Convertible Bond Pricing: TF Model and RL Approaches

Your Name Here

Thursday 6^{th} March, 2025

1 Introduction

This document outlines a project comparing:

- A PDE-based convertible bond pricing approach (Tsiveriotis-Fernandes, TF).
- An online Reinforcement Learning (RL) agent.
- A static (offline) RL or supervised learning approach.

We generate synthetic data, apply the TF model for bond pricing, and train RL agents (both online and offline) to see how closely they align with the PDE-based solution. Recent experiments produced two plots (`online_rl_comparison.png` and `static_rl_comparison.png`), which we analyze here.

2 Project Overview

2.1 Data Pipeline

(a) **Data Generation:** A script `data_generation.py` creates synthetic convertible bond parameters (bond price, IVOL, CDS, etc.) for thousands of samples.

(b) **Data Preparation:** `data_preparation.py` renames columns, preserves `ttm_days`, and stores the result in `prepared_data.xlsx`.

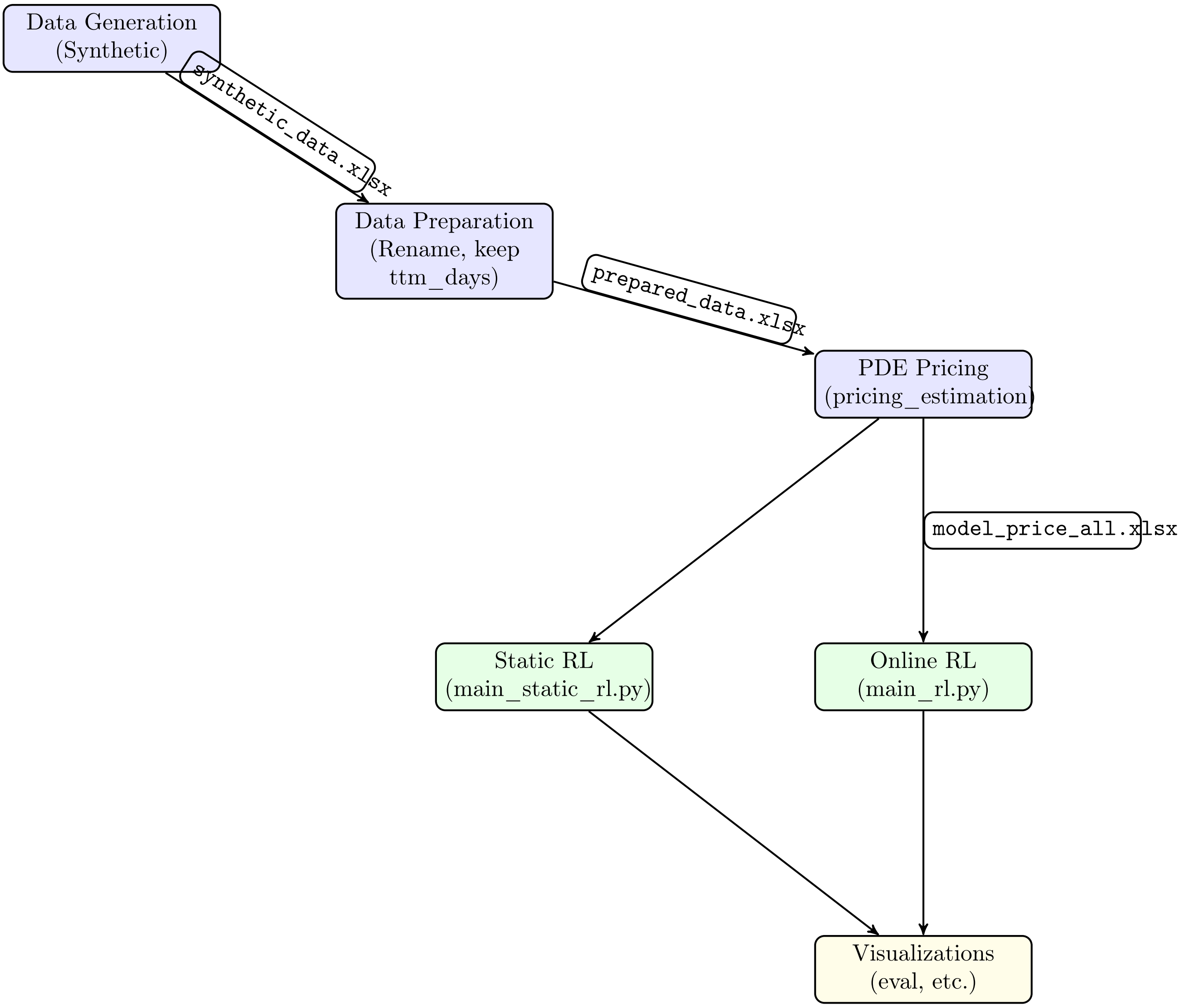
(c) **Pricing Estimation (TF Model):** `pricing_estimation.py` reads `prepared_data.xlsx`, applies a PDE solver (`explicit_FD`), and computes `Estimated_Price`, saving `model_price_all.xlsx`.

2.2 RL Approaches

(a) **Online RL (main_rl.py):** A `PolicyGradientAgent` interacts with `ConvertibleBondEnvTF`, which uses the PDE engine to compute a bond price at each time step. The agent learns via a REINFORCE-style algorithm whether to "hold" or "convert."

(b) **Static RL (main_static_rl.py):** A classifier is trained offline using data from `model_price_all.xlsx`. It labels each data point with 1 (convert) if the immediate conversion value > PDE-based price, else 0. The static RL approach imitates that policy.

3 Architecture Diagram



4 Key Observations

4.1 Online RL Plot

From Figure 1, we observe:

- **Stock Price Trajectory:** The price evolves from roughly 100 to around 110 with some fluctuations.
- **Sparse Conversion (Red X's):** The agent rarely converts, suggesting that the PDE price is typically higher than the immediate conversion value $\left(\frac{S}{K} \times \text{par}\right)$.
- **High PDE Price Near Maturity:** The PDE curve rises to tens of thousands by $t \approx 1$. This might be due to boundary condition choices, large $S_{\max}$, or coupon/spread misconfigurations in the PDE solver.

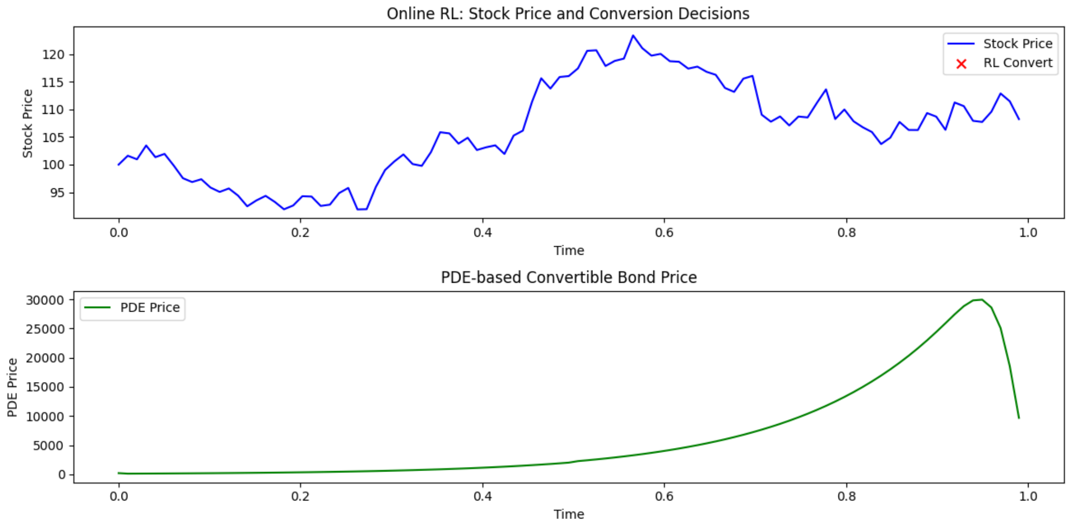


Figure 1: Online RL: Stock Price (top) and PDE-based Convertible Bond Price (bottom).

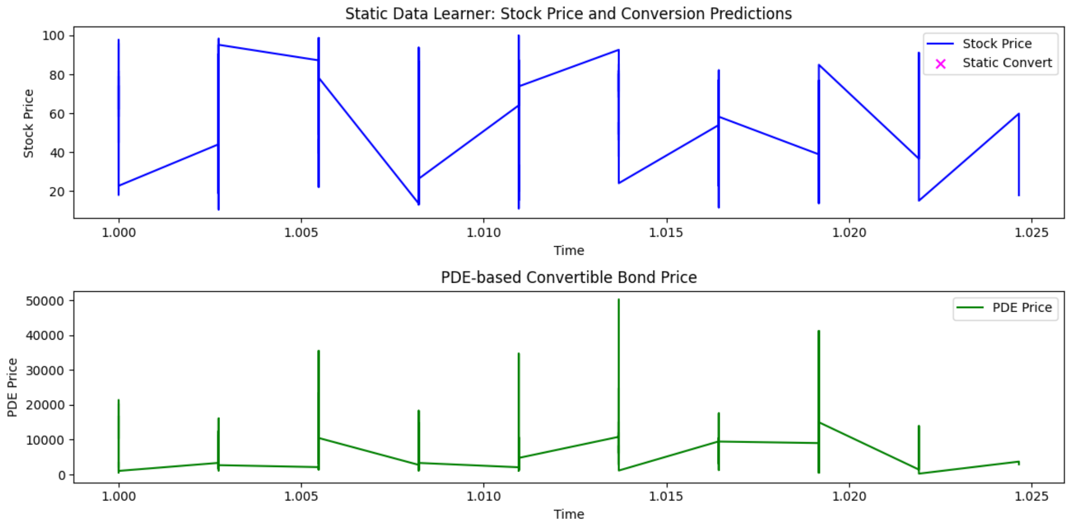


Figure 2: Static Data Learner: Stock Price (top) and PDE-based Price (bottom).

4.2 Static Data Learner Plot

From Figure 2, we observe:

- **Time Range Near Maturity:** The x-axis hovers around $t = 1.0$ to $t = 1.025$. The dataset used for plotting likely has times mostly at or near maturity.
- **Discrete Stock Price Jumps:** We see vertical lines, possibly because the data has only a few distinct times or is sorted in a non-chronological manner.
- **Frequent Convert Predictions (Pink X's):** Despite PDE values being large, the static learner is predicting convert often. This might indicate a labeling mismatch or that the PDE price is overshadowed by some data anomaly.
- **High PDE Values Again:** PDE in the thousands or tens of thousands suggests parameter/boundary condition issues, consistent with the online RL plot.

4.3 Bridging the Gap Between RL and TF

Both the online RL and the static RL approaches exhibit discrepancies from the PDE model (or each other) when PDE prices become unrealistically large. Potential ways to reduce the gap:

1. **Refine PDE Parameters:** Ensure interest rate, spread, coupon, and boundary conditions (like $S_{\max}$) are realistic so that the PDE solution yields more plausible bond prices.
2. **More Data Variation:**
   - *Time Variation:* Provide the static learner with data across a wide range of *ttm_days* (e.g., $t \in [0, 1]$).
   - *Market Cycles:* For online RL, simulate multiple regimes of stock volatility or drift so the agent sees more variety.
3. **Hybrid Approaches:** Combine PDE-based guidance (e.g., reward shaping) with RL exploration. The PDE model can act as a teacher or baseline to compare actions against.
4. **Check Labeling Logic (Static):** If PDE is $\gg$ conversion value, label should be 0 (hold). Validate that `label = (conversion_value > Estimated_Price).astype(int)` is correct for your data.

5 Usage Instructions

5.1 Installing Dependencies

```
pip install -r requirements.txt
```

5.2 Data Generation and Preparation

```
python src/data_generation.py
python src/data_preparation.py
```

5.3 TF Model Pricing

```
python src/pricing_estimation.py
```

Saves `model_price_all.xlsx` with an `Estimated_Price` column (and now includes `ttm_days`).

5.4 Online RL Training

```
python src/main_rl.py
```

Trains a policy network (REINFORCE) and saves `policy_model.pth` to `data/`.

5.5 Static RL Training

```
python src/main_static_rl.py
```

Trains a classifier on `model_price_all.xlsx`, labeling each row based on immediate conversion value vs. PDE price. Saves `static_policy_model.pth` to `data/`.

5.6 Evaluation and Plotting

```
python src/eval_rl.py
```

Generates the two plots shown in Figures 1 and 2, saved to `images/`.

6 Additional Notes

- **Coupon Logic:** If your convertible bond pays discrete coupons, ensure the PDE solver implements them accurately. Otherwise, PDE prices can inflate unexpectedly.
- **Environment vs. Static Data:** Online RL sees a time-evolving environment, while the static learner uses a snapshot-based dataset. If the dataset is narrow in time, it won't reflect early or mid-lifecycle decisions.
- **Future Extensions:**
  - Multi-factor models (stochastic interest rates, credit spreads).
  - Jump-diffusion processes for the stock price.
  - Real historical data integration to calibrate PDE parameters.

7 Conclusion

We have demonstrated how both online RL and static RL can diverge from the PDE model if certain parameters lead to excessively high bond prices. By refining PDE inputs, adding more diverse data, and possibly adopting hybrid RL–PDE reward shaping, one can bring the RL decisions closer to TF-based valuations. This pipeline provides a flexible foundation for exploring and comparing convertible bond pricing approaches.