

## Project Proposal: Optimal Option Hedging and Pricing

---

**Authors.** Arif Ansari (aa4433@princeton.edu), Christos Avgerinos Tegopoulos (ct3125@princeton.edu), Jeremy Jun-Ping Bird (jb9895@princeton.edu),

**Project type:** 2. Applying RL to a new problem

**Introduction:** We aim to apply RL to the well-established financial problem of option pricing. Option pricing is underpinned by the idea that in complete and frictionless markets, one can perfectly replicate an option via a continuously rebalanced portfolio consisting of the underlying stock and a bond (the "replicating portfolio"). It follows from this that the price of an option will equal the cost of this replicating portfolio, subsequently making the problems of optimal hedging and option pricing synonymous.

The Black-Scholes-Merton's (BSM) model analytically solves for this replicating portfolio, effectively solving the problem of optimal option hedging and pricing. Yet option trading is still a multi-billion dollar business where traders are employed to (typically manually) price options and manage their risk. From this alone, it is clear that the traditional BSM approach to option hedging and pricing fails under realistic conditions faced outside of these "idealized" markets.

### Pitfalls of BSM:

- **Non-Continuous Reheding:** In reality, rebalancing of the replicating portfolio is done at a finite frequency (e.g. daily). Without continuous rebalancing, perfect replication is no longer feasible and the trader will be exposed to some degree of risk.
- **Transaction Costs:** Traders face transaction costs (e.g. brokerage, market impact) each time they rebalance. However, these costs are entirely neglected within the classical BSM model.
- **Stochastic Volatility:** Realistic markets exhibit non-stationary asset return volatility. BSM model, however, assumes constant volatility for stock price dynamics.

**RL Approach:** Determining optimal hedging practices requires one to balance minimizing transaction costs while trying to hedge as perfectly as possible (so as to minimize hedged portfolio PnL variance). This is an incredibly difficult problem to solve for analytically. This introduces the potential to apply RL to learn optimal policy (e.g. hedging strategy). Formally we have:

- **Actions  $a_t$ :** The amount of the underlying hedged at each timestep  $t$
- **States  $S_t$ :** The price of the underlying (or some time transformed variable  $X_t$  which normalizes for drift in the underlying dynamics)
- **Dynamics:** The underlying follows geometric Brownian motion with stochastic volatility (SABR model)
- **Reward Function  $R_t$ :** Here our reward at each period  $0 < t < T$  is equal to our change in wealth  $\Delta w_t = S_t(a_{t-1} - a_t) - \kappa|S_t(a_{t-1} - a_t)|$  minus the variance of  $\Delta w_t$  scaled by some risk-aversion parameter  $\lambda$

$$R_t = \Delta w_t - \lambda \mathbb{V}[\Delta w_t] = \Delta w_t - \lambda (\Delta w_t - \mathbb{E}[\Delta w_t])^2$$

Note that at  $t = 0$  the initial change in wealth  $\Delta w_0 = -S_0 a_0 - \kappa|S_0 a_0|$  as we must buy the initial replicating portfolio and at maturity we have  $\Delta w_T = S_T a_T - \kappa|S_T a_T| - G(S_T)$  where  $G(S_T)$  is the option payoff.

This gives final Bellman Equation

$$V^\pi(S_t) = \mathbb{E}_t^\pi[R_t(S_t, a_t, S_{t+1}) + \gamma V^\pi(S_{t+1})]$$