We study financial portfolio optimization in the presence of unknown and uncontrolled system variables referred to as stochastic factors. We propose FaLPO (factor learning portfolio optimization), a framework that interpolates between deep policy learning and continuous-time finance models. We explicitly formulate the dynamics of Assets prices $S$ with $\gamma = \gamma \exp(-\gamma)$ with $\gamma = \gamma$ and $\gamma \neq 1$; and the exponential utility $U(z; \gamma) := \frac{\gamma}{\gamma}(1 - \gamma)^{-1}$ with $\gamma = \gamma$ and $\gamma > 0$.

**Problem Definition**

Notations are provided below:

- Assets prices $S_i := [S_1^i, S_2^i, \ldots, S_n^i]^{\top}$ and factors $Y_i$.
- Risk-free return as zero.
- Terminal wealth under a policy: $Z_\gamma^t$.

Portfolio optimization aims to maximize the expected terminal utility: $E[U(Z_\gamma^T)]$, with two examples: the power utility $U(z; \gamma) := \frac{1}{1-\gamma} \gamma^{1-\gamma}$ with $\gamma = \gamma$, $\gamma > 0$, and $\gamma \neq 1$; and the exponential utility $U(z; \gamma) := \frac{\gamma}{\gamma}(1 - \gamma)^{-1}$ with $\gamma = \gamma$ and $\gamma > 0$.

**Background**

DDPG directly maximizes the following performance objective:

$$\max_{\theta_D} V(\theta_D) := \mathbb{E}[U(Z_\gamma^T; \theta_D)]$$

**Stochastic Factor Models** explicitly formulate the dynamics

$$\frac{dS_i^t}{S_i^t} = f_S^t(Y; \theta_S^t)dt + \sum_{j=1}^{dW} S_j^t dW_j^t,$$

$$\frac{dY_i}{Y_i} = f_Y^t(Y; \theta_Y^t)dt + g_Y^t(Y; \theta_Y^t) dW_i,$$

**Theorem 1** Define $V_\gamma^* := V(\pi^*)$, where $\pi^*$ is an optimal discrete-time admissible policy with time interval $\Delta t$, and $\theta_\gamma^* := (\theta_{\gamma, \Delta t}^*, \theta_{\gamma, \pi}^*, \theta_{\gamma, D}^*) \in \arg\max_{\theta_D, \theta_{\gamma, \pi}, \theta_{\gamma, D}} H(\theta_D, \theta_{\gamma, \pi}, \theta_{\gamma, D})$ with the policy functional form (3). With assumptions above,

$$\lim_{\Delta t \to 0} \left( V_\gamma^* - V(\theta_\gamma^*) \right) = 0.$$