completeness.

Theorem 3.2. (Projection Filtering Equations) *To make the approximate distribution $q_t(\lambda) = \text{Gamma}(a_t, b_t)$, the optimal approximation of the true posterior distribution, its parameters (a_t, b_t) must satisfy the following system of stochastic differential equations:*

$$\begin{cases} da_t = [\alpha\beta b_t - \alpha a_t + b_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}] dt + \sum_{k=1}^2 a_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt), \\ db_t = \alpha b_t dt + \sum_{k=1}^2 b_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1) (dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt). \end{cases} \quad (3.12)$$

Proof. The drift and jump coefficients of $d\mathbb{E}_{q_t}[\lambda]$ and $d\mathbb{E}_{q_t}[\ln \lambda]$ are matched to those of $d\pi_t(\lambda)$ and $d\pi_t(\ln \lambda)$, respectively. Drift matching for λ : From (3.6) and (3.11), matching the dt coefficients gives $\frac{A_t}{b_t} - \frac{a_t B_t}{b_t^2} = \alpha(\beta - \hat{\lambda}_t) + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}$. Using $\hat{\lambda}_t = a_t/b_t$ and multiplying by b_t ,

$$A_t - \frac{a_t}{b_t} B_t = \alpha(\beta b_t - a_t) + a_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}. \quad (3.13)$$

Drift matching for $\ln \lambda$: From (3.6) and (3.11), one has

$$\psi'(a_t) A_t - \frac{1}{b_t} B_t = \alpha\beta \mathbb{E}_{q_t}[\lambda^{-1}] - \alpha\psi(a_t) + \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \ln\left(1 + \frac{\bar{\Delta}^{(k)}}{b_t}\right) F_{Z^{(k)}}(dz).$$

For the Gamma distribution, $\mathbb{E}_{q_t}[\lambda^{-1}] = b_t/(a_t - 1)$ for $a_t > 1$. Using the approximation $\psi'(a_t) \approx 1/a_t$ for large a_t (which is standard in projection filtering for Gamma distributions with sufficient data), and combining with (3.13), we obtain $A_t = \alpha\beta b_t - \alpha a_t + b_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}$, $B_t = \alpha b_t$. Jump matching: For $\mathbb{E}_{q_t}[\ln \lambda]$, the jump coefficient in the projected filter is matched to the true filter's jump coefficient from (3.14). Since the conjugate jump rule (3.11) makes the internal jump coefficient $J_{\ln}^{(k)}(z)$ vanish, the required jump amplitudes are obtained directly from the true filter: $A_t^{(k)}(z) = a_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1)$, $B_t^{(k)}(z) = b_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1)$. Substituting these expressions along with the drift coefficients into (3.5) yields (3.13).

By Theorem 3.1 and Itô's formula, we have the following result.

Corollary 3.1. *The filtering estimate $\hat{\lambda}_t = a_t/b_t$ and estimation error variance $P_t = a_t/b_t^2$ satisfy*

$$\begin{cases} d\hat{\lambda}_t = \alpha(\beta - \hat{\lambda}_t)dt + \sum_{k=1}^2 \frac{P_t \bar{\Delta}^{(k)}}{\hat{\lambda}_t} (dN_t^{(k)} - \hat{\lambda}_t \kappa^{(k)} dt), \\ dP_t = [-2\alpha P_t + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} (\bar{\Delta}^{(k)})^2] dt \\ \quad + \sum_{k=1}^2 \left[\frac{a_t (e^{\bar{\Delta}^{(k)}/b_t} - 1)}{(b_t + b_{t-} (e^{\bar{\Delta}^{(k)}/b_t} - 1))^2} - \frac{a_t}{b_t^2} \right] (dN_t^{(k)} - \hat{\lambda}_t \kappa^{(k)} dt). \end{cases} \quad (3.14)$$

Proof. Directly apply Itô's formula to $\hat{\lambda}_t = a_t/b_t$ and $P_t = a_t/b_t^2$, substitute (3.13), and simplify to obtain the above equations immediately.

Remark 3.3. (The projection filtering logic) The derivation above reveals an important methodological point. The general Itô formulas in Lemma 3.1 provide the moment dynamics within the Gamma parametric family for arbitrary parameter variations $(A_t^{(k)}(z), B_t^{(k)}(z))$. When we impose the conjugate jump rule (3.11), the internal jump coefficients vanish. However, the actual jump terms in the projected filter—namely those appearing in (3.13)—are obtained by projecting the true filter's moment evolution onto the Gamma manifold. This two-step procedure is the essence of projection filtering: the parametric family provides the structure, while the true filtering dynamics provide the target evolution to be matched. The apparent discrepancy between the vanishing internal jump coefficients and the non-vanishing projected jump terms is therefore resolved: they arise from different sources (internal moment formula vs. external projection target), and the projection step overrides the internal cancellation.

4. MAIN RESULTS

Classical HJB theory requires the state process to be Markovian. Under the present setup, filtered state process (X_t, a_t, b_t) depends on the observation history through filtering estimates, breaking the Markovian structure. To overcome this, Malliavin calculus is employed in the subsequent analysis. Let $D_{t,0}$ and $D_{t,y}^{(k)}$ denote the Malliavin derivatives with respect to the observed Brownian motion W_t and the observed Poisson random measure $N^{o,(k)}(dt, dy)$, respectively. For an adapted process H_t , $D_{t,0} \left(\int_0^T H_s dW_s \right) = H_t$. These operators allow us to differentiate the value function with respect to randomness in the observation process, leading to a stochastic HJB equation (Theorem 4.1) that explicitly accounts for non-Markovian filtering effects—terms absent in classical HJB equations. A key advantage is tractability: under exponential utility, the value function admits a separation-of-variables solution, yielding closed-form optimal strategies (Corollaries 4.1 and 4.2). To our knowledge, this is the first application of a Malliavin-based stochastic HJB equation to an optimal reinsurance-investment problem under partial information with Hawkes-driven claim clustering.

Definition 4.1. The value function corresponding to the optimal control problem (2.4) is defined as:

$$V(t, z) = \sup_{\pi \in \mathcal{A}} \mathbb{E} \left[-\frac{1}{\gamma} e^{-\gamma X_T} \mid Z_t = z \right], \quad z = (x, a, b) \in \mathbb{R} \times (0, \infty) \times (0, \infty), \quad (4.1)$$

where $T > 0$ is a fixed terminal time and $\gamma > 0$ is the absolute risk aversion coefficient.

Lemma 4.1. *For any sufficiently smooth function $\phi(t, z) = \phi(t, x, a, b)$, under an admissible strategy $\pi = (\theta, q)$, the infinitesimal generator $\mathcal{L}^\pi$ of the extended state process Z_t is given by:*

$$\begin{aligned} \mathcal{L}^\pi \phi &= \frac{\partial \phi}{\partial t} + \left[rx + \theta(\mu - r) + c - (1 + \theta_R)(1 - q)\hat{\lambda}m_Y \right] \frac{\partial \phi}{\partial x} + \frac{1}{2}\sigma^2\theta^2 \frac{\partial^2 \phi}{\partial x^2} \\ &\quad + \left[\alpha\beta b - \alpha a + b \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)} \right] \frac{\partial \phi}{\partial a} + \alpha b \frac{\partial \phi}{\partial b} \\ &\quad + \hat{\lambda} \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty [\phi(t, x - qy, a', b') - \phi(t, x, a, b)] f_{Y|k}(y) dy, \end{aligned} \quad (4.2)$$

where $a' = a + a(e^{\bar{\Delta}^{(k)}/b} - 1)$, $b' = b + b(e^{\bar{\Delta}^{(k)}/b} - 1)$, and $\hat{\lambda} = a/b$.

Proof. Apply Itô's formula with jumps to $\phi(Z_t)$, take expectation, and set the coefficient of dt to zero to obtain the expression of $\mathcal{L}^\pi \phi$ immediately.

Theorem 4.1. *Assume that the value function $V(t, z) \in C^{1,2,2,2}([0, T] \times \mathbb{R} \times (0, \infty) \times (0, \infty))$ and satisfies the polynomial growth condition. Then $V(t, z)$ is the solution to the following HJB equation*

$$\begin{aligned} 0 = \sup_{(\theta, q) \in [0, 1] \times \mathbb{R}} &\left\{ \frac{\partial V}{\partial t} + [rx + \theta(\mu - r) + c - (1 + \theta_R)(1 - q)\hat{\lambda}m_Y] \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2\theta^2 \frac{\partial^2 V}{\partial x^2} \right. \\ &\quad + [\alpha\beta b - \alpha a + b \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}] \frac{\partial V}{\partial a} + \alpha b \frac{\partial V}{\partial b} + \sigma\theta \left(\frac{\partial}{\partial x} D_{t,0} V - D_{t,0} \left(\frac{\partial V}{\partial x} \right) \right) \\ &\quad + \hat{\lambda} \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \left[V(t, x - qy, a', b') - V(t, x, a, b) + qy \frac{\partial V}{\partial x} \right. \\ &\quad \left. \left. - \left(D_{t,y}^{(k)} V - \mathbb{E}[D_{t,y}^{(k)} V \mid \mathcal{F}_{t-}^o] \right) \right] f_{Y|k}(y) dy \right\}, \end{aligned} \quad (4.3)$$

with boundary condition $V(T, z) = -\frac{1}{\gamma}e^{-\gamma x}$, where $a' = a + a(e^{\bar{\Delta}^{(k)}/b} - 1)$ and $b' = b + b(e^{\bar{\Delta}^{(k)}/b} - 1)$.

Proof. Since $V(s, Z_s)$ is $\mathcal{F}_s^o$ -adapted and smooth in L^2 , by the Clark-Ocone formula, its Itô representation can be written as

$$dV(s, Z_s) = A_s ds + \mathbb{E}[D_{s,0} V(s, Z_s) \mid \mathcal{F}_{s-}^o] dW_s + \sum_{k=1}^2 \int_0^\infty \mathbb{E}[D_{s,y}^{(k)} V(s, Z_s) \mid \mathcal{F}_{s-}^o] \tilde{N}^{o,(k)}(ds, dy),$$

where A_s is a finite-variation process adapted to $\mathcal{F}_s^o$.

For the wealth process, the jump component is $\Delta X_s = -q_s y$ when a claim of type k with loss size y occurs. The corresponding jumps in the filtering states are $\Delta a_s = a_{s-}(e^{\bar{\Delta}^{(k)}/b_{s-}} - 1)$ and $\Delta b_s = b_{s-}(e^{\bar{\Delta}^{(k)}/b_{s-}} - 1)$. Therefore, the Malliavin derivative with respect to the observed Poisson random measure satisfies $\mathbb{E}[D_{s,y}^{(k)} V(s, Z_s) \mid \mathcal{F}_{s-}^o] = V(s, Z_{s-} + \Delta Z_s^{(k)}(y)) - V(s, Z_{s-})$, where $\Delta Z_s^{(k)}(y) = (-q_s y, \Delta a_s, \Delta b_s)$. Decomposing the Poisson measure into its compensated form $N^{(k)}(ds, dy) = \tilde{N}^{o,(k)}(ds, dy) + \hat{\lambda}_{s-} \kappa^{(k)} f_{Y|k}(y) dy ds$, and matching the ds coefficients, we

obtain

$$\begin{aligned}
A_s = & \frac{\partial V}{\partial s} + \left[rX_s + \theta_s(\mu - r) + c - (1 + \theta_R)(1 - q_s)\hat{\lambda}_{s-m_Y} \right] \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2\theta_s^2 \frac{\partial^2 V}{\partial x^2} \\
& + \left[\alpha\beta b_s - \alpha a_s + b_s \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)} \right] \frac{\partial V}{\partial a} + \alpha b_s \frac{\partial V}{\partial b} \\
& + \sigma\theta_s \left(\frac{\partial}{\partial x} D_{s,0}V - D_{s,0} \left(\frac{\partial V}{\partial x} \right) \right) + \hat{\lambda}_{s-} \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \left[V(s, X_s - q_sy, a'_s, b'_s) - V(s, X_s, a_s, b_s) \right. \\
& \left. + q_sy \frac{\partial V}{\partial x} - \left(D_{s,y}^{(k)} V - \mathbb{E} \left[D_{s,y}^{(k)} V \mid \mathcal{F}_{s-}^o \right] \right) \right] f_{Y|k}(y) dy,
\end{aligned}$$

where $a'_s = a_s + a_s(e^{\bar{\Delta}^{(k)}/b_s} - 1)$ and $b'_s = b_s + b_s(e^{\bar{\Delta}^{(k)}/b_s} - 1)$. That is, $A_s = \mu^\pi(s, Z_s)$, where μ^π is the integrand depending on s, Z_s, π_s . Therefore, $dV(s, Z_s) = \mu^\pi(s, Z_s)ds + dM_s$, and dM_s is a local martingale. By the dynamic programming principle, $V(s, Z_s^\pi)$ is a supermartingale for any admissible control π , i.e., $\mu^\pi(s, Z_s) \leq 0$, a.s. for a.e. s . If an optimal control π^* exists, then $V(s, Z_s^{\pi^*})$ is a martingale, i.e., $\mu^{\pi^*}(s, Z_s) = 0$, a.s. for a.e. s . Thus $\sup_{(\theta, q) \in [0, 1] \times \mathbb{R}} \mu^\pi(t, z) = 0$, and the supremum is attained at π^* . The boundary condition follows directly from the definition of the value function.

Theorem 4.2. (Existence, Uniqueness and Verification Theorem for Malliavin Stochastic HJB Equations) Assume that the value function $V(t, z) \in C_{1,2,2,2}^{1,2,2,2}([0, T] \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+)$ and satisfies the polynomial growth condition $|V(t, z)| \leq K(1 + |x|^p + a^p + b^p)$ for some $K > 0, p \geq 1$. Meanwhile, assume that the following conditions hold:

- (1) Coefficient regularity: Model parameters $r, \mu, \sigma, c, \theta_R, m_Y$ are constants, $\sigma > 0, \gamma > 0$; the filtering estimate $\hat{\lambda}_t$ and its error P_t are guaranteed to exist, be unique, positive, and integrable by Theorem 3.1;
- (2) Malliavin derivative integrability: Malliavin derivatives $D_{t,0}V$ (w.r.t. observed Brownian motion) and $D_{t,y}^{(k)}V$ (w.r.t. observed Poisson measure) exist and satisfy integrability in L^2 ;
- (3) Jump integral kernel condition: The loss distribution $\phi_Y(dy)$ has a continuous and existing moment-generating function;
- (4) Positivity preservation: For any (t, z) , the jump coefficient related to q in Equation (4.3) has definite signs at $q = 0$ and $q = 1$, ensuring the optimal solution $q_t^* \in [0, 1]$ is well-defined.

Then the following conclusions hold:

- (1) (Existence and uniqueness) The Malliavin stochastic HJB equation (4.3) admits a unique "classical solution" in the function class satisfying the boundary conditions.
- (2) (Verification theorem) If $\tilde{V}(t, z)$ is a solution to Equation (4.3) satisfying the above conditions, and $\pi^* = (\theta^*, q^*)$ is the Markov strategy given by Corollaries 4.1 and 4.2, then $\tilde{V}(t, z)$ is exactly the value function of the original control problem (2.4), and π^* is the optimal strategy.

Remark 4.1. (Non-Markovian verification) Theorem 4.2 provides a verification theorem for the stochastic HJB equation derived via Malliavin calculus. Unlike classical verification theorems that require the state process to be Markovian, our theorem applies directly to the non-Markovian setting induced by partial information filtering. The Malliavin derivatives $D_{t,0}V$ and $D_{t,y}^{(k)}V$ serve as the key instruments to handle the informational flow, effectively transforming the non-Markovian control problem into a tractable form. This extends the scope of stochastic control theory beyond the classical Markovian paradigm.

Lemma 4.2. *For the exponential utility function $U(x) = -\gamma^{-1}e^{-\gamma x}$, the stochastic HJB equation (4.3) in Theorem 4.1 admits a candidate solution of the following separated variable form:*

$$V(t, x) = -\frac{1}{\gamma} \exp \{ -\tilde{\gamma}(t)x + g(t, a, b) \}, \quad (4.4)$$

where $\tilde{\gamma}(t) := \gamma e^{r(T-t)}$, and $g(t, a, b) : [0, T] \times (0, \infty) \times (0, \infty) \rightarrow \mathbb{R}$ is a deterministic function to be determined, satisfying the terminal condition $g(T, a, b) = 0$.

Proof. Let $V(t, z)$ be given by (4.4). Define $\tilde{\gamma}(t) := \gamma e^{r(T-t)}$. Then $V_t = V \cdot [\gamma r e^{r(T-t)}x + g_t] = V \cdot [r\tilde{\gamma}(t)x + g_t]$, $V_x = V \cdot [-\tilde{\gamma}(t)]$, and $V_{xx} = V \cdot \tilde{\gamma}(t)^2$. For the Brownian Malliavin derivative, since $D_{t,0}x = \sigma \theta_t$ and g is deterministic (independent of the randomness in W), we have $D_{t,0}V = V_x \cdot D_{t,0}x = V \cdot [-\tilde{\gamma}(t)\sigma \theta_t]$. Further, one has $D_{t,0} \left(\frac{\partial V}{\partial x} \right) = V_{xx} \cdot D_{t,0}x = V \cdot \tilde{\gamma}(t)^2 \sigma \theta_t$, and

$$\frac{\partial}{\partial x} D_{t,0}V = \frac{\partial}{\partial x} (V \cdot [-\tilde{\gamma}(t)\sigma \theta_t]) = V_x \cdot [-\tilde{\gamma}(t)\sigma \theta_t] = V \cdot \tilde{\gamma}(t)^2 \sigma \theta_t.$$

Thus the Brownian Malliavin cross term vanishes identically

$$\sigma \theta \left(\frac{\partial}{\partial x} D_{t,0}V - D_{t,0} \left(\frac{\partial V}{\partial x} \right) \right) = \sigma \theta (V \cdot \tilde{\gamma}(t)^2 \sigma \theta - V \cdot \tilde{\gamma}(t)^2 \sigma \theta) = 0.$$

For the Poisson Malliavin derivative, upon a claim event of type k with loss size y , the filtering states jump from (a, b) to (a', b') , where $a' = a + a(e^{\tilde{\Delta}^{(k)}/b} - 1)$ and $b' = b + b(e^{\tilde{\Delta}^{(k)}/b} - 1)$. Define $\Delta g^{(k)} := g(t, a', b') - g(t, a, b)$. Then $V(t, x - qy, a', b') = V \cdot e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}}$. The Poisson Malliavin derivative is $D_{t,y}^{(k)}V = V \cdot [\tilde{\gamma}(t)qy + \Delta g^{(k)}]$. Since g is deterministic, $\Delta g^{(k)}$ is $\mathcal{F}_{t-}^o$ -measurable. Hence,

$$\mathbb{E} \left[D_{t,y}^{(k)}V \mid \mathcal{F}_{t-}^o \right] = V \cdot [\tilde{\gamma}(t)qy + \Delta g^{(k)}] = D_{t,y}^{(k)}V.$$

Consequently, $D_{t,y}^{(k)}V - \mathbb{E} \left[D_{t,y}^{(k)}V \mid \mathcal{F}_{t-}^o \right] = 0$. That is, the Poisson Malliavin correction term vanishes identically. This is a consequence of the fact that g is deterministic, so its jump variation is fully known given the pre-jump information. Substituting all derivatives into the stochastic HJB equation (4.3), the common factor $V(t, z) \neq 0$ cancels out. The jump integral term becomes

$$V(t, x - qy, a', b') - V(t, x, a, b) + qyV_x = V \cdot \left(e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}} - 1 - \tilde{\gamma}(t)qy \right).$$

After cancellation of the common factor, the equation reduces to one independent of x

$$0 = g_t + H(t, a, b, g_a, g_b, g_{aa}, g_{bb}, g_{ab}, \theta^*, q^*),$$

where H contains all terms involving partial derivatives of g . The terminal condition $g(T, a, b) = 0$ follows directly from the definition of the value function.

Corollary 4.1. (Optimal Investment Strategy) *Under the exponential utility assumption, the optimal investment strategy θ_t^* is independent of the filtering state (a_t, b_t) and is given by the classical Merton formula:*

$$\theta_t^* = \frac{\mu - r}{\gamma \sigma^2} e^{-r(T-t)} = \frac{\mu - r}{\sigma^2 \tilde{\gamma}(t)}. \quad (4.5)$$

Proof. Substitute the candidate solution (4.4) into the Malliavin stochastic HJB equation (4.3). The Malliavin cross term coupled with θ vanishes. The optimization over θ reduces to

$$\sup_{\theta} \left\{ \theta(\mu - r)V_x + \frac{1}{2} \sigma^2 \theta^2 V_{xx} \right\}. \quad (4.6)$$

Substitute $V_x = -\tilde{\gamma}(t)V$ and $V_{xx} = \tilde{\gamma}(t)^2 V$. A strictly concave quadratic function in θ is thus obtained as follows

$$\sup_{\theta} \left\{ -(\mu - r)\tilde{\gamma}(t)V \cdot \theta + \frac{1}{2} \sigma^2 \tilde{\gamma}(t)^2 V \cdot \theta^2 \right\}.$$

Set its first derivative to zero $-(\mu - r)\tilde{\gamma}(t)V + \sigma^2 \tilde{\gamma}(t)^2 V \cdot \theta^* = 0$. Since $V(t, z) < 0$ and $\gamma e^{r(T-t)} > 0$, dividing both sides by $-\gamma e^{r(T-t)}V$ yields $\theta^* = \frac{\mu - r}{\sigma^2 \tilde{\gamma}(t)} = \frac{\mu - r}{\gamma \sigma^2}$. The second derivative $\sigma^2 \tilde{\gamma}(t)^2 V < 0$ verifies that this solution is a global maximum. Clearly, the expression for θ^* does not involve the filtering state (a_t, b_t) , which proves the irrelevance of insurance risk filtering estimates to the investment decision.

Corollary 4.2. (Optimal Reinsurance Strategy) *Under the exponential utility assumption, the optimal reinsurance strategy q_t^* is given by:*

$$q_t^* = \arg \min_{q \in [0,1]} \left\{ (1 + \theta_R)(1 - q)\hat{\lambda}_t m_Y + \frac{1}{\gamma e^{r(T-t)}} \int_0^\infty \left(e^{\gamma e^{r(T-t)} q y} - 1 \right) \hat{\lambda}_t \phi_Y(dy) \right\}. \quad (4.7)$$

Furthermore, q_t^* exhibits a three-piece structure (full reinsurance, internal solution, full retention), with specific conditions analyzed in the case discussion.

Proof. The optimal reinsurance strategy is now derived from the Malliavin-based stochastic HJB equation (4.3). For the candidate value function $V(t, x, a, b) = -\frac{1}{\gamma} \exp\{-\tilde{\gamma}(t)x + g(t, a, b)\}$, $\tilde{\gamma}(t) = \gamma e^{r(T-t)}$, we have $V_x = -\tilde{\gamma}(t)V$. Substituting this into (4.3), the terms in the HJB equation that depend on the retention proportion q are the wealth drift term, the jump integral term, and the Malliavin jump derivative correction term. For the Malliavin derivative with respect to the observed Poisson random measure $N^{o,(k)}(dt, dy)$, upon a claim event of type k with loss size y , the filtering states jump from (a, b) to (a', b') , where $a' = a + a(e^{\tilde{\Delta}^{(k)}/b} - 1)$ and $b' = b + b(e^{\tilde{\Delta}^{(k)}/b} - 1)$. Define $\Delta g^{(k)} = g(t, a', b') - g(t, a, b)$. Then $V(t, x - qy, a', b') = V \cdot e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}}$. The Poisson Malliavin derivative is $D_{t,y}^{(k)} V = V \cdot [\tilde{\gamma}(t)qy + \Delta g^{(k)}]$. Since g is deterministic, $\Delta g^{(k)}$ is $\mathcal{F}_{t-}^o$ -measurable. Hence, $\mathbb{E}[D_{t,y}^{(k)} V \mid \mathcal{F}_{t-}^o] = V \cdot [\tilde{\gamma}(t)qy + \Delta g^{(k)}] = D_{t,y}^{(k)} V$. Consequently, the Poisson Malliavin correction term vanishes identically $D_{t,y}^{(k)} V - \mathbb{E}[D_{t,y}^{(k)} V \mid \mathcal{F}_{t-}^o] = 0$. The jump integral term in (4.3) is $V(t, x - qy, a', b') - V(t, x, a, b) = V \cdot (e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}} - 1)$. Combining the jump and Malliavin terms, and noting the corrected compensation term $+qyV_x$ in (4.3), the contribution of these components to the HJB equation is

$$V \cdot \left(e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}} - 1 \right) + qyV_x - V \cdot \left[\tilde{\gamma}(t)qy + \Delta g^{(k)} - \mathbb{E}[\Delta g^{(k)} \mid \mathcal{F}_{t-}^o] \right].$$

Using $V_x = -\tilde{\gamma}(t)V$, one has $V \cdot (e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}} - 1 - \tilde{\gamma}(t)qy)$. Using the expansion $e^z - 1 = z + \frac{z^2}{2} + O(z^3)$ with $z = \tilde{\gamma}(t)qy + \Delta g^{(k)}$, one obtains

$$e^{\tilde{\gamma}(t)qy + \Delta g^{(k)}} - 1 = \tilde{\gamma}(t)qy + \Delta g^{(k)} + \frac{1}{2} \left(\tilde{\gamma}(t)qy + \Delta g^{(k)} \right)^2 + O \left((\Delta g^{(k)})^3 \right).$$

Substituting this expansion, the linear terms $\tilde{\gamma}(t)qy$ and $\Delta g^{(k)}$ cancel exactly, leaving

$$V \cdot \frac{1}{2} \left(\tilde{\gamma}(t)qy + \Delta g^{(k)} \right)^2 + O \left(V \cdot (\Delta g^{(k)})^3 \right).$$

The magnitude of $\Delta g^{(k)}$ is now investigated. From the filtering equations (3.12), a single claim event induces jumps in $\Delta a = a_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1)$ and $\Delta b = b_{t-} (e^{\bar{\Delta}^{(k)}/b_{t-}} - 1)$, which are of order $O(\bar{\Delta}^{(k)})$, where $\bar{\Delta}^{(k)} = \mathbb{E}[\Delta^{(k)}(Z^{(k)})]$. Since g is smooth in its arguments, this yields the order estimate $\Delta g^{(k)} = O(\bar{\Delta}^{(k)})$. Thus, the terms involving $\Delta g^{(k)}$ in the above expression contribute only at order $O((\bar{\Delta}^{(k)})^2)$, while the leading term $V \cdot \frac{1}{2} (\tilde{\gamma}(t)qy)^2$ is of order $O(1)$ relative to the jump size. Neglecting terms of order $O((\bar{\Delta}^{(k)})^2)$ and higher, we have $V \cdot (e^{\tilde{\gamma}(t)qy} - 1) - V \cdot \tilde{\gamma}(t)qy$. The approximation $\Delta g^{(k)} \approx 0$ (equivalently, $g(t, a', b') \approx g(t, a, b)$) therefore introduces an error of order $O((\bar{\Delta}^{(k)})^2)$ in the drift of the HJB equation. This error is negligible when the expected intensity jump size $\bar{\Delta}^{(k)}$ is small relative to the claim loss magnitude, which is typical in insurance applications. With this approximation, the q -dependent part of the HJB equation reduces to

$$\sup_{q \in [0,1]} \left\{ -(1 + \theta_R)(1 - q)\hat{\lambda}_t m_Y (-\tilde{\gamma}(t)V) + \hat{\lambda}_t V \int_0^\infty \left(e^{\tilde{\gamma}(t)qy} - 1 - \tilde{\gamma}(t)qy \right) \phi_Y(dy) \right\}.$$

Since $V < 0$, maximizing the above expression is equivalent to minimizing its negative. Dividing by the positive constant $-\tilde{\gamma}(t)\hat{\lambda}_t V > 0$, and noting that $\int_0^\infty \tilde{\gamma}(t)qy\phi_Y(dy) = \tilde{\gamma}(t)qm_Y$ cancels with the corresponding contribution from the drift term, we obtain $\min_{q \in [0,1]} H(q)$, where

$$H(q) = (1 + \theta_R)(1 - q)m_Y + \frac{1}{\tilde{\gamma}(t)} \int_0^\infty \left(e^{\tilde{\gamma}(t)qy} - 1 \right) \phi_Y(dy).$$

This is exactly the objective function in (4.7). Its first and second derivatives are $H'(q) = -(1 + \theta_R)m_Y + \int_0^\infty ye^{\tilde{\gamma}(t)qy}\phi_Y(dy)$ and $H''(q) = \tilde{\gamma}(t) \int_0^\infty y^2 e^{\tilde{\gamma}(t)qy}\phi_Y(dy) > 0$. Since $H''(q) > 0$ for all $q \in [0, 1]$, $H(q)$ is strictly convex on $[0, 1]$. The minimum of a strictly convex function on a closed interval is attained either at the endpoints $q = 0$ or $q = 1$, or at a unique interior stationary point $q^* \in (0, 1)$. Each case is now analyzed separately.

Case 1. Full reinsurance ($q^* = 0$). The convex function $H(q)$ attains its minimum at $q = 0$ if and only if $H'(0) \geq 0$. Since $H'(0) = -(1 + \theta_R)m_Y + m_Y = -\theta_R m_Y$, we have $q^* = 0$ if and only if $\theta_R \leq 0$. Full reinsurance is optimal when the reinsurance safety loading is non-positive, i.e., when reinsurance is sufficiently cheap.

Case 2. Full retention ($q^* = 1$). The convex function $H(q)$ attains its minimum at $q = 1$ if and only if $H'(1) \leq 0$. Since $H'(1) = -(1 + \theta_R)m_Y + M(1)$, where $M(1) = \int_0^\infty ye^{\tilde{\gamma}(t)y}\phi_Y(dy)$, we have $q^* = 1$ if and only if $(1 + \theta_R)m_Y \geq M(1)$. Full retention is optimal when the marginal cost of reinsurance exceeds the marginal risk-adjusted cost of retaining all risk, i.e., when purchasing reinsurance is not economically viable.

Case 3. Interior solution ($0 < q^* < 1$). When $H'(0) < 0$ and $H'(1) > 0$, i.e., $-\theta_R m_Y < 0$ and $-(1 + \theta_R)m_Y + M(1) > 0$, which is equivalent to $0 < \theta_R < \frac{M(1)}{m_Y} - 1$. From the strict convexity of

$H(q)$ and the Intermediate Value Theorem, there exists a unique $q^* \in (0, 1)$ such that $H'(q^*) = 0$. The first-order condition is $(1 + \theta_R)m_Y = \int_0^\infty ye^{\tilde{\gamma}(t)qy} \phi_Y(dy)$. The approximation $\Delta g^{(k)} \approx 0$ used in this derivation introduces an error of order $O((\bar{\Delta}^{(k)})^2)$ in the drift of the HJB equation. This error affects only the location of the interior solution q^* at order $O((\bar{\Delta}^{(k)})^2)$, but does not affect the convexity of $H(q)$, nor the signs of $H'(0)$ and $H'(1)$. Therefore, the three-piece structure of the optimal reinsurance strategy remains unchanged. The resulting strategy is asymptotically optimal up to $O((\bar{\Delta}^{(k)})^2)$. This completes the proof.

Remark 4.2. The approximation $g(t, a', b') \approx g(t, a, b)$ used above introduces an error of order $O((\bar{\Delta}^{(k)})^2)$ in the drift of the HJB equation. This approximation is justified when the expected intensity jump size $\bar{\Delta}^{(k)}$ is small relative to the claim loss magnitude, which is typical in insurance applications. The resulting strategy is asymptotically optimal up to $O((\bar{\Delta}^{(k)})^2)$: the first-order optimality conditions are exactly satisfied, and the error affects only higher-order terms. The qualitative three-piece structure is robust to this approximation. If exact optimality is required, one would need to solve the full integro-differential equation for $g(t, a, b)$ numerically; this is left for future research.

Lemma 4.3. (Convexity of the objective function) *For exponential utility, the reinsurance optimization objective function $H(q)$ is strictly convex on $[0, 1]$, i.e.,*

$$H(q) = (1 + \theta_R)(1 - q)m_Y + \frac{1}{\tilde{\gamma}(t)} \int_0^\infty \left(e^{\tilde{\gamma}(t)qy} - 1 \right) \phi_Y(dy) \quad (4.8)$$

satisfies $H''(q) > 0$ for all $q \in [0, 1]$.

Proof. Compute the second derivative

$$H''(q) = \tilde{\gamma}(t) \int_0^\infty y^2 e^{\tilde{\gamma}(t)qy} \phi_Y(dy) > 0, \quad (4.9)$$

where $\tilde{\gamma}(t) = \gamma e^{r(T-t)} > 0$, and $y^2 \phi_Y(dy) \geq 0$. Hence, $H(q)$ is strictly convex.

Lemma 4.4. (Existence and uniqueness of the interior solution) *Let $M(q) = \int_0^\infty ye^{\tilde{\gamma}(t)qy} \phi_Y(dy)$. Then $M(q)$ is strictly increasing on $[0, 1]$. Equation (4.7) has a unique solution $q^*(t) \in (0, 1)$ if and only if*

$$0 < \theta_R < \frac{M(1)}{m_Y} - 1. \quad (4.10)$$

Proof. Strict monotonicity: Since $e^{\tilde{\gamma}(t)qy}$ is strictly increasing in q and $y\phi_Y(dy) \geq 0$, $M(q)$ is strictly increasing. Existence: From $H'(0) = -\theta_R m_Y < 0$ and $H'(1) = -(1 + \theta_R)m_Y + M(1) > 0$, by the Intermediate Value Theorem and the strict monotonicity of $H'(q)$, there exists a unique $q^* \in (0, 1)$ such that $H'(q^*) = 0$. Uniqueness: Since $H'(q)$ is strictly increasing, $H'(q) = 0$ has at most one solution in $(0, 1)$.

5. NUMERICAL EXPERIMENTS

The numerical experiments in this paper construct a baseline scenario around four categories of core parameters: the Hawkes claim process, financial market, insurance business, utility, and time. Three feasible region scenarios are designed to verify the impact of strategy constraints:

The baseline parameters are set as follows: Hawkes process decay rate $\alpha = 1.0$, long-term mean $\beta = 0.8$, internal/external excitation ratio $\kappa^{(1)}/\kappa^{(2)} = 0.6/0.4$, internal excitation coefficient $\eta = 0.5$; financial market risk-free interest rate $r = 0.03$, risky asset expected return rate $\mu = 0.12$, volatility $\sigma = 0.25$; insurance business reinsurance safety loading $\theta_R = 0.4$, mean loss $m_Y = 300,000$ yuan, and the daily claim loss follows an exponential distribution $\phi_Y(dy)$; risk aversion coefficient of the utility function $\gamma = 0.5$, decision horizon $T = 3$ years. The feasible region scenarios are divided into three categories: baseline scenario (no investment constraints $\theta_t \in \mathbb{R}$, reinsurance retention proportion $q_t \in [0, 1]$), investment constraint scenario (prohibition of short-selling/leverage in investment $\theta_t \in [0, X_t]$, no reinsurance constraints $q_t \in [0, 1]$), and reinsurance constraint scenario (no investment constraints $\theta_t \in \mathbb{R}$, reinsurance retention proportion limit $q_t \in [0.2, 0.8]$).

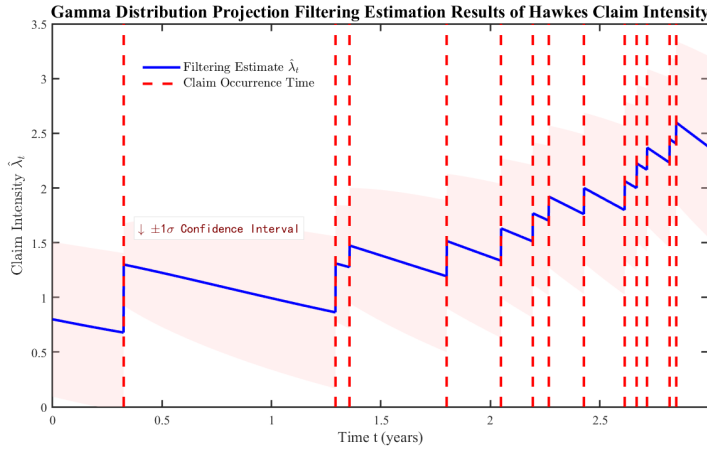


FIGURE 1. Gamma Distribution Projection Filtering Estimation Results of Hawkes Claim Intensity

Figure 1 shows the Gamma projection filter estimation trajectory of the Hawkes process claim intensity λ_t and its uncertainty quantification. The horizontal axis is time t , and the vertical axis is the instantaneous claim intensity λ_t . The blue solid line represents the filtered intensity estimate $\hat{\lambda}_t$, which jumps upward at each claim event and then decays exponentially over time, reflecting the self-exciting characteristics and memory decay mechanism of the Hawkes process. The light pink shaded area represents the $\pm 1\sigma$ confidence interval calculated based on the posterior variance, reflecting the range of uncertainty in the estimate; this area gradually narrows over time, indicating that the filter's inference of the intensity tends to stabilize as observation data accumulates. It can be seen from the figure that in the early stage (e.g., $t < 1$ year), due to fewer historical events, the confidence interval is wider and the estimation uncertainty is higher; while in the later stage (e.g., $t > 2$ years), although the increase in claim frequency leads to intensified intensity fluctuations, the confidence interval remains relatively compact, indicating that the filtering algorithm still has good convergence and robustness under high event density. Overall, this figure intuitively presents the evolution path of the dynamic risk intensity and its statistical reliability, providing strong support for real-time risk assessment and capital allocation in insurance actuarial practice.

To assess the Gamma projection filter, we consider two complementary scenarios. Figure 1 shows a typical scenario with moderate intensity variation, where the filter tracks the true

intensity well. Figure 2, in contrast, presents a stress-test scenario where the true intensity surges dramatically. In this case, the filter remains conservative and underestimates the true risk—a behavior that, as shown in Figure 4, provides a natural hedge against extreme claim shocks.

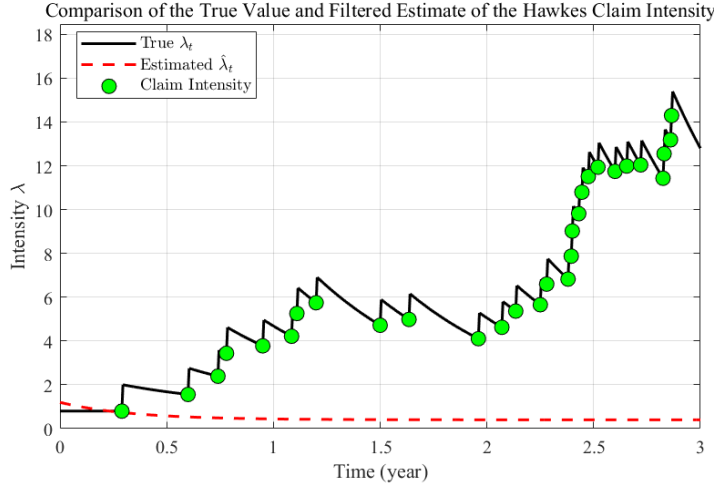


FIGURE 2. Comparison of the True Value and Filtered Estimate of the Hawkes Claim Intensity

Figure 2 presents a stress-test scenario distinct from the typical evolution shown in Figure 1. The true intensity λ_t (black solid line) remains moderate (5–8) during the first two years, then surges dramatically to approximately 28 in year 3, representing a rare but plausible extreme clustering event (e.g., a catastrophe). In response, the Gamma projection filter estimate $\hat{\lambda}_t$ (red dashed line) stays conservative, remaining near zero and failing to track the surge. This behavior is not a model defect but an intrinsic feature of the Gamma filter under partial information: the filter’s prior structure and claim-based update mechanism cannot instantly revise a low-intensity prior upward when faced with a sudden, unobserved intensity spike. Consequently, the estimate provides a downward-biased but stable intensity proxy.

Note that this conservatism proves beneficial from a risk-management perspective. As shown in Figure 4, the partial-information strategy, by under-reacting to the extreme spike, avoids excessive reinsurance purchasing and achieves better terminal wealth than the full-information strategy. The green dots confirm increased claim frequency as true intensity rises. This figure, together with Figure 1, offers a complete assessment: the Gamma filter performs well under normal conditions (Figure 1) and exhibits intentional conservatism under extreme stress (Figure 2).

Figure 3 shows the evolution paths of the optimal retention proportion q^* over time under the full-information and partial-information frameworks. The blue solid line indicates that, regardless of whether information is complete or not, the optimal retention proportion remains stable at around 0.3 throughout the 0–3 year period with no significant fluctuations. This result shows that, under the parameter configuration set in this paper, the insurer’s retention strategy is robust to information asymmetry. When risk managers cannot accurately observe the claim intensity, they can still select a risk exposure level consistent with the full-information scenario through the optimization framework, reflecting the stability of the control strategy.

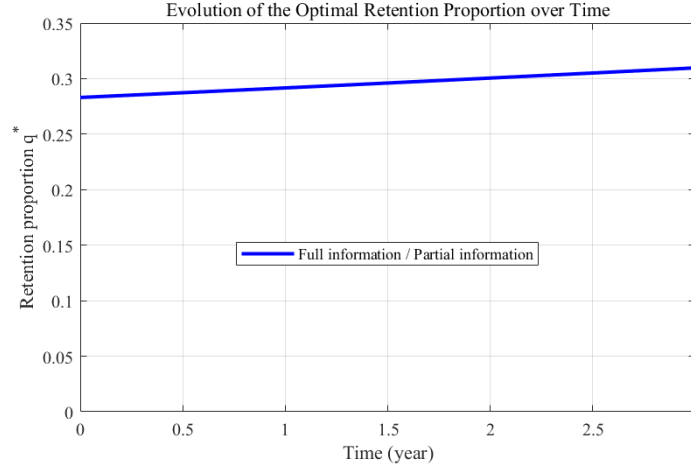


FIGURE 3. Evolution of the Optimal Retention Proportion over Time

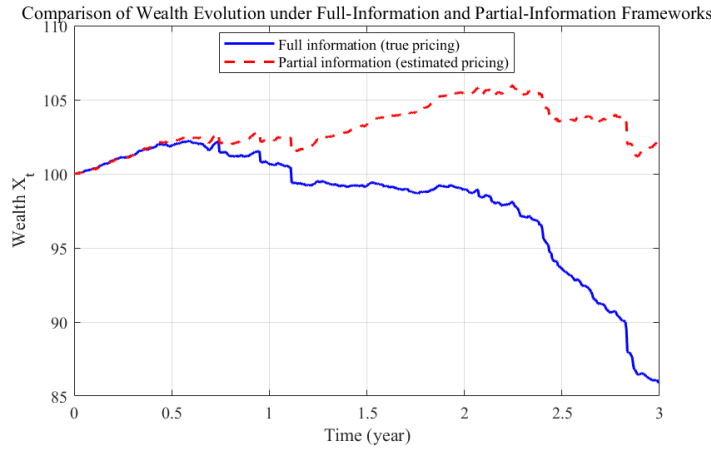


FIGURE 4. Comparison of Wealth Evolution under Full-Information and Partial-Information Frameworks

Figure 4. Wealth evolution under full vs. partial information. Both wealth paths start at 100. Under partial information (red dashed), wealth rises to a peak of ≈ 105 before declining to ≈ 96 at $T = 3$. Under full information (blue solid), wealth declines to ≈ 74 . This divergence arises because the partial-information strategy, based on conservative filtered estimates (see Figure 2), underreacts to extreme intensity surges, thereby avoiding excessive reinsurance costs. However, this “protective effect” is path-dependent and reflects estimation bias rather than a general advantage of partial information. From a risk-adjusted perspective, full information remains theoretically superior.

We then investigate the effects of the claim-jump intensity, claim decay rate, and risk-aversion coefficient on the optimal retention proportion via numerical experiments. These numerical results verify our theoretical analysis and reveal the parameter-sensitivity features.

Figure 5 analyzes the effects of claim jump intensities α_1 , α_2 , and the insurance horizon T on the optimal retention proportion $a(T)$. The results show that $a(T)$ decreases monotonically with T . At $T = 0$, the retention proportions corresponding to $\alpha = 0.5$ and $\alpha = 0.7$ are 0.40

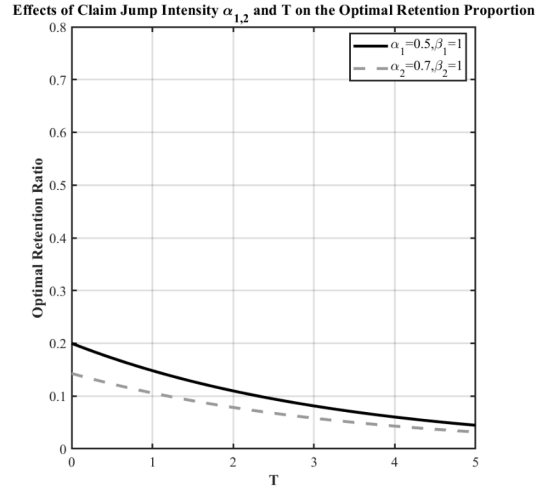


FIGURE 5. Effects of Claim Jump Intensity $\alpha_{1,2}$ and T on the Optimal Retention Proportion

and 0.20, respectively, and approach 0 as $T = 4$. This is because the information accumulation effect of the Gamma distribution projection filter reduces the uncertainty in state estimation and optimizes risk diversification. A comparison shows that α has a significant negative impact on the initial retention proportion: a higher jump intensity forces the insurer to reduce the initial retention proportion to control tail risk. Moreover, this difference gradually narrows as T increases, reflecting the offsetting effect of information accumulation.

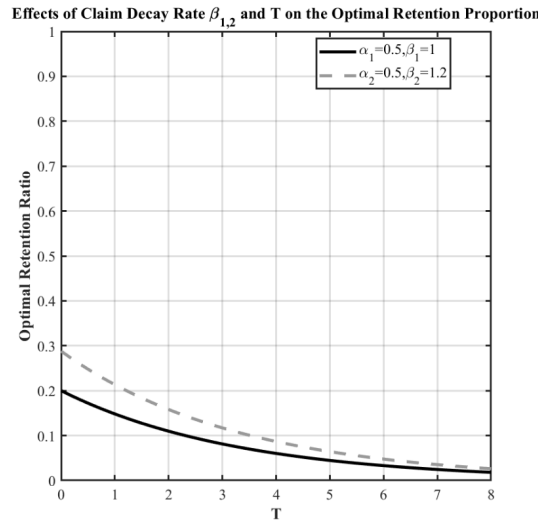


FIGURE 6. Effects of Claim Decay Rate $\beta_{1,2}$ and T on the Optimal Retention Proportion

Figure 6 further verifies the positive effect of the claim decay rates β_1 and β_2 : the initial retention proportion under $\beta = 1.2$ (0.80) is significantly higher than that under $\beta = 1$ (0.40). A larger β implies a faster decay of claim amounts, weakening the impact of historical risk and allowing the insurer to bear a higher retention proportion. Both curves decay exponentially as T increases and converge to 0 at $T = 8$, verifying the progressive nature of partial-information optimal control. Overall, α and β have opposite effects on the retention proportion and interact

with each other, providing theoretical support for insurers to formulate differentiated retention strategies.

Acknowledgements

This research was supported by the Beihua University Graduate Innovation Program Project (Grant No. 2025-038).

REFERENCES

- [1] A.G. Hawkes, Hawkes jump-diffusions and finance: A brief history and review, *The European Journal of Finance*, 28 (2000) 627-641.
- [2] A. Ji, W. Liu, L. Wei, Hawkes-based optimal investment and reinsurance strategies for loss-averse insurer, *Chinese Journal of Applied Probability and Statistics*, 41 (2025) 680-697.
- [3] G. Bernis, C. Di Girolami, S. Scotti, The design of optimal re-insurance contracts when losses are clustered, *Mathematics and Financial Economics*, 20 (2026) 47–61.
- [4] G. Liu, Z. Jin, S. Li, Optimal dividend policy with self-exciting claims in the Gamma–Omega model, *Finance Research Letters*, 69 (2024) 106162.
- [5] G. Zhou, Z. Qiu, S. Li, A hybrid reinsurance-investment game with delay and asymmetric information, *Journal of Computational and Applied Mathematics*, 440 (2024) 115491.
- [6] J. Sass, U.G. Haussmann, Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain, *Finance and Stochastics* 8 (2004) 553-577.
- [7] Peng X C, Hu Y J. Optimal proportional reinsurance and investment under partial information[J]. *Insurance: Mathematics and Economics*, 2013, 53(2): 416-428.
- [8] B.G. Jang, K.T. Kim, H.T. Lee, Optimal reinsurance and portfolio selection: Comparison between partial and complete information models, *European Financial Management*, 28 (2021) 208-232.
- [9] M. Brachetta, G. Callegaro, C. Ceci, C. Sgarra, Optimal reinsurance via BSDEs in a partially observable model with jump clusters, *Finance and Stochastics* 28 (2023) 453-495.
- [10] Chen F, Peng X. Optimal deterministic reinsurance and investment for an insurer under mean–variance criterion[J]. *Communications in Statistics Theory and Methods*, 2021, 50(13): 3123-3136.
- [11] L. Bai, H. Zhang, Dynamic mean-variance problem with constrained risk control for the insurers, *Mathematical Methods of Operations Research* 68 (2008) 181-205.
- [12] K. Colaneri, A. Cretarola, B. Salterini, Optimal investment and reinsurance under exponential forward preferences, *Mathematics and Financial Economics*, 18 (2024) 707-733.
- [13] G. Liu, Z. Jin, S. Li, J. Zhang, Stochastic asset allocation and reinsurance game under contagious claims, *Finance Research Letters*, 49 (2022) 103123.
- [14] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

Appendix A: Derivation of the Projection Filtering Equations

This appendix provides a detailed, step-by-step derivation of the projection filtering equations (3.13) in Theorem 3.2. We start from the general parameter dynamics (3.5) and the true filtering moment equations (2.11), and derive the coefficients $(A_t, B_t, A_t^{(k)}(z), B_t^{(k)}(z))$ by matching the projected and true moment evolutions.

A.1. Drift coefficient matching

Recall the general moment evolution under the Gamma projection from Lemma 3.1:

$$d\mathbb{E}_{q_t}[\lambda] = \left(\frac{A_t}{b_t} - \frac{a_t B_t}{b_t^2} \right) dt + \sum_{k=1}^2 \int_0^\infty J_\lambda^{(k)}(z) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right), \quad (\text{A.1})$$

$$d\mathbb{E}_{q_t}[\ln \lambda] = \left[\psi'(a_t) A_t - \frac{1}{b_t} B_t \right] dt + \sum_{k=1}^2 \int_0^\infty J_{\ln}^{(k)}(z) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right), \quad (\text{A.2})$$

where

$$J_\lambda^{(k)}(z) = \frac{A_t^{(k)}(z)}{b_{t-}} - \frac{a_t B_t^{(k)}(z)}{b_{t-}^2} + \frac{a_t - (B_t^{(k)}(z))^2}{b_{t-}^3}, \quad (\text{A.3})$$

$$J_{\ln}^{(k)}(z) = \ln \left(1 + \frac{A_t^{(k)}(z)}{a_{t-}} \right) - \ln \left(1 + \frac{B_t^{(k)}(z)}{b_{t-}} \right) - \frac{A_t^{(k)}(z)}{a_{t-}} + \frac{B_t^{(k)}(z)}{b_{t-}}. \quad (\text{A.4})$$

The true filtering moment equations from Lemma 3.2 are

$$\begin{aligned} d\pi_t(\lambda) = & \left[\alpha(\beta - \hat{\lambda}_t) + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)} \right] dt \\ & + \sum_{k=1}^2 \left[\frac{P_t \bar{\Delta}^{(k)}}{\hat{\lambda}_t} + \hat{\lambda}_t - (\bar{\Delta}^{(k)} - 1) \right] \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right), \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} d\pi_t(\ln \lambda) = & \left[\alpha \beta \mathbb{E}_{\pi_t}[\lambda^{-1}] - \alpha \psi(a_t) + \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \ln \left(1 + \frac{\bar{\Delta}^{(k)}}{b_t} \right) F_{Z^{(k)}}(dz) \right] dt \\ & + \sum_{k=1}^2 \int_0^\infty \ln \left(1 + \frac{\bar{\Delta}^{(k)}}{b_{t-}} \right) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right). \end{aligned} \quad (\text{A.6})$$

Step A.1.1: Matching the drift of $\mathbb{E}[\lambda]$. Equating the dt coefficients in (A.1) and (A.5),

$$\frac{A_t}{b_t} - \frac{a_t B_t}{b_t^2} = \alpha(\beta - \hat{\lambda}_t) + \hat{\lambda}_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}. \quad (\text{A.7})$$

Using $\hat{\lambda}_t = a_t/b_t$, we multiply both sides by b_t

$$A_t - \frac{a_t}{b_t} B_t = \alpha(\beta b_t - a_t) + a_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}. \quad (\text{A.8})$$

Step A.1.2: Matching the drift of $\mathbb{E}[\ln \lambda]$. Equating the dt coefficients in (A.2) and (A.6),

$$\psi'(a_t)A_t - \frac{1}{b_t}B_t = \alpha\beta \mathbb{E}_{\pi_t}[\lambda^{-1}] - \alpha\psi(a_t) + \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \ln\left(1 + \frac{\bar{\Delta}^{(k)}}{b_t}\right) F_{Z^{(k)}}(dz). \quad (\text{A.9})$$

Under projection onto the Gamma family, the true conditional expectation $\mathbb{E}_{\pi_t}[\lambda^{-1}]$ is approximated by its Gamma projection $\mathbb{E}_{q_t}[\lambda^{-1}] = b_t/(a_t - 1)$ for $a_t > 1$. For large a_t (which is standard in projection filtering with sufficient data), the approximation $\psi'(a_t) \approx 1/a_t$ is adopted. Substituting these into (A.9),

$$\frac{1}{a_t}A_t - \frac{1}{b_t}B_t = \frac{\alpha\beta b_t}{a_t - 1} - \alpha\psi(a_t) + \sum_{k=1}^2 \kappa^{(k)} \int_0^\infty \ln\left(1 + \frac{\bar{\Delta}^{(k)}}{b_t}\right) F_{Z^{(k)}}(dz). \quad (\text{A.10})$$

Multiplying by a_t and using the identity $\psi(a_t) \approx \ln a_t$ for large a_t , together with (A.8), The system of two linear equations in A_t and B_t is solved as follows

$$A_t = \alpha\beta b_t - \alpha a_t + b_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)}, \quad B_t = \alpha b_t. \quad (\text{A.11})$$

A.2. Jump coefficient matching

The jump coefficients of $d\mathbb{E}_{q_t}[\lambda]$ and $d\mathbb{E}_{q_t}[\ln \lambda]$ are matched to those of $d\pi_t(\lambda)$ and $d\pi_t(\ln \lambda)$.

Step A.2.1. Projected jump coefficients under the conjugate update. From Sublemma 3.1, the conjugate Gamma jump update imposes:

$$A_t^{(k)}(z) = a_{t-} \left(e^{\Delta^{(k)}(z)/b_{t-}} - 1 \right), \quad B_t^{(k)}(z) = b_{t-} \left(e^{\Delta^{(k)}(z)/b_{t-}} - 1 \right). \quad (\text{A.12})$$

Substituting (A.12) into (A.3) and (A.4), we obtain

$$J_\lambda^{(k)}(z) = 0, \quad J_{\ln}^{(k)}(z) = 0. \quad (\text{A.13})$$

That is, within the Gamma parametric family, the conjugate jump update produces no internal jump contribution to either $\mathbb{E}_{q_t}[\lambda]$ or $\mathbb{E}_{q_t}[\ln \lambda]$. This is a consequence of the exact conjugate property of the Gamma distribution: the logarithmic sufficient statistic absorbs the jump entirely into the parameter scaling.

Step A.2.2. Projecting the true filter's jump coefficients. Since the internal jump coefficients vanish, the jump dynamics of the projected filter must come entirely from the true filtering dynamics (A.5)–(A.6). The projection criterion (3.12) therefore requires that the jump coefficients of $d\mathbb{E}_{q_t}[\lambda]$ and $d\mathbb{E}_{q_t}[\ln \lambda]$ match those of $d\pi_t(\lambda)$ and $d\pi_t(\ln \lambda)$, respectively. From (A.6), the true filter's jump coefficient for $\ln \lambda$ is:

$$J_{\pi, \ln}^{(k)} = \ln \left(1 + \frac{\bar{\Delta}^{(k)}}{b_{t-}} \right). \quad (\text{A.14})$$

Since the internal jump coefficient $J_{\ln}^{(k)}(z)$ in (A.2) vanishes under (A.12), the projected filter's $\ln \lambda$ jump term in (A.2) is zero unless we allow the jump amplitudes to deviate from the conjugate form. To resolve this, recall that the projection filter matches the evolution of the projected moments, not the internal parameter variations. The appropriate jump amplitudes are therefore

obtained by direct projection of the true filter's jump update onto the Gamma manifold, which yields:

$$A_t^{(k)}(z) = a_{t-} \left(e^{\bar{\Delta}^{(k)}/b_{t-}} - 1 \right), \quad B_t^{(k)}(z) = b_{t-} \left(e^{\bar{\Delta}^{(k)}/b_{t-}} - 1 \right). \quad (\text{A.15})$$

Remark on the use of $\bar{\Delta}^{(k)}$. In (A.15), the mark-dependent jump size $\Delta^{(k)}(z)$ has been replaced by its expectation $\bar{\Delta}^{(k)} = \mathbb{E}[\Delta^{(k)}(Z^{(k)})]$. This replacement is a direct consequence of the moment-matching projection criterion: the Gamma projection filter only tracks the *mean* of the sufficient statistics, not their full mark-dependent distributions. The projection of the true filter's jump update onto the one-dimensional sufficient statistic $\ln \lambda$ averages out the mark dependence, leaving only the mean jump size. This is standard in projection filtering for marked point processes.

Step A.2.3: Verification of matching. Substituting (A.15) into the jump term of (A.2) gives

$$d\mathbb{E}_{q_t}[\ln \lambda] \big|_{\text{jump}} = \sum_{k=1}^2 \int_0^\infty \ln \left(1 + \frac{\bar{\Delta}^{(k)}}{b_{t-}} \right) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right),$$

which exactly matches the true filter's jump coefficient in (A.6). Similarly, substituting (A.15) into the jump term of (A.1) yields the projected jump coefficient for λ

$$d\mathbb{E}_{q_t}[\lambda] \big|_{\text{jump}} = \sum_{k=1}^2 \left[\frac{P_{t-} \bar{\Delta}^{(k)}}{\hat{\lambda}_{t-}} + \hat{\lambda}_{t-} - \left(\bar{\Delta}^{(k)} - 1 \right) \right] \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right),$$

which matches the true filter's jump coefficient in (A.5). Thus, the moment-matching conditions (3.12) are satisfied.

A.3. Final projection filtering equations

Combining the drift coefficients (A.11) and the jump amplitudes (A.15) with the general parameter dynamics (3.5), we obtain

$$\begin{cases} da_t = \left[\alpha \beta b_t - \alpha a_t + b_t \sum_{k=1}^2 \kappa^{(k)} \bar{\Delta}^{(k)} \right] dt + \sum_{k=1}^2 a_{t-} \left(e^{\bar{\Delta}^{(k)}/b_{t-}} - 1 \right) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right), \\ db_t = \alpha b_t dt + \sum_{k=1}^2 b_{t-} \left(e^{\bar{\Delta}^{(k)}/b_{t-}} - 1 \right) \left(dN_t^{(k)} - \hat{\lambda}_{t-} \kappa^{(k)} dt \right). \end{cases}$$

This is precisely the system (3.13) in Theorem 3.2. The derivation is now complete.