

Dynamic Optimization of Autocallable Product Trigger Mechanisms: A Distributionally Robust MDP Approach with Chinese Market Validation

Yuhang Jiang

Wuhan University, Wuhan, Hubei, 430072, China

ABSTRACT

Snowball products, also known as autocallable structured notes, have grown rapidly in China, but their fixed knock-in and knock-out thresholds often fail under market stress. This paper proposes a dynamic optimization framework that integrates distributionally robust optimization (DRO) with Markov decision processes (MDP) to adapt thresholds to market volatility, skewness, liquidity, and tail risk. It also develops a layered Delta-Gamma-Vega hedging strategy suitable for China's less liquid options market. Using Monte Carlo simulations and historical backtesting on Chinese market data from 2018 to 2024, the study compares fixed-threshold, rule-based dynamic, and DRO-MDP designs. The DRO-MDP approach reduces conditional value-at-risk (CVaR) by 28%, increases the Sharpe ratio by 55%, and keeps the additional hedging cost below 0.5 percentage points. The results provide practical implications for structured-product designers, risk managers, and regulators.

KEYWORDS

Autocallable products; Distributionally robust optimization; Markov decision processes; Dynamic threshold; Chinese financial market; Hedging.

1 Introduction

Snowball products have become an important segment of China's structured retail market. They pay high coupons as long as the underlying asset remains within a prespecified range, with discrete observation dates and path-dependent knock-in and knock-out barriers. By early 2022, the total notional outstanding reportedly exceeded 150 billion yuan, with many products linked to the CSI 500 index. This market scale has raised regulatory concerns about systemic risks from clustered knock-in events and the associated forced hedging.

Conventional designs employ fixed thresholds, such as an 80% knock-in threshold and a 100% knock-out threshold. Such static mechanisms cannot adapt to changing market conditions. In a market as volatile and policy-sensitive as China, fixed thresholds may cause premature knock-outs or deep knock-ins, harming investors and complicating hedge management. The 2020 COVID-19 shock and the 2022 downturn illustrate this problem. During early 2022, a wave of CSI 500-linked snowball products triggered knock-ins, forcing issuers to cut delta hedges and potentially amplifying downward pressure. This procyclical pattern underscores the need for more adaptive product designs.

Existing literature models autocallable products as discrete barrier options under normality and frictionless markets [1]. These models ignore path dependence, non-normal returns, and distributional uncertainty, all of which are particularly important in China because of frequent jumps, liquidity constraints, and policy interventions. Moreover, most existing models assume fixed thresholds and therefore do not exploit the possibility of adapting product terms to market states. Although distributionally robust optimization (DRO) [3], [4] and Markov decision processes (MDP) have been used separately in finance, their combination for sequential autocallable threshold design remains underexplored.

This paper fills the gap by proposing a unified DRO-MDP framework that dynamically adjusts knock-in and knock-out thresholds. The product is modeled as a discrete barrier option with a state space comprising implied volatility, skewness, liquidity, and tail risk. Threshold adjustment is formulated as a sequential decision problem under uncertainty and solved using a Wasserstein DRO-MDP hybrid that minimizes CVaR under the worst-case distribution. The paper also designs a layered Delta-Gamma-Vega hedging strategy tailored to China's limited options liquidity. The framework is validated through Monte Carlo simulations and historical backtesting on Chinese market data from 2018 to 2024.

The rest of the paper is organized as follows. Section II reviews the literature. Section III presents the theoretical framework and methodology. Section IV reports the empirical results. Section V discusses mechanisms, implications, and limitations. Section VI concludes.

2 Literature Review

This section synthesizes prior work on autocallable pricing, tail risk measurement, and robust optimization,

highlighting the research gap that motivates this study.

2.1 Autocallable Product Pricing and Modeling

Early studies treat autocallable products as discrete barrier options and apply Black-Scholes extensions, finite-difference methods, binomial trees, or Monte Carlo simulation under assumptions of geometric Brownian motion and fixed thresholds^[1]. These methods provide a useful baseline but do not fully capture path dependence or the effect of changing market conditions. Recent work by Cui et al.^[11] uses Markov chain approximation to better handle discrete observation and hedging, but it still focuses mainly on a pricing and hedging framework rather than dynamic product-term optimization. Many existing models are also calibrated to developed markets and pay limited attention to jump risk, volatility clustering, and liquidity constraints that are especially relevant in China.

2.2 Non-Normal Risk Measurement

Financial returns exhibit heavy tails, volatility clustering, and occasional jumps. To address these stylized facts, researchers have developed GARCH-type models and jump-diffusion processes. McNeil and Frey^[6] combine GARCH with extreme value theory (EVT) to estimate tail-related risk measures. Conditional value-at-risk (CVaR), also known as expected shortfall in related risk-management literature, has become a standard coherent risk measure^[2,5]. However, these approaches often rely on a fixed historical distribution and do not fully adapt to structural breaks or distributional shifts, which is a critical issue in emerging markets where policy changes can abruptly alter return dynamics.

2.3 Robust Optimization and Dynamic Decision Making

Distributionally robust optimization has gained traction as a tool for handling model uncertainty. Delage and Ye^[3] and Ben-Tal et al.^[4] develop frameworks that optimize over ambiguity sets defined by moment constraints or statistical distances. DRO has been applied to portfolio optimization and risk management, but its application to structured-product design remains limited. MDPs are natural for sequential decisions, yet standard MDPs typically assume known transition probabilities. Combining DRO with MDP provides a principled way to handle both model uncertainty and sequential decision making. To the best of our knowledge, this combined framework has not been systematically applied to autocallable threshold optimization.

2.4 Research Gap and Contribution

Despite progress in individual areas, no existing work fully integrates path-dependent autocallable modeling, non-normal tail-risk measurement, and distributionally robust dynamic optimization. Specifically, fixed thresholds remain the norm; tail-risk models often do not adjust to changing volatility or jump intensity; and robust optimization has rarely been applied to sequential knock-in and knock-out threshold design. This paper contributes a unified DRO-MDP framework, a layered hedging strategy that accounts for limited liquidity, and a validation exercise based on Chinese market data.

3 Methods

3.1 Product Modeling and State Space

Consider a snowball product on the ChinaAMC SSE 50 ETF (510050). The initial price is $S_0 = 100$. The maturity is $T = 12$ months, with monthly observation dates $t = 1, \dots, 12$. The knock-out threshold $B_u(t)$ and knock-in threshold $B_d(t)$ are time-varying. The annual coupon rate is $c = 15\%$. The payoff rules are as follows:

Knock-Out: If $S_t \geq B_u(t)$ at any observation date t , the product terminates, and the investor receives principal 100 plus accrued coupon $100 \times c \times t/12$.

Knock-In: If $S_t \leq B_d(t)$ at any observation date t , the product knocks in. If no later knock-out occurs, the investor loses $\max(0, 100 - S_T)$ at maturity.

Neither: If neither knock-in nor knock-out occurs by maturity, the investor receives principal plus the full coupon $100 \times c$.

To capture path dependence, the state vector is defined as

$$X_t = \left(\frac{S_t}{S_0}, \sigma_{\mathcal{N}}(t), \text{Skew}(t), \text{Liq}(t), \text{MDD}(t) \right) \in \mathcal{X} \subset \mathbb{R}^5 \quad (1)$$

Here, $\sigma_{\mathcal{N}}(t)$ is the implied volatility of the 50 ETF, $\text{Skew}(t)$ is return skewness, $\text{Liq}(t)$ is daily turnover as a liquidity proxy, and $\text{MDD}(t)$ is the historical maximum drawdown. The state space $\mathcal{X}$ is continuous and five-dimensional.

$$\text{MDD}(t) = \max_{0 \leq s \leq t} \frac{S_s - S_t}{S_s}$$

3.2 Non-Normal Risk Measurement

3.2.1 GARCH-Jump Process

To simulate realistic price dynamics, the log return $r_t = \ln(S_t/S_{t-1})$ follows a GARCH-jump process:

$$r_t = \mu + \sqrt{h_t} z_t + J_t q_t, \quad (2)$$

$$h_t = \omega + \alpha(r_{t-1} - \mu)^2 + \beta h_{t-1} \quad (3)$$

In this specification, $z_t \sim N(0,1)$ is i.i.d.; $q_t \sim \text{Bernoulli}(\lambda)$ with $\lambda = 0.05$; and $J_t \sim N(\mu_j, \sigma_j^2)$ with $\mu_j = -0.02$ and $\sigma_j = 0.01$. Calibration using daily returns from January 2018 to December 2024 gives $\mu = 0.0002$, $\omega = 1.2 \times 10^{-6}$, $\alpha = 0.12$, and $\beta = 0.85$. These parameter values are consistent with recent evidence on SSE 50 ETF return dynamics^[9].

3.2.2 Tail Risk: Dynamic CVaR

At the 95% confidence level, CVaR is defined as $\text{CVaR}_{0.95}(L) = E[L | L \geq \text{VaR}_{0.95}(L)]$ where L denotes loss. EVT is applied with a threshold at the 95th percentile, and excess losses are fitted to a generalized Pareto distribution (GPD)^[6]. To make risk tolerance market-dependent, the following dynamic CVaR constraint is imposed:

$$\text{CVaR}_{0.95}^{(t)}(X_t) \leq \theta_t, \theta_t = \theta_0 \left(1 - \frac{\sigma_N(t) - \bar{\sigma}}{\bar{\sigma}} \right) \quad (4)$$

Here, $\theta_0 = 20$, and $\bar{\sigma}$ is the historical average implied volatility. When current volatility exceeds $\bar{\sigma}$, θ_t decreases, which tightens the risk limit.

3.3 Dynamic Optimization: DRO + MDP

3.3.1 MDP Formulation

Product management is formulated as a finite-horizon MDP. The state is $s_t = (X_t, I_t)$, where $I_t \in \{0, 1\}$ indicates whether a knock-in has occurred, with 1 denoting yes. The process terminates upon knock-out. The action is $a_t = (B_u(t+1), B_d(t+1))$ with bounds $B_{d,\min} = 75$, $B_{d,\max} = 85$, $B_{u,\min} = 95$, and $B_{u,\max} = 105$, subject to $B_d(t+1) < 100 < B_u(t+1)$. Transition probabilities are approximated by Monte Carlo simulation with 10,000 trajectories. The immediate cost is

$$c(s_t, a_t) = \text{HedgeCost}_t(a_t) + \lambda_1 \text{CVaR}_{0.95}^{(t)} + \lambda_2 \|a_t - a_{t-1}\|^2 \quad (5)$$

where $\lambda_1 = 0.5$ and $\lambda_2 = 0.1$. The objective is to minimize the expected discounted cumulative cost:

$$\min_{\pi} \mathbb{E}^{\pi} \left[\sum_{t=0}^{N-1} \gamma^t c(s_t, a_t) \right], \gamma = 0.95 \quad (6)$$

3.3.2 Distributionally Robust Optimization

The true return distribution is unknown and may drift. Let $\hat{P}_N$ denote the empirical distribution. The ambiguity set is defined as $\mathcal{P}_{\varepsilon} = \{P: W_1(P, \hat{P}_N) \leq \varepsilon\}$, where W_1 is the 1-Wasserstein distance. The robust MDP problem becomes

$$\min_{\pi} \max_{P \in \mathcal{P}_{\varepsilon}} \mathbb{E}^P \left[\sum_{t=0}^{N-1} \gamma^t c(s_t, a_t) \right] \quad (7)$$

Using duality^[3,4], this problem can be reformulated as a finite-dimensional convex program. It is solved through scenario approximation with 10,000 scenarios and stochastic gradient descent. Cross-validation selects $\varepsilon = 0.1$.

3.3.3 Solution Algorithm

Because the state space is continuous, approximate dynamic programming (ADP) is used. A neural network with three hidden layers and 128 ReLU units in each layer approximates the value function $V(s)$. Training uses rolling simulations with 10,000 paths per iteration and Bellman residual minimization with the Adam optimizer at a learning rate of 10^{-3} . Convergence is achieved after 500 iterations. The deterministic policy is

$$a_t = \arg \min_a \left[c(s_t, a) + \gamma \mathbb{E}[V(s_{t+1})] \right] \quad (8)$$

3.4 Layered Hedging Strategy

The issuer hedges Delta, Gamma, and Vega:

Delta Hedging: Hold $\Delta_t = \partial V / \partial S$ units of at-the-money (ATM) 50 ETF options with one-month maturity. Delta is computed using the Black-Scholes model with $\sigma_N(t)$.

Gamma Hedging: Use options with different strikes to keep portfolio Gamma near zero.

Vega Hedging: Use variance swaps or option combinations to manage volatility exposure.

Because options liquidity in China is limited, the substitute hedging portfolio consists of 50 ETF options, CSI 300 index futures, and spot ETFs. Rebalancing is triggered when $|\Delta_t - \Delta_{t-1}| > 0.05$ or $|\sigma_N(t) - \sigma_N(t-1)| > 2\%$. Transaction costs are set at 0.05% for options and 0.02% for futures.

3.5 Experimental Design

3.5.1 Product Scenarios

Three product designs are compared:

Scenario I (Benchmark): Fixed thresholds $B_u \equiv 100$ and $B_d \equiv 80$, representing the standard 80-100 structure.

Scenario II (Rule-Based Dynamic): Monthly adjustments based on $\sigma_N(t)$ and $\text{Skew}(t)$:

If $\sigma_N(t)$ is above the 80th percentile, set $B_u(t+1) = 102$ and $B_d(t+1) = 78$.

If $\sigma_N(t)$ is below the 20th percentile, set $B_u(t+1) = 98$ and $B_d(t+1) = 82$.

Otherwise, keep 100/80.

Scenario III (DRO + MDP): The optimal dynamic policy from Section III.C.

3.5.2 Monte Carlo Simulation

The study generates 10,000 price paths at monthly steps using the calibrated GARCH-jump model. For each path and scenario, it computes the knock-out rate, knock-in rate, non-trigger rate, extreme-loss probability, mean and standard deviation of annualized returns, Sharpe ratio, hedging cost as a percentage of principal, maximum drawdown, VaR(95%), and CVaR(95%). Extreme loss is defined as a loss greater than 20% after knock-in. The risk-free rate is set at 3%. Results are reported as means with 95% bootstrap confidence intervals based on 1,000 resamples.

3.5.3 Historical Backtesting

Historical backtesting is conducted using daily data from January 2018 to December 2024 with 12-month rolling windows shifted monthly. Three sub-periods are examined: the COVID-19 shock from February to August 2020, the systematic downturn from January to October 2022, and the calm period from January to December 2019.

3.5.4 Sensitivity Analysis

Sensitivity analysis varies three key parameters: Wasserstein radius $\varepsilon \in \{0.05, 0.1, 0.15, 0.2\}$, jump intensity $\lambda \in \{0.02, 0.05, 0.10, 0.15\}$, and observation frequency (monthly, bi-weekly, weekly). The analysis tracks effects on CVaR, knock-out rate, and hedging cost, with two-sample t-tests used to assess statistical significance.

4 Results

4.1 Monte Carlo Simulation Results

Table 1 summarizes outcomes over 10,000 simulated paths. The knock-out, knock-in, and non-trigger rates sum to 100% for each scenario.

Relative to the fixed benchmark, the rule-based design lowers the knock-in rate by 23.4%, from 26.5% to 20.3%, and reduces the extreme-loss probability by 18.4% ($p < 0.01$). The DRO + MDP strategy achieves larger reductions: the knock-in rate falls by 34.0%, and the extreme-loss probability falls by 27.2%. The Sharpe ratio increases from 0.38 to 0.59, an improvement of 55.3%, while the mean annualized return increases by 1.21 percentage points. Hedging cost rises by only 0.4 percentage points. CVaR(95%) falls by 28.2% ($p < 0.01$). For a 1 million yuan investment, the expected loss in the worst 5% of scenarios is about 68,000 yuan lower under DRO than under the fixed design. The non-trigger rate increases from 5.3% to 20.2%, meaning that more paths avoid knock-in and earn the full coupon.

Table 1 Performance Comparison of the Three Product Scenarios (Mean $\pm$ 95% CI)

Metric	Scenario I (Fixed)	Scenario II (Rule-Based)	Scenario III (DRO + MDP)
Knock-Out Rate (%)	68.2 $\pm$ 0.9	64.8 $\pm$ 1.0	62.3 $\pm$ 1.1
Knock-In Rate (%)	26.5 $\pm$ 0.9	20.3 $\pm$ 0.8	17.5 $\pm$ 0.8
Non-Trigger Rate (%)	5.3 $\pm$ 0.4	14.9 $\pm$ 0.6	20.2 $\pm$ 0.7
Extreme-Loss Probability (%)	12.5 $\pm$ 0.7	10.2 $\pm$ 0.6	9.1 $\pm$ 0.6
Mean Annualized Return (%)	8.42 $\pm$ 0.32	9.15 $\pm$ 0.28	9.63 $\pm$ 0.25
Std. of Annualized Return (%)	14.3 $\pm$ 0.5	12.7 $\pm$ 0.4	11.2 $\pm$ 0.4
Sharpe Ratio	0.38 $\pm$ 0.03	0.48 $\pm$ 0.03	0.59 $\pm$ 0.03
Hedging Cost (%)	4.2 $\pm$ 0.2	4.4 $\pm$ 0.2	4.6 $\pm$ 0.2
Maximum Drawdown (%)	28.6 $\pm$ 1.2	22.3 $\pm$ 1.0	19.8 $\pm$ 0.9
VaR(95%) (%)	18.4 $\pm$ 0.6	15.2 $\pm$ 0.5	13.7 $\pm$ 0.5
CVaR(95%) (%)	24.1 $\pm$ 0.8	19.8 $\pm$ 0.7	17.3 $\pm$ 0.6

4.2 Historical Backtesting Results

Table 2 shows performance across market regimes.

During the COVID-19 shock, the fixed-threshold product loses 3.2% with a 32.1% drawdown. The rule-based product returns 1.8% with a drawdown of 23.5%, while the DRO + MDP product returns 3.5% with a drawdown of 20.8%. In the

2022 downturn, when many snowball products knocked in, the fixed-threshold product loses 8.5% with a drawdown of 37.6%. The rule-based product still loses 2.1%, but the DRO + MDP product achieves a positive 0.6% return with a drawdown of 24.5%. In the calm 2019 period, all three designs perform well, but DRO + MDP delivers the highest return, at 12.8%, and the lowest drawdown, at 8.9%. Differences between DRO + MDP and the fixed-threshold benchmark are statistically significant at the 5% level for the crisis periods.

Table 2 Historical Backtesting: Annualized Return / Maximum Drawdown by Sub-period

Sub-Period	Scenario I	Scenario II	Scenario III
COVID-19 Shock (2020.02-2020.08)	-3.2% / 32.1%	1.8% / 23.5%	3.5% / 20.8%
Systematic Downturn (2022.01-2022.10)	-8.5% / 37.6%	-2.1% / 28.9%	0.6% / 24.5%
Calm Period (2019.01-2019.12)	11.2% / 11.7%	12.5% / 9.8%	12.8% / 8.9%

4.3 Sensitivity Analysis

4.3.1 Wasserstein Radius

Increasing ε from 0.05 to 0.2 reduces CVaR from 18.5% to 16.1% but also lowers the knock-out rate from 64.5% to 60.8%. This trade-off reflects the conservatism introduced by a larger ambiguity set. The baseline value $\varepsilon = 0.1$ is selected as a balance between tail-risk reduction and product attractiveness.

4.3.2 Jump Intensity

As λ increases from 0.02 to 0.15, the advantage of DRO + MDP expands. The extreme-loss reduction relative to the fixed design rises from 23% to 41%. At the baseline $\lambda = 0.05$, the reduction is 27.2%. Thus, DRO provides the greatest benefit when jumps are frequent, which is precisely where fixed-threshold designs perform poorly. This finding is particularly relevant for China, where policy announcements can trigger sharp market movements.

4.3.3 Observation Frequency

Weekly observation increases hedging cost by about 1.5 percentage points and shrinks the Sharpe-ratio advantage of DRO + MDP to 0.12. Bi-weekly observation strikes the best balance, preserving most of the tail-risk reduction while keeping transaction costs moderate. Monthly observation is the cheapest but reacts too slowly to market changes, leading to higher CVaR during crisis periods.

5 Discussion

5.1 Mechanisms

Dynamic thresholds improve performance because they adapt to market conditions. When implied volatility spikes or skewness turns negative, the strategy raises the knock-out threshold to avoid premature redemption and lowers the knock-in threshold to reduce downside-trigger probability. This design weakens the procyclicality of fixed thresholds, under which falling markets may cause knock-ins, force issuers to reduce hedges, and further worsen market pressure.

The DRO component adds robustness. Instead of relying on a single estimated distribution, it considers a family of distributions within a Wasserstein ball. Under high jump risk, the DRO policy selects more conservative thresholds that perform well even under the worst-case distribution in that ambiguity set. This explains why DRO + MDP outperforms the rule-based dynamic threshold most clearly when jumps are frequent or markets are under severe stress, such as in 2022.

A simulated example based on the March 2020 drop illustrates the difference. The fixed-threshold product knocks in at 80 and does not recover, producing a 25% loss. The DRO product observes the volatility surge and proactively lowers the knock-in threshold to 78. If the underlying price remains above the adjusted threshold and later recovers, the product can avoid knock-in, knock out later, and pay the coupon.

5.2 Comparison with Prior Work

Compared with Cui et al. ^[1], who focus on pricing and hedging autocallable products using Markov chain approximation, this paper adds non-normal jump dynamics and dynamic threshold optimization. The resulting DRO + MDP design achieves a 28% reduction in CVaR relative to the fixed-threshold benchmark. The layered Delta-Gamma-Vega hedging strategy also refines conventional practice by actively managing Gamma and Vega exposures, which helps reduce rebalancing pressure during market stress. Methodologically, this paper is among the first to apply data-driven DRO to autocallable threshold optimization.

5.3 Practical Implications

Issuers: The rule-based dynamic threshold is straightforward to implement, while the DRO + MDP model offers greater sophistication for large exposures. Parameters such as $\varepsilon = 0.1$ and bi-weekly observation can be calibrated to an issuer's risk appetite. The dynamic CVaR constraint can also be automated to adjust thresholds in real time.

Investors: Fixed-threshold products may obscure tail risk. The DRO + MDP product's higher Sharpe ratio (0.59 vs. 0.38) and lower maximum drawdown (19.8% vs. 28.6%) provide quantitative support for dynamic designs. Investors should seek clear disclosure on how thresholds adjust under stress.

Regulators: The dynamic CVaR constraint can support forward-looking stress testing. Requiring issuers to report expected losses under different volatility and jump scenarios would improve transparency and systemic stability. Encouraging dynamic threshold mechanisms may also reduce the risk of clustered knock-ins.

5.4 Limitations and Future Work

Several limitations remain. First, credit risk is not modeled. Future work could incorporate credit valuation adjustment (CVA) and jointly model underlying-asset dynamics and issuer credit spreads. Second, liquidity is measured only by aggregate turnover. Product-level liquidity indicators, such as bid-ask spreads and order depth, may be more informative during stress. Third, computation is heavy and can take hours on a workstation. Deep reinforcement learning methods, such as proximal policy optimization (PPO) and soft actor-critic (SAC), may provide more scalable alternatives. Fourth, the empirical validation relies partly on simulated data. Proprietary data on actual redemptions and issuer hedging would allow more precise calibration.

Future research can extend the framework in several directions. Machine learning models, such as LSTM networks and random forests, can be used to forecast volatility and jumps, making thresholds more forward-looking. Multi-asset snowball products linked to the CSI 500 or STAR 50 can be studied with correlation and tail-dependence models. Regulatory constraints, including leverage limits and concentration caps, can be added to the optimization problem. Finally, alternative ambiguity sets, such as ϕ -divergence balls, can be compared with the Wasserstein approach.

6 Conclusion

This paper proposes a dynamic optimization framework for snowball products that integrates path-dependent modeling, non-normal risk measurement, and distributionally robust optimization with Markov decision processes. Using Monte Carlo simulation and backtesting on Chinese market data from 2018 to 2024, the analysis shows that the DRO + MDP approach substantially reduces extreme losses, improves risk-adjusted returns, and keeps additional hedging costs modest relative to fixed-threshold and rule-based designs. These findings hold across different market regimes, including the COVID-19 shock and the 2022 downturn. The proposed framework offers practical guidance for product designers, risk managers, and regulators. Future work should extend the framework to multi-asset settings, incorporate machine learning for state forecasting, and validate the model with proprietary product-level data.

References

- [1] Y. Cui, L. Li, and G. Zhang, "Pricing and hedging autocallable products by Markov chain approximation," *Review of Derivatives Research*, vol. 27, no. 3, pp. 259–303, 2024.
- [2] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.
- [3] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.
- [4] A. Ben-Tal, D. den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen, "Robust solutions of optimization problems affected by uncertain probabilities," *Management Science*, vol. 59, no. 2, pp. 341–357, 2013.
- [5] C. Acerbi, C. Nardio, and C. Sirtori, "Expected shortfall as a tool for financial risk management," *arXiv:cond-mat/0102304*, 2001.
- [6] A. J. McNeil and R. Frey, "Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach," *Journal of Empirical Finance*, vol. 7, no. 3–4, pp. 271–300, 2000.
- [7] D. P. Bertsekas, *Reinforcement Learning and Optimal Control*. Belmont, MA, USA: Athena Scientific, 2019.
- [8] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*. Boca Raton, FL, USA: Chapman & Hall/CRC, 2004.
- [9] M. Shi, C. Zhang, Q. Chen, and W. K. Härdle, "Modeling SSE 50 ETF returns and option pricing: Evidence from a score-driven GARCH-jump approach," *Mathematics*, vol. 13, no. 20, Art. no. 3332, 2025.